SOUTH ATLANTIC FISHERY MANAGEMENT COUNCIL

SCIENTIFIC AND STATISTICAL COMMITTEE: MRIP WORKSHOP

Town and Country Inn
Charleston, South Carolina

August 19-21, 2019

<u>Summary Minutes</u>

<u>Scientific and Statistical Committee Members</u>

Dr. George Sedberry, Chair   Dr. Carolyn Belcher
Dr. Jeffrey Buckel   Dr. Scott Crosson
Dr. Chris Dumas   Dr. Churchill Grimes
Anne Lange   Dr. Yan Li
Dr. Genny Nesslage   Dr. Marcel Reichert
Dr. Frederick Scharf   Dr. Tracy Yandle

<u>Council Members</u>

Mel Bell   Steve Poland

<u>Council Staff</u>

Gregg Waugh   John Carmichael
Myra Brouwer   Julia Byrd
Cierra Graham   Dr. Mike Errigo
Christina Wiegand

<u>Other Observers and Participants</u>

Drew Cathey   Brad Floyd
Chris Wilson   Dr. Wilson Laney
Dave Dan Vorhees   Michael Larkins
Beverly Sauls   Shep Grimes
Dawn Franco   John Foster
Lauren Dolinger-Few   Rob Andrews
Vivian Matter   Ryan Rindone
Kyle Dettloff   Chris McDonough
Richard Cody   David Deltor
Kyle Shertzer   Jim Estes

Other observers and participants attached.

Attachment 2: SSC October 2019 Meeting

The Scientific and Statistical Committee of the South Atlantic Fishery Management Council convened an MRIP Workshop at the Town and Country Inn, Charleston, South Carolina, on Monday, August 19, 2019, and was called to order by Dr. George Sedberry.

INTRODUCTION

DR. SEDBERRY: Good afternoon, everybody. My name is George Sedberry, and I am Chairman of the South Atlantic SSC, and we'll have everybody introduce themselves shortly, but, first, I wanted to welcome everybody to the SSC MRIP Data Workshop. As you know, we have recently -- In our fall SSC meeting, we looked at some revised assessments, and we had some problems, the SSC had some problems, understanding what was done and coming to consensus on whether they were suitable for setting catch levels.

Since that time, we've had kind of an internal SSC webinar to go over the details of those assessments, and then, at the March South Atlantic Council meeting, the council directed staff and the SSC to form this workshop to provide more input to the SSC to help us better understand the data process and the data that go into the MRIP estimates, so that we can in fact make some decisions on assessments that are in the works, and future assessment as well, and so that's what we're here today to do, is to -- Or for the next couple of days, is to get more information from the folks involved in MRIP about how the survey is done and what data are collected and how they are post-processed within MRIP and within the Southeast Fisheries Science Center, so that we can better understand what is going on and move forward with assessments, some of which have been put on hold until we have a better understanding of how the MRIP data works, and so that's why we're here.

I wanted to make sure that everybody understands that this is an SSC meeting, but we have many invited guests, and we have some invited speakers, and we have state representatives that are involved with MRIP data collection, and I want to make sure that our invited guests are seated at the table, so that we get input from everybody here. We want to hear suggestions and clarifications and information from everybody here that has been invited, including SSC members, of course, and then please -- I invite everybody to participate in the discussions.

At some point, the SSC will have to make some recommendations and come to some conclusions and to make up -- Not make up. We're not making them up, but we're coming up with consensus statements about how to move forward, and those will be SSC decisions, but we really want our invited experts to help us to make those decisions, and so please take a seat at the table, if you're not already there.

Once we have everybody at the table, I would like to go around and have everyone introduce themselves, and that also brings up the point that this meeting is being webcast, it's a webinar, and it's also being recorded, and so be mindful of that and identify yourself, but we also want to, during this introductory time here, is to have everybody introduce themselves, so we can get your voice on the record, so we can kind of match up names and faces with voices, if we need to later, and so let's start in my far-left over here, and we'll go around the table. State your name and your affiliation, please. Thanks.

DR. CODY: Richard Cody, supporting the Office of Science and Technology.

MS. DOLINGER FEW: Lauren Dolinger Few, also from the Office of Science and Technology.

MR. ANDREWS: Rob Andrews, also with the MRIP Program and the Office of Science and Technology.

MR. DETLOR: Good afternoon. David Detlor, Acting Director of NOAA Fisheries Office of Science and Technology.

MS. LANGE: Anne Lange, SSC.

DR. LANEY: Wilson Laney, SSC, and I guess my affiliation, for SSC purposes, is North Carolina State University Department of Applied Ecology.

DR. SEDBERRY: Before we move on, I would like to welcome Wilson as a new member of the South Atlantic SSC, and so he will be joining us many times in the future, and so welcome, and we will enjoy working with you over the next few years.

DR. LANEY: We will see how many last, but thanks, George.

DR. SEDBERRY: Okay. Moving on.

DR. DUMAS: Chris Dumas, SSC, and I'm a natural resource economist at UNC-Wilmington.

DR. WILLIAMS: Erik Williams, Southeast Fisheries Science Center, Beaufort Lab.

DR. SHERTZER: Kyle Shertzer, Southeast Fisheries Science Center, Beaufort Lab.

DR. BELCHER: Carolyn Belcher, Georgia Department of Natural Resources.

DR. CROSSON: Scott Crosson, SSC member and also Southeast Fisheries Science Center.

DR. REICHERT: Marcel Reichert, South Carolina Department of Natural Resources, SSC member.

DR. SEDBERRY: George Sedberry, SSC Chair.

DR. ERRIGO: Mike Errigo, council staff, and, if I could, as far as the invited participants go, there should be at least one person from each of the states sitting at this table, and so, if you're not, you should have someone come up and sit at the table. Thank you. Sorry we weren't clear about that from the beginning. I was a little preoccupied fixing the agenda.

MR. FOSTER: John Foster, NOAA Fisheries, Office of Science and Technology.

DR. BUCKEL: Jeff Buckel, NC State University.

DR. SCHARF: Fred Scharf, UNC-W, SSC member.

DR. GRIMES: Churchill Grimes, SSC member.

DR. YANDLE: Tracy Yandle, SSC member, Emory University.

MS. MATTER: Vivian Matter, Southeast Fisheries Science Center, Miami Lab.

DR. VAN VOORHEES: Dave Van Voorhees, Chief of the Fisheries Statistics Division of the NOAA Office of Science and Technology.

MS. SAULS: Beverly Sauls, Florida Fish and Wildlife Conservation Commission.

DR. NESSLAGE: Genny Nesslage, Chesapeake Biological Lab and SSC member.

DR. LI: Yan Li, SSC member and NC Division of Marine Fisheries.

DR. CATHEY: Drew Cathey , North Carolina Division of Marine Fisheries.

MR. WILSON: Chris Wilson, North Carolina Division of Marine Fisheries.

MR. FLOYD: Brad Floyd, South Carolina Department of Natural Resources.

DR. SEDBERRY: Thanks, everybody. I think we also have a couple of SSC members that are on the webinar, and I don't know if they can introduce themselves, if they need to, but I think we have Luiz Barbieri, Fred Serchuk, and who else is online? Do you know?

DR. ERRIGO: Alexei is online.

DR. SEDBERRY: Rob Ahrens, the SSC Vice Chair, is also on the webinar.

DR. BARBIERI: Hello everybody. This is Luiz Barbieri, Florida Fish and Wildlife.

DR. AHRENS: Rob Ahrens, SSC Vice Chair, University of Florida.

DR. ERRIGO: Excellent. I believe that is all of our SSC members online who are not here.

DR. SEDBERRY: Okay, and Dawn has joined us at the table since we went around, and so if you could please introduce yourself for the record, please.

MS. FRANCO: Thanks, George. Dawn Franco, Georgia DNR.

DR. SEDBERRY: Welcome. Okay. At our April SSC meeting, a steering committee was formed to help organize this workshop, and that steering committee consisted of Fred Scharf, Chris Dumas, Luiz Barbieri, Yan Li, and myself, and Mike Errigo and John Carmichael, council staff, helped us to organize this, and so we appreciate everybody's effort that went into getting this going, and that steering committee came up with terms of reference for this data workshop.

If you look at the agenda, it's kind of organized around those terms of reference, and so, today, we're going to have some overview presentations of the MRIP process and procedures, and then,

starting tomorrow morning, we'll address the TORs individually. I would like to, as we address these agenda items, assign South Atlantic SSC members to take notes and write up the report sections for each of those agenda items, and what I'm going to do is just start at the top of my list of SSC members and assign four to each agenda item. We will do that shortly.

Usually, we have action items at the end of each of our presentations, and we divvy up those action items to SSC members, but we don't have action items for this, and so I'm just going to take the agenda topic items and assign them to folks as we go along. I didn't have time to do this in advance, like I usually do, and so we'll just do it on the fly here.

I should also mention that the Vice Chair, Rob Ahrens, is on the webinar, but he's unable to attend, and so Marcel Reichert, again, who is the Past Chair, and thought he was done with all of this kind of stuff, has been recruited, or volunteered, actually, to help me kind of stay organized and help take notes and keep things on track, and so I appreciate that, Marcel.

DR. REICHERT: You're welcome.

DR. SEDBERRY: I believe we are at the public comment period now.

DR. REICHERT: What about approval of the agenda?

DR. SEDBERRY: Is it public comment or approval of the agenda? Let me look.

DR. ERRIGO: First is Introductions and then Approval of the Agenda.

APPROVAL OF AGENDA

DR. SEDBERRY: Yes, you're right. Sorry. We do need to approve the agenda, and there have been some changes to the agenda since the version that's in the briefing book. We have some changes in speakers, and then we have, I think, an inserted agenda item that -- Mike could you address those two items, or I can do it.

One of them is an additional presentation by Kyle Shertzer, who is going to talk about I think it's TOR 1A, applied to black sea bass, and it's a presentation that's going to be tomorrow morning, and then the presentation, also dealing with 1A, on Wednesday is going to be given by Jay Breidt via webinar, and he's also at Colorado State University. Were there any other changes to the agenda, Mike?

DR. ERRIGO: Yes, and let me just briefly go over it. There has been a few changes to the attachment numbering, and I now have all the presentations together, labeled with a "P", and all the documents are labeled with an "A", for attachment, and so the numbering is a little different, and I am getting updated presentations and things, and I will get them out to you as soon as I can, and I'm about to send around the updated agenda and overview to everyone, and we'll get that up as soon as we can.

DR. REICHERT: The overview is only different because the agenda is different, or were there additions to the overview?

DR. ERRIGO: It's one document. Because there wasn't a ton of different bullet points here that needed fleshing out, and we had the terms of reference to do that, and so it's basically a single document that has the agenda and overview in it.

DR. SEDBERRY: Okay, and so, given those changes to the agenda, we need to approve the agenda. Is there anyone on the SSC that has any objections to the revised agenda, as it's just been outlined, and you will receive an email version of it shortly? Any objections to changes to the agenda? The agenda is approved.

Now we can move on to Public Comment. Usually, at our SSC meetings, we have public comment at the beginning and the end, and then after each presentation, but, for this kind of special session of the SSC, we're just going to have two public comment periods, and we had written public comments, and I don't believe any were received, but we would like to take public comment now, at the beginning, and then on Wednesday afternoon as well, and so, if there is any members of the public that would like to comment at this time, please come forward and find a mic.

PUBLIC COMMENT

MR. ESTES: Thank you. My name is Jim Estes, and I'm with the Florida Fish and Wildlife Conservation Commission, and I work in the Division of Marine Fisheries Management, and, along with the council, the State of Florida, my group, uses this information to manage our fisheries, and so I want to thank you for having this workshop, and I want to thank the MRIP folks for making, I think, widely-recognized changes in how we do our effort survey and our APAIS. I think most of us recognize that those are important.

I have been around doing fisheries work, mostly freshwater, but marine for the last seven or eight years, for thirty years, thirty-three years, actually, and I'm known as a skeptic, and I'm sorry, but I am a little bit skeptical, and I think that's a good thing, because that's what science is about. Science is about debate, and I hope to hear some debate here about this.

I wrote one of the letters that I think you have in your packet, questioning not the work that's been done, not the need for the work, but perhaps the accuracy of the results, and that's what we question, is the accuracy of the results, because we have to, frankly, use the results, and so the way that we look at it -- I don't have any baseline for knowing how much fishing effort there is on private boats. I have no clue about that, but I do know that I drive along the beach, and I fish, and it seems to me like some of the estimates that we have, that we have gotten out of our shoreline mode, seem, to me, incredible. Perhaps they are right, but, again, I'm a skeptic.

If you will indulge me for just a minute, I want to tell you a quick, short story. Every year, I go to 10,000 Islands and fish for snook, and, about three weeks ago -- We kind of changed where we were staying at, and, about three weeks ago, I went down there, and we stayed in Marco Island, and, in Marco, if any of you are familiar with Marco, there is lots of canals, and there is lots of residences, and there's lots of people all packed into a little space, and, unfortunately, we had to stay at a place that was in the very backend of a series of canals, and so it took us about forty minutes each morning to ride slowly through the canals.

Well, because of this issue, and because I was wondering if I really should be skeptical about some of these things, I thought, and I think you'll see a presentation later about this, that we have all these people packed together, and they have lots of docks, and maybe people are fishing there. Well, I can tell you, in eight days of fishing, and we paid attention every single day, I saw two kids three times, and that's all I saw fishing in those areas, and so I look forward to listening to what you have to -- I would just challenge you to be a little bit skeptical in your mind and try to find ways, if there needs to be, if it's needed, try to find ways to improve this or try to find some independent methods to show me, the skeptic, and other folks that might be skeptical and to give our fishermen, our public, some confidence that what we're doing is accurate. Thank you.

DR. SEDBERRY: Thank you, and the letters that you mentioned are -- The council solicited from the state agencies letters outlining their concerns with the MRIP data, and those were included in our briefing books. They are not public comment though. They were solicited by the council, and I just wanted to clarify that point, but thank you for your public comment.

DR. CROSSON: I had a question for you, and I guess I should have -- This kind of came to mind when I was looking at the briefing book or the slide shows that were sent. Because one of the big factors in these new MRIP numbers are large increases in state-managed fisheries, like drum and speckled seatrout and things like that, we may end up going into discussions with -- I mean, it seems to me that the state agencies are going to be more drastically affected by some of these numbers than even the South Atlantic Council, and so are we going to be taking care not to go too far down that, or what is our thoughts on that? I guess we'll just see how it comes?

DR. SEDBERRY: Well, I hadn't thought about that until right now. I think my thought is we will see how it goes, but we are dealing here with -- The reason we're here is because of the South Atlantic Fishery Management Council. At their March meeting, they passed a motion requesting this workshop and that the SSC participate to deal with South Atlantic Council issues, and so that doesn't mean we can't consider those other things, but we really need to focus on what the South Atlantic Council has requested that we do. Any other questions or comments before we move on? Any additional public comments? I don't see any, and so we will close the public comment period for now, and we will open up to public comments again on Wednesday. The first agenda item is the Introductory Presentations from MRIP, and we have several of those this afternoon, and I think all are being presented by John Foster. No?

DR. ERRIGO: We will jump back and forth between John Foster and Rob Andrews.

DR. SEDBERRY: Okay. These are background information, so that we can begin to address the terms of reference tomorrow, but these should be informative to help us -- To give us some background information so that we can address those terms of reference.

As I mentioned earlier, I would like the SSC members to participate in taking notes and preparing the report from this workshop, and I need to give a report to the South Atlantic Council at their September meeting, and so it would be nice to have a clear, concise summary of what happened and what we learned, or what we concluded, any consensus statements that the SSC makes regarding moving forward.

As I said, I'm just going to go down the list here and assign people for this first agenda item, for these introductory presentations this afternoon, and so I have the first four people on the list here

as Rob Ahrens, Scott Crosson, Luiz Barbieri, and Carolyn Belcher. If you all could work together to take notes. Again, Mike will keep his usual bullets, but we need to make sure that we summarize all the discussions, and particularly any kind of action items or consensus statements or recommendations that the SSC comes up with as we go through the agenda.

With that, are we ready for the first presentation? These presentations are in the briefing book, under the presentations tab, and, if there's any updates, Mike has emailed them to us. We also have, in the briefing book, five documents that go along with this afternoon's presentations, and those are Attachments 1 through 5, and I don't believe the numbering has changed on those attachments. That was just the Attachments 6 and after.

DR. ERRIGO: The numbering has not changed on those attachments. However, the order of the presentations has changed slightly. John Foster talked to me and said that he thought that it would be a better flow of things if they took the second presentation, the calibration one, and moved it to be fourth in line, and so after the next two after it, and so we're going to just shift that one, and, other than that, everything else should stay in the same order.

DR. SEDBERRY: Go ahead, John. Thanks.

INTRODUCTORY PRESENTATIONS FROM MRIP

MR. FOSTER: First, thanks very much for the opportunity to come and share this information. We tried to compile quite a bit of information into these five presentations, and so hopefully it will all be useful. Some of it may be a little too in-depth, some of it not as much as folks would like, but we'll certainly, of course, take any questions as we go along.

The first one I will be giving on, again, sort of a detailed overview of MRIP design and estimation procedures, and then we'll shift gears to Rob Andrews, who will walk through the FES development, Fishing Effort Survey development, the mail survey, and then the detailed presentation comparing the new mail survey to the previous Coastal Household Telephone Survey, and Rob is the lead in our office for the Fishing Effort Survey, as he was previously for the Coastal Household Telephone Survey. Then we'll switch gears, and I will present on the calibration methods for the Access Point Intercept Survey as well as the transition to the Fishing Effort Survey, and then we'll have a fifth presentation sort of going in-depth on the changes in the catch estimate series for the five species that were identified as being of interest, and so, again, that's a lot that we'll try to get through, but, again, taking questions as we go along.

I will start now with the design and estimation presentation, and so I want to give a little bit of background, and I will try to get through this pretty quickly, on sort of the overall transition from MRFSS, the previous Marine Recreational Fisheries Statistics Survey, to MRIP, and then I'll go through the APAIS design, the FES design, and then the catch and effort estimation procedures.

I will back up just for a second. Because the big change here is with the Fishing Effort Survey, and that provides the effort estimates for private boats and shore modes, I won't specifically be touching on the estimation for charter mode, charter boat mode, or headboats, and headboats really are covered by the separate headboat program in the Southeast anyway, but, if there are questions on those, I am happy to take them as well.

Starting with some background information, originally, we had MRFSS, or the Marine Recreational Fisheries Statistics Survey, for us, starting in roughly 1979, and it ran through the end of 2006. For most of its time, it had two component surveys, the intercept survey, very similar to what we now call the APAIS, or Access Point Angler Intercept Survey, interviewing anglers as they finished their fishing trips at access points to collect information on catch rates and other effort characteristics, and then we had the Coastal Household Telephone Survey, a random-digit-dial telephone survey of residential households that provided the base effort estimates, and, for most of the time, it provided those effort estimates for private boat, shore mode, and charter boat as well.

In 2006, a review by the National Research Council identified a number of serious problems with the MRFSS program that created the potential, the serious potential, for biases, primarily with the intercept survey, but there were also serious concerns with the Coastal Household Telephone Survey as well. Specifically, for the intercept survey, the survey design was complex, and I will get into that a little bit when we talk more about APAIS, but it was complex, but the estimation procedures made a number of simplifying assumptions, assumptions which could introduce bias, either directional bias or just additional variable error.

In addition, the MRFSS intercept survey focused on the most productive times of day, limiting temporal coverage, which could also introduce bias, to the extent that trips returning at off-peak times might differ in important ways from those returning during the peak hours. Then, finally, it was only conducted at a subset of all fishing access sites, those generally publicly available. Now, that is still present in the current design. It's just not feasibly logistically possible to sample everywhere someone might take a fishing trip, primarily due to the private access issue.

With the Coastal Household Telephone Survey, again, for its entire time period, it was generally an inefficient design. Frankly, a low proportion of the public fish, take saltwater fishing trips. We estimate now that it's somewhere generally in the neighborhood of 5 percent of the population generally fishes in any given two-month wave. It's higher in some states, like Florida for instance, and lower in others. Just to randomly make phone calls to households is a fairly inefficient way to try to find those anglers.

Additionally, in more recent years, since about 2000, it had increasing -- It had accelerating decreasing coverage, as people abandoned landline phones for cellphones, which were outside of the Coastal Household Telephone Survey frame. It only covered households with landline service. Then, sort of as kind of the third major issue, not just affecting the Coastal Household Telephone Survey, but, in general, telephone surveys more broadly, it had seriously declining response rates, with less than 10 percent response in the final years of the survey.

For all of those reasons, and others, in 2007, or really beginning in 2006, the Marine Recreational Information Program, or MRIP, was established, first to evaluate the current state of the methods that were in play, to begin to develop, through a number of pilot tests and projects, superior methods that could be used to replace what was currently done, have those tested and peer reviewed and certified, and then, finally, implement them. After roughly a decade or more, we have implemented both a revised intercept survey as well as the new Fishing Effort mail survey.

Specifically, in 2013, after several years of development of the new intercept survey design, as well as the pilot test, which was conducted for a year in North Carolina, in 2013, we implemented the new intercept survey, the APAIS, in all states, where it is conducted, and it ended the MRFSS intercept survey. Then, in 2018, we implemented the Fishing Effort Survey formally in all states, where it's conducted, and it ended the Coastal Household Telephone Survey. It ended at the end of 2017.

Both surveys have a number of improvements compared to their predecessors. They have greater coverage. For example, the Fishing Effort Survey, and Rob will talk more about this, covers the entire in-state resident population for coastal states, whereas the Coastal Household Telephone Survey just covered the coastal counties, and a large expansion factor was used to account for all of the remaining effort.

We now have appropriate complex estimation methods that account for or match to the complex sampling designs, using essentially weighted estimation, and I will talk in some detail about how those sample weights are calculated and how they are used in estimation, and these address the causes, or the potential, for bias in the old design.

Both of the new surveys, as well as the rest of the MRIP program, was, again, reviewed. What had been the National Research Council, in the interim, had become the National Academy. This review was done in 2016, with a report published early in 2017, and the changes that had been made were reviewed very favorably in this peer review for both the APAIS as well as the Fishing Effort Survey.

That is the background, and now I will go into detail on the current Access Point Angler Intercept Survey design, or APAIS. Again, just starting at kind of a general high level, these are in-person interviews conducted at fishing access sites at the end, or conclusion, of fishing trips, whether they are in boat modes or shore modes, and, again, these are publicly-accessible sites.

The survey is conducted continuously, year-round, in one-month -- Sampling is stratified by month, but we produce estimates in two-month waves, combining two months' worth of data in a weighted estimation, and it's done year-round in the Gulf states and in Florida, on the Atlantic coast, and in North Carolina. In the other states, it's done for five waves, excluding just January and February. Again, we collect primarily catch information from the intercept survey, but some other trip effort characteristics that I will speak to a little bit later as well, and these are used in calculating our standard two-month wave level estimates.

The design is complex. It's both stratified and clustered, and it has a multistage design, as well as an equal selection probability, and I will sort of walk through all of these different pieces. Again, the stratification allows us to essentially focus on different sectors of the overall recreational fishing, with the ability to target sample to those specific sectors.

The multistage cluster design essentially within each of the different stratification dimensions, that I will speak to in more detail in just a minute, the angler trips that we sample are clustered in various ways, and so, within the primary stage units, which are sites on specific dates during specific six-hour time windows, those represent sort of the highest level cluster of angler trips that return to that site within that window of time.

Within that cluster, in a second stage, you have, in the boat modes, fishing groups, or fishing parties, fishing on the same boat, and so those anglers are clustered within individual boat trips, and then you can have even additional clustering within an individual angler trip, when we get down to the catch level, and sub-sampling can occur at each of those levels, which all has to be accounted for in how the sample weights are calculated.

Just a bit on how we calculate the frame, or construct the frame, for APAIS. Again, it's a list of public -- It starts with a list of public access fishing sites, and these lists are maintained by our state and regional data collection partners, essentially the resource management state agencies that conduct the actual field sampling component of APAIS, maintain the lists, add sites, retire sites, update the detailed information on fishing activity at the sites, and so this is the beginning sort of point for the frame. The sites then can be clustered together, and we can have one or two-site clusters, and then those clusters are crossed with a calendar of days to make the frame for a specific month, as well as six-hour time intervals, and I will speak more about this in just a slide or two.

Backing up to stratification then, it has a geographic dimension, and so sampling, of course, is stratified by state and then, within states, particularly the larger states, there can be sub-state stratification, and so, in this example, Florida, which is stratified by coast, and then, within each coast, it has additional stratification, but, also, North Carolina, for example, has three sub-state regions within the overall state, and so sampling is formally stratified in these cases within a state, such that there are separate sample allocations within these sub-regions as well as independent sample selections made within the sub-regions.

Additionally, there is a temporal dimension of stratification, and so I mentioned that each wave is comprised of two months, and so sampling is stratified by month within each wave. Within month, it's further stratified by kind of day, and so weekends are separate from weekdays. Weekdays we define as Monday through Thursday, and weekends are Friday through Sunday. In addition, federal holidays are considered to be a weekend day type, and so weekdays versus weekends have, again, separate sample allocations and sample selections.

Then, within day, the twenty-four-hour day is stratified into a number of -- It's four principal six-hour time intervals, and then there is a fifth interval that you see there, and the four primary ones are A, B, C, and D. There are two nighttime intervals and two daytime intervals, and then there is a fifth interval, P, which stands for peak, which is created to help us in a way that can be accounted for in the sample weighting. It helps us improve the productivity of the intercept survey by targeting, or allocating, some sample to the sort of peak activity window of time, but, again, it's being done in a way that can be accounted for in the sample weighting process, so that the estimates are still representative and not biased towards a certain window of time, as they had been previously in the old MRFSS intercept design.

Another dimension of stratification in the new design is site group, and so this is used to help us target sample to specific modes of fishing. Under the old design, fishing mode was a formal stratification variable, and so, when an assignment was conducted, there was a specific mode of fishing that could be intercepted on that assignment, and there were some procedures to allow for alternate mode interviewing, but, in principle, an assignment had a single mode that interviews could be obtained in.

With the new design, that was no longer efficient. We had to come up with a better way to make the assignments more productive, and so, now, we formally separate the sites based on whatever the primary mode of fishing activity is at the site into site groups, and a sample can still then be allocated to target specific modes of fishing, but, when assignments are now conducted, intercepts can be obtained in any mode that is present at the site, which helps, again, improve the productivity in the new design, but doing so in a way, again, that can be accounted for with sample weighting and removing the potential for bias from oversampling one mode or another.

Specifically now, at a given site, intercepts could be obtained in shore mode, private boat mode, charter boat mode, if it's present, and it just depends on what is actually going on at the site on the given day of sampling.

A little bit of specific information for North Carolina. In North Carolina, the shore mode historically has been sort of partitioned into two separate sub-modes, or more detailed modes, beach/bank, sort of natural shoreline, and manmade structure, such as fishing piers, jetties, bulkheads, things like that.

Under the previous MRFSS intercept, they were sampled separately, with some allowances for mixing in specific conditions, and separate estimates were produced for beach/bank separate from manmade. In the new design, there is a single shore mode set of estimates, but the split between beach/bank and manmade is still maintained in the sampling and in terms of how the sample weights are calculated, such that, when those data are then combined to come up with the shore mode estimates, they have been weighted appropriate to reflect the different selection probabilities between beach/bank and manmade as well as the different activity levels that are at the sites on any given day of sampling. This split, while it's not present in the estimates that are produced on our website or in our estimate datasets, the split at the sampling level for the intercept survey is still present.

Now a little more detail on the different stages. Again, it's a multistage design. At the highest stage, the primary stage, is what we actually draw in our sample selection procedure, and it is a combination of sites and times, and so it's a site cluster, which, again, can be a one or two-site cluster or group, a specific day of the month, and a six-hour time interval. Those dimensions define the primary stage unit.

Within that cluster then, the next stage down is the amount of time that the sampler is actually present sampling at each site, or at the site, and so, if, for example, the sampler is present, if it's a one-site cluster and sampler is present, sampling for all six hours, then there is no sub-sampling at this stage. They have, in a sense, censused it, but, if it's, for example, a two-site cluster, they can only spend up to half the time at each site, and so there is a sub-sampling of time for each of the sites.

Nested within that then, returning to a site, are angler trips, and these are -- You have a total number of trips that return to the site. Some, or all of those, may be sampled, and so there could be a sub-sampling at this stage, if the sampler is -- If it's a busy site, they are not able to interview everyone, and so there is a sub-sampling there, and then, specific to the boat modes, a sampler may not be able to interview everyone on a single boat trip, and so there can be a sub-sampling within the boat at this stage as well. Then, again, because samplers don't have the time, necessarily, to observe or

measure all of the catch, there can be a sub-sampling within the catch for a specific angler trip as well, for things like lengths and weights.

Just a bit more on the site clustering, and I have spoken a little bit to this. Again, sites can either stand alone as a single-site cluster, and that's all of the high-pressure sites, or they could be clustered with a second site, if they're lower pressure, meaning they have considerably less fishing activity, and this -- The clustering is not fixed, and so, again, because the activity levels at sites can vary seasonally, and they can vary temporally within the day, morning versus peak versus afternoon, and they can vary by fishing mode, and it might be a high-activity private boat site, but very little charter effort, for example, or vice versa, and the clusters are done within the individual strata units. Two sites, if they are clustered together, they may not be always clustered together, and it depends on the specific fishing activity and what that is for an individual stratum.

There is a couple of other details. If two sites are clustered together, they can't be more than sixty minutes apart, driving time. Sites can only be clustered with other sites within the same county, for example, and, again, only sites with lower activity levels can be clustered. High-pressure sites will always sort of stand alone as an individual site cluster.

Some on the sample selection, and this is probably an area where folks might like to have a little more detail, and I've got one more slide on this. I will speak to it, but then I'm happy, again, to take questions, if folks would like more detail, and so the sample selection process uses unequal selection probabilities in a process known as probability proportional size, and so high-activity sites have a higher probability of being selected, and low-activity sites have a lower probability of being selected.

The table here sort of gives you kind of the categorical scale that we use to give the size measure to the sites and then sort of in kind of a ranking of what those sizes equate to, or how the numeric values come to, and so, on the left, you have sort of what is labeled as an expected number of angler trips to be returning to the site during the window of sampling, and you can see it's in angler trips, and there is a range, ranges, and then that is converted into a size measure for the sample selection process, which, in most cases, is just the floor of the range that you see on the left, but, for the first two, it's actually halved, because the lower-productivity sites, historically, are even less productive, potentially, than what we see, and so we don't want to be going to them too frequently, because it's an inefficient use of the sample.

Probability proportional to size takes these individual size measures and calculates from them a selection probability, which, when you then combine it with the number of samples that you're actually going to be selecting from the stratum, it gives you the inclusion probability for an individual sample unit.

Formally, what we use is known as probability proportional size without replacement, and so an individual primary stage unit cannot be selected more than once. This is a fairly standard approach. Now, we have a complication here though, because there are logistical field sampling constraints that we have to take into account, and so, for example, the state agencies conducting the field work only have so many people available. They only have so many samplers available to conduct assignments on a given day, and so we can't randomly select more units on one day than we have samplers available to conduct them, and so we use a form of controlled selection, which, again, is

not something that we came up with. It is a standard survey methods practice when you have things like sampling constraints that need to be accounted for.

Essentially, what we do is -- The way we incorporate the constraints are to generate a very large number of replicate draws, and it's sort of a bit of a bootstrapping approach, or a replication approach, and so we will draw 200,000, 500,000, however many -- A very large number of draws that have no constraints. They are just standard PPS selections.

We will then take the constraints that we have and essentially filter that large number of initial draws through the constraints and have what we will call a survivor set of draws that meet all of the constraints. Again, just by chance they do, and then, from that survivor set, we will select one at random that will become the official sample draw, and this is -- Again, it's a procedure that is used, and it's not just our approach, but it allows us to incorporate the constraints.

The bottom line is it allows us to incorporate the constraints in a way that can be appropriately accounted for in the sample weighting, and so, instead of simply running our draw procedure over and over again until we happen to get one that meets the constraints, doing it this way allows us to calculate appropriate sample weights that also take into account the sampling constraints that we have in the field. Let me back up. Again, all of this was reviewed in the reviews that both surveys have been through, and it's also stated in much more detail in the reference documentation on the sampling and estimation procedures, the MRIP sampling estimation procedures, document.

A quick slide on QA/QC. These are sort of evolving as we transition into direct electronic data collection in the field, and so, on the Atlantic coast, samplers are now using tablets in the field to collect data, and so some of these procedures have changed as more quality control measures can be pushed directly into the data collection -- At the point of data collection, but, historically, as data were collected, they would be reviewed by sort of a lead biologist or program manager, project manager, within each of the states, either at the state agencies or, previously, by the contracting company that would be conducting the intercept survey for us, and so the raw data coming in would be checked sort of as close to the field as possible, and any issues identified there could be resolved at that level.

Once that was completed, data would then be transmitted up to sort of the regional level, and, at this point, that would be ACCSP for the Atlantic and GulfFIN in the Gulf. They would then run sort of statistical software programs on the data that had a number of checks that had been developed over the thirty-plus years of the program, and these would do logic checks, and they would do range checks, looking at fish lengths and weights, fish distribution ranges, incorporating regulations, when possible, and just a number of -- Looking at a number of -- Essentially all data elements, just to try to identify errors as quickly as possible.

Data would then be, after that process had been completed, data is delivered to NMFS, and we would run the same checks again, the same statistical programs, as sort of a double-check, and anything that had been identified as a problem or unusual -- Again, a similar type of follow-up, down to the regional and state level would occur, and I will talk a little bit more later on, in the estimation section, about incorporating -- I have got another slide on outliers and sort of what we do with those, how we identify those, but, essentially, any -- When errors are identified, if estimates have already been run, and the errors are corrected, the estimates will essentially be re-run at whatever point those data errors or data issues are identified.

That's what I have for the APAIS design. Again, we'll be doing the actual estimation piece in the last section of this presentation, and so now the next is the FES design, and I don't know if we want to stop and take questions, or do you want me to keep going?

DR. SEDBERRY: I think it might be good to take some questions now, before we get too far into the FES and things get a little muddled.

DR. LANEY: John, one question relative to the shore mode for North Carolina, and I know it's split. Is it the only state that is split like that, and, if so, why is that the case?

MR. FOSTER: Yes, it is the only state that is split like that. I will ask -- I think Dave Van Voorhees might have a better insight on that. To my knowledge, it has just historically maintained that split, but, Dave, if you don't mind, if you have more insight on sort of the genesis of that.

DR. VAN VOORHEES: The stratification of the State of North Carolina into three geographic strata is something that we did with the new design.

MR. FOSTER: Dave, I think he's asking about the split between manmade and beach/bank for North Carolina.

DR. VAN VOORHEES: Yes, and that's been done since back in the 1980s, before I started in 1991.

DR. LANEY: Chris and Drew, is that because we have so many manmade structures, because we have all those coastal fishing piers?

DR. CATHEY: We have had, at one time, over twenty-five ocean fishing piers, and I think we're running around twenty now, but the split was for species of Spanish, king, and cobia, stuff that the catch rate may be a little bit different off the pier versus the shore.

DR. SEDBERRY: Any other questions on Sections 1 or 2 before we move on? Okay, John. I think we're ready to move forward.

MR. FOSTER: Okay. Thanks very much. We'll move into the Fishing Effort Survey, or FES, design now. Again, sort of just a high-level summary of the design, which many of you may already be pretty familiar with, but it's a self-administered mail survey, and so, again, we're mailing out questionnaires, and it's fairly standard household survey methodology, mail survey methodology, and, again, it's conducted for either five or six two-month waves annually, again depending on the Gulf states, including the east coast of Florida and North Carolina, and it's conducted for six waves. In the other Atlantic coast states, it's conducted for five, excluding January and February, Wave 1.

The sample frame essentially consists of all residential households that receive mail from the post office, and that coverage is estimated at 94 to 99 percent of all residential households, depending on the particular study that's been used to evaluate that, and so it is essentially as close as we can get to full or universal coverage of residential addresses in the states, in the coastal states, and this

file is updated and maintained or updated by the U.S. Postal Service as new residences are built or destroyed, in a near real-time basis.

Again, this survey is used to estimate effort, in terms of total angler trips, for private boat and shore modes, and it is for all in-state resident anglers, and I will speak a little bit more to this when we get to the estimation section, but, essentially, this covers all effort -- All angler trips taken by in-state residents for the coastal states, and it does not cover trips that are taken by out-of-state residents, tourists coming to the state to fish.

That is accounted for a coverage adjustment factor coming from the APAIS, the intercept survey, and I will speak to that in a later slide, but this design -- That was true previously of the Coastal Household Telephone Survey as well. It had to account for not only the out-of-state residents, but it also had to account for the non-coastal residents. Again, this was for the Coastal Household Telephone Survey, and so there was a much larger correction factor with the old telephone survey that was coming from the intercept survey than we have now with the Fishing Effort Survey. We're getting a much larger component of the effort estimate coming directly from the effort survey, with the Fishing Effort Survey.

The FES is not as complex of a design as the intercept survey, but it is still stratified. For each coastal state, it is spatially stratified into coastal and non-coastal regions, and the coastal region matches up with what had historically been the coastal counties in the old telephone survey, and, in most states, the coastal counties are within twenty-five miles of the coast. In some states, like Florida for example, the entire state is coastal, and so there is no non-coastal region of the state.

Within the coastal versus non-coastal regions, the sample is further stratified into those addresses that match to a license list, the National Saltwater Angler Registry, which, at this point, for all the states here, that license information is coming directly from the states.

It is the state license list that is used to populate the National Saltwater Angler Registry, and so addresses are screened, essentially, against the license list to determine addresses that have an angler, and those go into the license match stratum, and those addresses that don't match the list, essentially that don't have a licensed angler, and that's in the unmatched stratum. Then selections are made within -- Independent selections are made just with random sampling within the individual strata. Again, just simple random sampling within each of the strata, and the sample allocation that's used to divide the sample into each of the strata -- I'm sorry. Go ahead.

DR. REICHERT: Just a clarification. You mentioned earlier that the out-of-state, or you mentioned tourists, and those are people from non-coastal states as well as people from abroad, or are they U.S. citizens from non-coastal states treated differently than others?

MR. FOSTER: Sure, and so the out-of-state component would be any trip that is ending in a specific state that was taken by someone that wasn't a resident of that state, and so regardless of how they're not a resident.

DR. REICHERT: Thank you.

MR. FOSTER: Okay, and so allocation for the FES is assessed annually, and it's done using a fairly standard approach and survey methods known as Neyman allocation, which, essentially, it

evaluates the variance, the variability, of estimates within each stratum and uses that as a way to allocate sample to minimize, or to meet, some precision goals, and so to bring variances down, either globally or within each stratum or in some intermediate domain, and that is how we allocate sample for the FES, to sort of achieve a uniform precision target across the states.

In terms of the data collection design, just a little detail on this. There are three mailings to households that have been selected. The initial mailing includes sort of a cover letter that describes the goals, the purpose of the survey, with reasons to encourage participation. The questionnaire is included, as well as a postage-prepaid envelope and then a two-dollar incentive.

Our research, as well as just a lot of wealth of more general survey research, has shown that not only do incentives increase response rates, but, in most cases, they actually pay for themselves, in terms of collecting the same amount of samples, and so, if we didn't include incentives, we would have to sample at a much higher rate to get the same absolute number of returns as we do using the incentive, and it would be more expensive, more costly, to not include the incentive. We've done a number of pilot tests, and I won't go into a lot of detail, because that will be some of what Rob will speak to, but the incentives are highly effective at increasing and maintaining response rates.

The second mailing is simply a simple postcard reminder or thank you that follows one week after the initial mailing, and then, finally, at the third week, there is a full follow-up mailing that includes a letter that's been tailored to try to convert folks that might not respond, sort of addressing the usual arguments of why they may not care to respond, but trying to convince them that it's important for them to do so. A second questionnaire, in case they have thrown away or lost the original one, and then, again, a postage-prepaid envelope.

Just quickly on FES QA/QC procedures, as an initial step, once the data come in to us from the survey contractor, we verify all the deliverables, in terms of the fields being present, formats, following coding procedures, things like that, and the next step is to go through and look for data missingness or item non-response in individual return questionnaires, or the records, rather, and, to the best we can, fill in those missingness, using essentially simple logic checks.

At this point, there is really no significant major edits to the data, and then, in the next step, we are looking -- We are potentially making larger edits, if, for example, there are very far out of range values that are inconsistent with other information in the questionnaire, and the questionnaire is sort of split into household-level questions and then individual person-level questions, and, if information is inconsistent between the household-level and the person-level, edits could be made to the data elements, or the entire record may be removed from the sample, if there is no logical way to correct the errors.

Then, as a final step, there may be adjustments made to the sample weights, and I will speak in more detail about how the sample weights are calculated, but, just in this QA slide, we use standard weight trimming techniques. If we have observations that are sort of highly influential and might be classified as outliers, there are standard methods to trim the weights on those individual observations and then redistribute the trimmed weight to the remaining responses, and we use an approach that is a mean square error approach, and so it looks at minimizing potential bias by doing these weight trimmings while minimizing the overall variance, or increasing the overall

precision, in the estimates, and, again, it's not something that we came up with specifically for the FES. It's a standard survey methods approach.

The next section is the catch and effort estimation, but, again, I will stop for a second and see if there are any questions, and I should also say that Rob will be giving a lot of additional detail on the FES, and so maybe, if you have questions now, they will be addressed in his presentations as well.

DR. GRIMES: This is informational, but the original -- You were talking about stratification, and you said coastal and non-coastal, you stratified it that way, and, on a county-by-county basis, and all of Florida was considered coastal?

MR. FOSTER: Yes.

DR. GRIMES: Which geographically isn't true, but it's closer to the ocean, and maybe that's rationale for doing this.

MR. FOSTER: Correct that certainly not every county in Florida has a coastline, but it's really more of just -- I would think of them more as labels, and it's really more about sort of activity levels and how much alike the counties are, in terms of their overall prevalence of fishing in households, and so all of Florida, in general, would have a higher level of fishing activity than say the non-coastal counties in other states.

It's really these labels are sort of a holdover from the CHTS design, where the non-coastal counties weren't sampled at all, because the probability of -- The fishing prevalence was so low that the chance of actually getting a fishing household by random digit dialing was essentially zero, and they would have to make an unaffordable amount of calls to get fishing households, and so the estimation was done differently for the Coastal Household Telephone Survey. It's really more about sampling efficiency, this particular stratification, but, yes, I agree that not every county in Florida is truly coastal.

MS. SAULS: I would also like to ask why was Florida not split regionally? I think the old Coastal Telephone Survey -- It asked about trips on the Atlantic and the Gulf coast, and the new survey -- It's just a simple random sample of any household that doesn't have a matched fishing license to it with no regional stratification, and I wonder why that decision was made, and I'm not necessarily sure that it's appropriate either to assume that someone who lives fifty miles from shore behaves similarly as our larger population of coastal people in Florida.

MR. FOSTER: Rob, would you like to respond, or do you want me to start?

MR. ANDREWS: I will address the first part of that, with the stratification in east and west Florida. You're correct that the Coastal Household Telephone Survey split Florida -- It stratified Florida into east Florida and west Florida and so we selected independent samples from those two strata, and then, when we conducted the interview, we could get details about the county in which the fishing trip occurred, and so, for east Florida, we could identify trips that were taken in east Florida, as well as in west Florida, but we would only produce an estimate for east Florida, for trips that were taken in east Florida.

A big difference between the mail survey and the telephone survey is that we tried to make the survey as short and as simple as possible, and so, right now, it's just a four-page survey, and it's actually -- I don't know the exact dimensions, but it's eleven-by-twenty-two, and so you fold it in half, and it's just a booklet, and so real estate is a premium, and so we can't -- We don't have the space on the questionnaire to ask about trips in east Florida and west Florida.

Now, we have talked in the past about creating an east Florida FES that would only ask about fishing activity in east Florida, and then also do the same thing for west Florida, and so ask about trips that specifically occurred in the Gulf of Mexico and then stratifying the same way we did in the past, but we just haven't been able to test that yet, and I guess, to your second point, we're not -- When you stratify, the purpose of stratification is to improve the efficiency in the variance of your estimate, and so, if you create strata that are very similar -- If the individuals within a strata are similar, with respect to fishing activity, and so you're maximizing the homogeneity within your strata, in terms of what you're measuring, you'll come up with a more precise estimate, and we're not stratifying Florida geographically, but we're not introducing a bias by not doing that.

It might not be the most efficient estimate, and there might be ways to stratify Florida to come up with a more precise estimate, and, again, these are all things -- Like we didn't develop the FES and put it out there and call it done. We are constantly looking at ways to improve the methodology and looking at stratification is something we're going to be looking at in the future, and geographic stratification of Florida is something I'm sure we'll be considering at some point.

MR. FOSTER: One additional sort of follow-up on that is, when we look at the geographic distribution of samples in Florida, even given that it is not formally stratified within any sort of sub-state regions, the distribution of sample very well matches the distribution of the population within Florida, and so we don't -- The distribution of the original drawn sample, as well as the distribution of the responses that we get back, and so we don't have, for example, a situation where we've got differential response rates, say in the Panhandle versus south Florida, where we're getting a disproportionate amount of the respondents coming from just one part of the state or another, where their fishing rates, in terms of trips they are taking, might be very different. The sample is well distributed geographically across the state, and that is something that we look at as the data come in.

DR. SEDBERRY: Any other questions on this FES section?

DR. CROSSON: I had a question about the timing of the mailing, and I was wondering when is that initial mailing sent relative to the wave that's being sampled? Is it sent towards the end of the wave or near the beginning?

MR. ANDREWS: We send the first mailing approximately a week before the end of the wave, and so the goal is to get it in the hands of the household by the end of the wave, and then all of the subsequent mailings are based upon that initial mailing date, and so the first mailing is a week before the end of the wave, and then we send a postcard a week after that, and then we send a second version of the questionnaire, or a second questionnaire, two weeks after the postcard, and so all the mailings are done within a three or four-week period.

DR. CATHEY: Under the Coastal Household Telephone Survey, North Carolina had a concession where we extended our coastal zone a hundred miles inland, and this was to account for large metro areas, particular Wake County, Raleigh/Durham, and is this still the case with FES?

MR. ANDREWS: Yes, we match the stratification. The FES, the current FES, stratification matches exactly with the CHTS, and that was by design, because we were making direct comparisons between the two, and we wanted to be able to have them line up exactly. We are continuing to do that, even though the CHTS has ended, but, again, that's something we'll continue to look at, but the goal is not to do what we always did with the CHTS. We moved away from the CHTS for a whole lot of reasons, and so the goal is to come up with the most precise and accurate estimates, and so, moving forward, we'll continue to evaluate the stratification and the sampling design, to try and improve the efficiency, but, currently, we're still doing it exactly the same. Our coastal stratum lines up exactly with the coverage of the CHTS in every state.

DR. SEDBERRY: Any other questions or comments on this FES section?

MS. SAULS: I just want to -- You said that the stratification matches, but the Coastal Household Telephone Survey was stratified by county, and so you had a much different distribution of your sample, I think, and I know you say that it matches the population, but we all know that fishing pressures are very different in different parts of the state, especially a state like Florida, and it can vary seasonally, and it varies by when different fishing seasons open and close, and so it's not necessarily that you are sampling proportional to where people live, but it should be where people fish, because you're sampling all households in Florida.

MR. ANDREWS: To address Drew's question, the coastal stratification for the FES matches the historical coastal coverage area of the CHTS. The coastal stratum in the FES matches the coverage area of the CHTS, yes, and we are not stratifying by county in the FES. Within strata, we're doing a simple random sample, and, yes, that is correct.

MR. FOSTER: To speak to that, I think, Bev, you're right that, if we could make the sampling more efficient, and potentially, again, assuming that there is funding to support the required sample sizes, because, as you create additional strata, you often have to increase the sample to have minimum sample sizes in each of the strata.

I think we could certainly make it more efficient with the right -- Creating the right county grouping strata, but even -- Without that, we're still not -- As Rob said earlier, we're not introducing a bias by not having that, because we expect that, again, as people are reporting -- As they are fishing differently in different parts of the state, their responses, the responses that are coming in from people in different parts of the state, should be reflecting that, and so, again, because we don't end up with all of the sample coming from one portion of the state, and it's coming in representatively from where people live, and they should be reporting as fishing changes across the state. Say the mix of zeroes in the data coming in should be changing from -- As the seasons change of fishing or the number of trips people are taking, on average, coming in from different parts of the state, that should be reflecting what they are doing.

Yes, I agree that, if we can stratify the state in a way that it can make the sampling more efficient and improve the precision on the estimates, but not having it doesn't directly equate to having some bias in the data.

MR. ANDREWS: To follow up on that, we are stratifying also by license match, and we can sample our strata at different rates, and so we sample the license strata at a much higher relative rate than we do the non-matched strata, and so we are targeting our sampling toward people who are likely to report fishing, but one thing to keep in mind, and I will talk, I think, a little bit more about this later, is the FES is not an angler survey. We are sampling households.

If we get more anglers responding, we're actually going to have biased estimates, and we're trying to get all households to respond and not just fishing households, but, if we can stratify in ways that allow us to identify and isolate, if we can, anglers in likely fishing households, then we can certainly put more sample into those strata, but the goal is not to just sample anglers. It's to sample households.

DR. SEDBERRY: Beverly, did you have a follow-up for that?

MS. SAULS: You said you sample the matched houses at a higher rate than the unmatched, because that's not what your current statement of work shows for Florida. I am looking at the 2018-2020 statement of work that I found online, and, just for Wave 1, the sample selection for the unmatched households is about a thousand, and the matched is about 206 households.

MR. ANDREWS: What was the matched and the unmatched again?

MS. SAULS: Unmatched was 1,000, and matched is 200.

MR. ANDREWS: Okay, and so what percentage of the population is licensed?

MS. SAULS: We have over a million license holders.

MR. ANDREWS: Right, and how many people live in Florida? It's over twenty-million, right? 5 percent of the population is matched, and approximately 10 percent of the sample is going toward the matched.

MS. SAULS: So you're saying a higher rate proportionately and not a higher sample.

MR. ANDREWS: Right, and, again, the goal is not to -- The goal is maximize precision for the state, and so we use historical information that includes fishing rates and mean trips per household, and that goes into this Neyman allocation model, and it distributes the sample in a way that maximizes the precision for the state wave, based upon historical information, and so it's not -- We're not saying that 50 percent of our sample is going to matched households. We're saying X percent of our population is going to matched households to maximize the precision of the state level estimate.

MS. SAULS: So a 30 percent response rate on 200 licensed households is a pretty low sample size, I feel, for the entire state of Florida.

MR. ANDREWS: It could be, but the response rates for the matched stratum can be as high as 65 percent. 30 percent or 35 percent is a weighted response rate across all households, but the response rates across strata are very different, and so we might have, in Florida, a 40 or 45 percent

response rate for the matched strata, and so, yes, I mean, it's still -- You still have non-response, and you still have a certain amount of sample that you can purchase, but it's still producing an estimate with a PSE of less than 20.

MS. SAULS: Okay, and so eighty responses for the entire state of Florida is representative of all the license holders?

MR. ANDREWS: If they are sampled in a representative fashion, then, yes, it's a representative sample, by definition and by design. It's a representative sample.

MR. FOSTER: We certainly wouldn't be opposed to increasing the sample size in the FES, in Florida or across-the-board, but it just requires additional funds to do that.

MR. WILSON: This is a question that kind of hits on something that Marcel asked about, and what it pertains to is the offering the non-coastal -- People who are out of state, and so your people are out of state, and you do an adjustment for those, but then you have public and private access that you're getting your APAIS sample from, or the APAIS sample is coming from public access, and you would think that the public access areas, like in Dare County, would have a higher rate of people from out of state versus folks fishing from private access would not probably have a -- Private access would probably not need much of a correction factor, whereas these coastal areas with public access would have a correction factor, and so I was just trying to figure out how you apply the adjustment, non-frame adjustment, to your estimate of trips when you don't know how many trips, what percentage, are going to be private access versus public access, because I think it would be a different rate that would need to be applied to those.

MR. FOSTER: That's an open question. We don't have an estimate of what trip characteristics are like, just in general, not just with regard to resident status, but catch rates and a whole number of things, for trips returning to private access, and so I would agree with you that, in some cases, in certain areas, there may not be any out-of-state folks taking trips from private access, but, in other parts, where you have seasonal rentals and things like that, there may be out-of-state folks that are fishing, but we don't have a way to quantify that, fishing in shore areas that may not have APAIS sites, where they can be sampled.

What I would also say though is that, in general, the percentages of trips in the APAIS that are out of state are small, and it's not like we see for shore and private boat mode. We don't see 50 or 75 or 80 percent of the trips being out-of-state. It's usually more like 15 percent, 10 percent, 5 percent, or less that is out-of-state, and so even -- Yes, I agree with you, and I think there could be differences there, but what we already measure with just the public access sites, the large majority of those trips for private boat and shore mode are already in-state. There is not that much, in general, as a general statement, that is out-of-state. In other words, the size of the potential bias there is already pretty well bounded, if it was -- If there is an extreme difference in the public versus private access.

DR. SEDBERRY: Yen, but, before you state your question, I should say that, if I don't call on you by name, don't take it personally. I am terrible with names, and sometimes it just escapes me in the moment, and I forget people's names, and so, if I do that, if I just say, hey, you, please identify yourself on the mic, so that we can get an accurate transcription. Thank you.

DR. LI: First, I am so glad to hear that, for comparison purposes, the spatial coverage and the spatial strata from these two surveys are matched, and so that's a relevant factor that we think may have caused the discrepancy between these two data streams. Then my question is, even though the spatial coverage and the spatial strata are matched, how about the population we survey within strata? Are they still matched?

For example, like, for example, within a coastal region, a coastal strata in North Carolina, for example, the sample population, they may have like a ten-day license, or they may have like a lifetime license, or they may have an annual license, and so the FES survey with the telephone survey, how about the percentage of those license types represented within the survey? Would that match between these two surveys, or something similar like this, some major questionnaires, how about the population of these surveys within each spatial strata? Are they still matched between these two surveys?

MR. FOSTER: I will start, and Rob, if I leave anything out, please fill it in for me. The short answer is the FES is providing as complete coverage of in-state residents as really can be achieved in standard survey methods. The CHTS, on the other hand, the landline telephone survey, and we will go into this, Rob will go into this, in great detail in about two presentations from now, that is one of the big drivers of the difference, because that population that was being covered was changing.

I mean, it was still people that had a landline telephone, but the demographics of people that have a landline telephone and are willing to answer it was changing dramatically over the last -- Certainly the last ten years, and really going almost all the way back to 2000. We have information, and there is information from a number of sources, not just our Fishing Effort Survey, but a number of other surveys have collected information on phone status. Big, national, federal surveys have collected information on phone status, to be able to monitor the changing demographics of people that have a landline telephone and, again, will answer it.

I would say, yes, the fact that the population that the CHTS was covering was changing over time is a big source of the difference in the estimates, but, really, it's -- We will present it to explain why the -- Because it was the declining coverage in the CHTS, and we'll show why it was introducing issues, problems, biases, into the CHTS estimates. Rob, do you want to follow on any?

MR. ANDREWS: Just quickly, because you mentioned the licenses, we didn't actually include any information about fishing licenses in the Coastal Household Telephone Survey design either, and so that was just a random sample of landline telephone numbers from within coastal counties, and so that wasn't a component of the telephone survey design, and so we don't really have that information available from the telephone survey to even see if the distributions of license types are comparable to what we did from the FES.

MR. WILSON: For the purposes of your license match with those strata, is it proportional to a license type, and so whether it be like a short-term or lifetime, but is a license just treated as a license, and it's a yes or no?

MR. ANDREWS: We match every residential address in North Carolina, for example, and so, in the coastal stratum within North Carolina, the sample frame is every residential address in that

strata. We take all of those addresses and match it to the entire North Carolina license database, and so whatever you guys, the state agency, provides, every wave, the list of licenses to our office. Then we dupe that, basically, and it's angler license database, but we sample households, and so we eliminate duplicate addresses, and we match the entire address frame to the entire license list, regardless of what type of licenses there are.

The list that you guys give us, or the list that we ultimately use, includes anyone who was licensed during any portion of the wave, and so, whether you had a ten-day license, an annual license, a lifetime license, if you were licensed during any portion of the wave, you are included in the list that we used for matching, and, again, that's just used to stratify the address base sample, and we're not selecting anything directly from the license list, and so we're only sampling from the address frame, and so, if there were more annual licenses, for example, because it's a probability survey, and it's a very high probability that we would have more of those types of licenses in the address sample, but we're not controlling that in any way. It's just whoever is licensed during the wave.

MR. WILSON: Okay, and so trying to piggyback on -- So had coastal, and so you had the coastal zone in both groups, and not so much about the license types, and this might be going all the way to the end, to figure out what -- Because I'm concerned about the shore mode, really, which I don't know if we can get down that path, but so, in the coastal population, you've got the coastal population with FES and the coastal population with the Household Telephone Survey. What changed? I mean, it's like the CHTS expanded with census data, and the FES constrains the growth of the estimate with census data, and is that correct? I don't know -- Is there a huge difference? Is there any change in the way that the census data is done? Also, it's like the avidity from the CHTS on positive hits within the coastal, is it much different than what the FES is showing?

MR. ANDREWS: I am going to be talking about the differences and the impact, and that's two presentations from now. We'll get to that on Thursday.

DR. SEDBERRY: I am going to suggest that we move on, and some of those presentations that are coming later might actually address some of the details that we're kind of getting into now, and so, if that's okay with everybody, and, since I'm the Chair, it's okay, I think, but I don't want to -- If you have some pressing, important question to understand what we've done so far, we can go ahead and ask that, but I'm going to suggest that we move on to the next section and the additional presentations, so that we can see if those details are covered in other presentation materials. Let's move on. Thank you, John.

MR. FOSTER: Okay, and so the next section is on catch and effort estimation. Again, it's going to focus on the way that the FES and APAIS data are combined, and so we're not really going to touch on charter mode here, and it will be limited to private boat and shore, but the methodology is very similar for charter estimation. If there are questions, I am happy to speak to it.

I'm sure folks, a lot of you, if not all of you, have seen this slide a bunch of times, and it's just to sort of give the high-level breakout of the estimation, and then we'll go into details, but, essentially, again, we have effort, in terms of total numbers of angler trips, estimated total numbers of angler trips, which is used to expand or multiply a catch rate, an estimated number of fish caught per angler trip, to give us then the estimate of total catch, and, again, this is done in a number of

different domains, by species, by catch type, landings versus releases, all of those things, and we'll get into those details in subsequent slides.

Again, for what we're discussing here, the bulk of the effort information is coming from the FES, again for private boat and shore modes, and I'm going to be talking about how -- It's weighted estimation, and so I will start with talking about how the sample weights, the components of the sample weights, are calculated, and then we'll talk about how those weights are used in what are essentially a standard estimator of weighted totals, and it's really just the Horvitz-Thompson estimator, and then there are two pieces of the effort, of the effort estimation, that do come from APAIS, and one we've talked about already a little bit, which is the FES only covers in-state residents, and so we need some way to account for the component of effort that is coming from those out-of-state residents, and so, again, that information comes from the intercept survey, from APAIS, and I will talk about how that is calculated and used.

Then a second component that comes from APAIS is the information that we use to partition the total effort estimate into the different fishing areas, and so inland waters, state territorial seas, which, for the most part, is open ocean, zero to three miles out, and then the EEZ, generally three miles out to 200 miles. That information, those proportions, are calculated from the APAIS data as well, and I will present those calculations.

Then, on the APAIS side, again, we mostly think of coming from APAIS is the catch rate information, the catch counts by angler trip, and I will talk about specifically how the APAIS sample weights are calculated, and, again, present how those weights are used in the calculation of the catch rates as well as these two other effort components.

Just revisiting the multistage nature of the -- I will start on the APAIS side, and so revisiting the sort of multistage design aspect of APAIS, each of those states potentially contribute to the final APAIS sample weight, because we can have sub-sampling occurring at each of those stages. Again, the highest stage, the primary stage unit, the PSU, that's the site cluster by day of the month, by time interval sample unit, the assignment we often call it, and so we select a sub-sample of those from the overall full frame, all of the access sites, all of the days of the month, all of the six-hour time intervals for each day.

Within the sample site cluster day time interval then, we have how long was the sampler on-site, and was it a two-site cluster, so that they were only at each site for up to three hours, up to half of the six-hour interval, or was it a one site and they were there for the full interval, in which case there might not be any sub-sampling, or were they not sampling for some reason, like they were in transit, or they had to take a break from sampling. Any kind of reduction in the time that they spent sampling is accounted for in this stage, a sort of temporal sub-sampling.

Then, finally, we have the angler trip stage, where, again, it's a busy site, and the samplers can't sample everybody, but they are counting the angler trips that they're missing, and so we know the sampling fraction at this stage, how many trips were returning to the site while the sampler was sampling and how many did they actually sample, and so there could be a sub-sample there as well, although, at a low-activity site, they may sample everybody, and so there might not be sub-sampling in some cases.

Sort of walking through these stages then in a little more detail, again, at the first stage -- These are the PSUs, the primary stage units, that we're sampling as part of our APAIS draw process, and we have an inclusion probability, really a chance that that particular unit was sampled, or selected, in the draw, and then the sample weight, at this stage, is just the inverse of that probability, the inclusion probability, and, again, this is standard sampling statistics, sample theory, and this is how the weighting works.

It's the inverse of the probability that a unit is selected becomes its sample weight, and so, again, in a simple, simple random sample, and let's say you had ten units, and each one has got an equal chance of being selected, and you only select one of them, you have selected one out of ten, and so its probability of being selected was one out of ten, and so its sample weight then is the inverse of that, or, again, ten, and so the one unit is representing all ten. That's a real simple example, but that's how you just invert the probabilities to get sample weights.

Stage II, again, it's very similar, and it's just, of the total six hours, what fraction of what were they sampling, again, at the given site, and then the weight becomes the reciprocal of that, and so, if they spent three hours at a site, the weight would then become six over three, or two, and so we're just doubling the weight at that stage, which makes sense. If they spent half the time at one site, you would want to double whatever they collected at that site to represent that full six-hour window of time.

Then, finally, at the third stage, again, this is angler trips within the sampled site. For shore mode, it's a little bit simpler. The Stage III weight is just essentially the reciprocal of the sampling fraction, and it's just the total count of angler trips that returned while the sampler was sampling divided by the total number that they intercepted, again, of angler trips, and the sampling fraction would just be the number they intercepted divided by the total count they had.

For boat modes, it's a little -- There is one additional complication, because there's another potential stage of clustering. On the left-hand side here, we're accounting for the fact that, for a given boat party, a given boat trip, they may not be -- The sampler may not be able to interview everyone in that trip, and so there is a sub-sampling at that point, and then the right-hand side here is, beyond that, there will be boat trips, entire boat parties, that they may not be able to sample, and so you have a total number of trips that return to the site divided by the number of interviews that they had across all of the boat parties, and so, again, it's just two stages of potential clustering, because we have -- We are sampling boat parties from all the boat parties that returned to a site and then, within an individual boat party, you may have sub-sampling there as well, and you may only get two or three of the four or five anglers, let's say, that were on a specific boat, and so, again, there's just sort of two potential sub-sampling steps there.

The final weight then becomes the product of the weights at each of the stages, and that final weight then is what is used in the weighted estimation, and now I'll sort of just walk through the calculation.

DR. ERRIGO: In that calculation for the boat modes, if you -- Let's say there is five people on a boat, and they come off, and you interview one person, but they give you the catch for the entire trip, do you then not have to account for not interviewing each person?

MR. FOSTER: Right, and so it can vary by the catch disposition, and so, for the unobserved catch, and that's catch that is just reported by the anglers to the samplers, and so these are things like released at-sea or dead discards, or even potentially landed fish, but the interviewer was not able to count them directly for any reason, those catch types are always by angler, and so, if there's a sub-sampling within the boat trip, the expansion for those catch counts will always be by angler, but Mike is right in pointing out that the observed catch can be grouped across anglers, and so, if it's the case that the observed catch that is -- The landed fish, observed landings, observed harvest, if that is fully enumerated by the sampler, and, if that reflects all of the anglers on the boat, then, yes, that would not be expanded by this piece, and that's right, but, if there are other people on the boat that were not part of that grouped catch, then it would be expanded to account for those folks, yes.

Again, just to walk through sort of the catch rate estimations, and these are just sort of standard weighted calculations for means, and these are done within domains, and so this is just sort of a standard domain estimation for survey methodology, where, to calculate the mean, you essentially have, for each sample unit, you have the final stage APAIS sample weight, you have Y here that is representing, in this case, the fish caught on that particular angler trip that corresponds to the sample weight, and then you have an indicator variable, as to whether or not this trip was in the particular domain that's being estimated here, and so what are our domains?

These are things that are defined by sort of as our standard estimation cells, and so domains are defined by management region, state, year, two-month wave, fishing mode, private boat or shore, area fished, inland, state seas, EEZ, and then, also, specifically for catch, you have species and catch type, and so are these observed landings, reported landings, released fish, things like that, and so that's the numerator.

Then the denominator then is just the sum of the weights, again, with the -- It's the sum of the product of the weights and the indicator variable, again, is the observation in the domain that we're currently estimating for or not, and so then that ratio gives you the weighted mean, and we -- Getting a little bit into the programming side of it, we use SASS for all of our standard production estimation coding, and, in particular, we use PROC SURVEYMEANS, which is its standard procedure for doing survey estimation of means and totals, standard estimators.

One thing that I want to sort of mention, in particular here, because sometimes we get questions on this, is there is often -- We need to clarify how we calculate the catch rates for a specific species, because, often, folks -- It's sort of intuitive to think we're going to calculate a catch rate for a species that is just based on the trips that have catch of that species, or perhaps have catch of a species or targeted that species, and then, if people think that, and then they hear that we expand those catches rates by the total effort across everything, then it seems like we're probably overestimating catch, potentially, but that would be a problem if we did it that way.

The way that the catch rates are calculated for a given species is using all of the trips that are intercepted within the rest of the domain, and so, for a given state and wave and mode and area fished, all of the trips are used to calculate the catch rate for a given species. Trips that don't have catch of the given species are just counted as zeroes, and so they are contributing zeroes into the catch rate estimation, into estimating the catch rate, and so, again, if we're calculating black drum -- If we're calculating a black drum catch rate, we'll have a few trips, maybe, that do have black drum catch, and we'll have a lot of trips that don't have black drum catch.

All of those trips that don't have black drum catch are still included in calculating the catch rate, but they just are contributing zero values into that numerator, or that Y, and, again, we're doing that because the effort that is going to be used to expand this catch rate includes all of the trips. It's not just black drum trips. It's everything, and so those zeroes have to be in there, or we would end up with a biased estimator of total catch.

Now, just to finish out on the APAIS side, I will walk through sort of the coverage adjustment factor for the FES, again for those out-of-state trips, and then we'll go to the area fished proportions.

Very similarly, it's a proportion, but it's calculated much like the catch rates. Again, it's the APAIS sample weight in the numerator multiplied by an indicator variable, as to whether this particular angler trip was taken by an in-state resident, in which case it's given a value of one, or if it's an out-of-state resident, and so the proportion that's being calculated here is the proportion of total trips within the domain that were taken by in-state residents, and that's this quantity, and, again, just as a weighted proportion. Those weights, again, are needed to account for the complex design of the APAIS.

Then, similarly, for the area fished proportions, again, it's a weighted mean here, where it's the weight for an individual angler trip multiplied by an indicator variable, which is indicating whether or not the trip was an inland trip or a state territorial seas trip or an EEZ trip, and we use that to calculate the proportions for each of those three areas, and then I will show how this proportion is -- How all the pieces are put together in just a moment.

We're almost there, but, just quickly, I will go through how the FES sampling weighting components are calculated and how the final FES sample weight is put together and then the calculating equations for the effort side, and so, for the FES sample weights, there is three main components, and there is sort of the base design weight, and that's the starting point, and we have a simple random sample within the strata, and so that weight, which is essentially just the -- I think I've got it on the next slide, and I will go to it then.

We have the base design weight, and then we make an adjustment for non-response, and I will talk about how we do that, and, again, this is to account for the fact that, for example, people with a license, people with a boat, people that are taking trips, they might be more likely to respond than people that don't take any trips, and so we want to make an adjustment to account for that, that difference in response rates, and then, after that non-response adjustment, we make sort of a third set of adjustments that help improve the representative nature of the sample, and so these are poststratification or raking adjustments, and, again, this is fairly standard practice, and it's essentially we collect a lot of demographic information in the sample, in the survey, and we can see how those demographics, the proportions or totals, match up to control totals, demographic control totals, from the U.S. Census Bureau, and so, in our particular sample, are we over or underrepresenting seniors or households with children or different demographic groups, and, based on -- Compared to information from the Census Bureau.

By making those comparisons, we can then adjust the sample weights so that the demographic estimates line up, our survey estimates will line up, exactly with the information from the Census Bureau, and, to the extent that those demographics correlate with differences in fishing effort, we

are improving the estimates. We are improving the accuracy in the estimates by making these adjustments.

Starting with the base weight, again, it's -- Again, it's just the inverse of the inclusion probability for a sample unit, but, in this case, within strata, we have simple random sampling, and so the inverse of the inclusion probability is just the total number of units in the frame, and, in this case, it's the total number of residential households in the stratum divided by the sample size in the stratum, and we haven't made any sort of non-response adjustment yet, and so this is just our total sample size, and we'll get to how we adjust for the folks that don't respond in just a minute, and so that's the base weight that we're starting from. If we had just 100 households in a stratum, and we selected a sample size of ten, then the weight here would be 100 over ten, or just ten. Of course, the big Ns here are much bigger than that, and the sample sizes are a small fraction, and so the weights here can be sizable.

To that, then we make a non-response adjustment, because, again, there will be a large number of folks that don't respond, and this is to try to minimize any non-response bias and make sure that, to the extent we can, we're accurately representing the target population, and so, essentially, we want to, as best we can, match up non-respondents to the respondents that are most like them, and we do that by creating what are called adjustment cells, or weighting class adjustment cells, and this is, again, not something that we came up with on our own. It is a fairly standard approach to making non-response adjustments.

The first step in this is to figure out meaningful ways to design or define these weighting adjustment cells, and so we use our stratification variables, and we have the coastal zones versus the non-coastal zones, and we have the license matched versus the unmatched, and then we have some additional variables, and so, for example, we bring in state vessel registration databases, and so we can see that, when we match, or when we link, vessel registrations to households, those folks that have a registered boat in their state, they are more likely to fish, and they are more likely to report fishing, and they are more likely to respond to the survey, and so it's important for us to take that into account in our non-response adjustment, because, if we didn't, we would have the potential for a positive bias there, and similarly with just folks that have a license and don't have a license.

The folks that have a license are more likely to respond, and they're more likely to report fishing when they do respond, but we're using that information here to define these adjustment cells, and, essentially, what this really means is, within these cells, the people that respond are upweighted to reflect the households that didn't respond in this same cell, and so people living in the coastal zone that have a license that respond, they are upweighted to only represent themselves and other people in the coastal zone that had a license that didn't respond.

Similarly, for folks that live in the non-coastal zone that don't have a license, the ones that respond are only upweighted to reflect the folks that also live in the non-coastal zone and also don't have a license, and so, in this way, we don't have people that do live in the coastal zone and do have a license and respond. We're not having them represent people that they are not like. They are not being upweighted to represent people that don't live in the coastal zone, people that don't have a license, and so it's a meaningful way to account for the non-response in a way that helps, two things, improve the variance, but also minimize the bias, or improve the accuracy.

This just speaks to sort of how this is done. It's essentially, within these weighting class adjustment cells, you take the initial weight, which was that simple random sample weight, and you simply divide it by the response rate, the weighted response rate, and that's just the math of how this works, but the meaningful part of it is that, within these adjustment cells, the respondents are upweighted to only represent the non-respondents in that same adjustment cell, and, by defining the adjustment cells, using variables that are well correlated with response and also reporting fishing, that both improves the variance and also reduces the bias. Rob may speak more on this in a later presentation, but I wanted to go ahead and present it here as well.

Then, finally, the post-stratification raking adjustments that I mentioned, this is another sort of ratio adjustment, where this w* that I am circling, that's the weight coming out of the non-response adjustment stage, and, to adjust that for this post-stratification raking approach, we simply multiply it by the ratio of our control total, this theta in the numerator, and that would be coming from the Census Bureau, for example, say the total number of households with children in the state, as just one example, and then divided by the current estimate of it from the FES, not using the final weights, but just using the non-response weight, and so, if these two things are already matched, already equal, then this is just one, and the final weight is our non-response weight, but, often, these are off by a few percentage points, and so there is a small adjustment, whether it's a little bit greater than one or a little bit less than one, and it just sort of fine-tunes the weight coming out of the non-response adjustment, and that gives us our final weight, and this isn't just one adjustment. This could be made for a number of different demographic estimates.

Then, just to walk through, the FES effort estimate is fairly simple, once we have the sample weights, and it's just a simple weighted total, where we have, for each individual responding household, we have their final sample weight, and then we have a total number of angler trips taken by the household, and it's the weighted sum of that product within the different domains, and so, again, state, wave, mode, year, and that gives us the FES total effort estimate, but, again, at this point, it still only represents in-state residents. We have to make that one adjustment for the out-of-state folks.

That's where we are now, and so, to get from the FES effort estimate that's the total for in-state to the full total for the state, which includes the out-of-state anglers, we simply take the FES effort estimate and we divide it by that proportion, the weighted proportion, of total trips that were taken by in-state residents, and so, again, a simple example is, if in-state residents -- It's much closer to 85 to 95 percent, but let's just say it's half. Half the trips were taken by in-state residents, and this proportion would be 0.5, and we would essentially be then doubling the FES in-state effort estimate to account for all those trips taken by the out-of-state folks, but, again, normally, these proportions are 80 to 95-plus percent, and so this really -- This adjustment doesn't tend to increase the FES estimates that much.

Then the last step is to take that total private boat or shore fishing effort estimate and partition it out into the different fishing areas, again, inland or state seas or EEZ, and that is simply done by taking the total and multiplying it by those individual area proportions, again which are coming from the weighted APAIS data, and that gets us to total private or private boat or shore fishing effort by area fished.

Then, finally, once we have that effort piece, which is now -- It shows all three of the components here on the left-hand side, in the parentheses, and we multiply -- We take that effort estimate,

which is, again, now adjusted for coverage, and it includes the out-of-state component, and it's been partitioned down to area fished, and we multiply the individual catch rates that have also been calculated again down at the area fished level.

It's for individual species, but remember that it includes all of the trips in the domain, and it's also by individual catch type, and so whether it's observed landings, reported landings, live releases, and so that product then gives us the corresponding total catch estimate, and, again, this is down at the lowest level, and so state, wave, mode, area fished, region, species, catch disposition, and then these would get summed up across waves, for example, to get you to an annual estimate, or across states, to get you to a regional estimate, or across modes, to get a full estimate for accounting for all the modes across areas, to account for all the landings, and those steps are just simple sums of these more detailed -- The estimates at this lowest level of detail.

I think this is the last slide. All right. Identifying outliers, and here is a place where I apologize in advance, and I'm sure folks would like a lot more detail, and I'm happy to walk through with questions, but I am trying to -- It doesn't seem like it, but I am trying to keep this presentation sort of contained a little bit, and so, anyway, in terms of outliers, the first step is sort of what do we do to review the estimates, just to begin with, regardless of whether it's an outlier or not?

The first step is, as the estimates -- Really, the individual estimate components are being produced, and so, before we put everything together even in terms of catch, total catch, we've got all those components that I mentioned, the catch rates, the area fished proportions, the coverage correction factors for state resident type, and even the sample weights that are calculated before those.

We sort of apply a time series method approach, where, essentially, time series modeling, if you're not familiar with it, one of its components is, instead of modeling the individual estimates in a time series, it will actually model, in part, the differences of the time series, and so what are the changes, say from wave to wave, consecutive waves, or what is the change from this wave in the current year from the same wave in the previous year, and so it will -- What we do is, instead of trying to look at the individual points as they are, we'll look at the differences, because that sort of combines both big relative changes, but also big absolute changes, whether it's a big increase or a big decrease, and it sort of gets added at the same time, and it also lets us know whether we're seeing a trend. It would take out the effects of trends over years or seasonal patterns within a year that are typical, and so, if we see a big change, whether it's a big increase or a big decrease, it's likely to be something that's not a usual pattern in the data. That is one of the main reasons for doing it.

This is done sort of for everything. All of the estimate series sort of get this treatment, and, essentially, if one of these differenced values for the current batch of estimates we're looking at is above or below two standard deviations from the mean of that differenced series, it gets flagged, and then it goes through a review process, and I will talk about that in just a second.

In addition to that step, we also manually review a list of sort of high-priority, high-priority from a management perspective, species that have regular assessments and that have ACLs and things like that that have high scrutiny, just in general, and those get a manual-level overview, initially internally by our staff, Science and Technology staff, but then NMFS staff out in the Regional Offices and the Science Centers that are familiar with the recreational estimates also do their own reviews of the estimates as well.

Then anything that gets flagged by these three sort of processes gets a data investigation, and so we go back and first look at what is likely driving or what is the driver for that big change, and generally it's a big increase, but sometimes it could be a large decrease, and, often, it's something coming from a single cell, a specific state or mode or area fished, and so we start digging and sort of winnowing it down to what the cause is, and then we go in and see if it's something on the data side, if it's a reported catch count value that is exceptionally high, or is it an influential cluster of observations, we had one boat trip that had several people in it that all reported maybe a high catch of something or a prohibited species or something like that, or is it a sort of rare-event situation, where it's a low-activity wave, and we don't have a lot of intercepts in a specific wave and state and area fished and mode, and what we happen to get, just by chance, must have been some unusual species or a high count in the limited amount of data that we did collect. All of those are sort of potential contributors to what we see.

Now, if it's an actual data error, if it can be corrected, it's corrected, and the estimates are re-run. If it's an error that there is no logical way to correct it, then the decision is generally made to just remove that record and re-run the estimation without that included. If there is an outlier in the data, but there is no obvious error, and the data have all been checked, all the way back down, ideally all the way back down to the sampler, which is usually the case, then, generally, those estimates are left in, and we don't try to make any decision on how should we make some ad hoc post-hoc adjustment to the estimates. We just leave them as they are and leave it to data workshops and assessment workshops and the SEDAR process to best decide how to handle them.

We do document the data investigation that is done, and also potentially not just the data investigation, but an estimation investigation, because sometimes there is not necessarily an influential value in the data, and it could be an influential value in the sample weights, and so we will investigate that as well, and we'll also investigate to make sure that just the estimation process worked as it is intended to.

A brief overview, that is essentially how outliers are handled. Now, in the fifth presentation, where I will walk through sort of the general changes, and specifically large changes, for specific point estimates for the five species that were identified in the terms of reference, I will talk a little bit more about this, but mostly focus on sort of -- I will focus on what the drivers are for those large changes, whether it was an outlier to begin with or whether it's one of the two calibrations or a combination, through an interaction with them, and so I'm not to go into specific point estimates here, but we will do that in the fifth presentation for all five of the species that were identified. With that, I think I'm done, and so thank you. Sorry that was so long.

DR. SEDBERRY: Thank you, John. That was a very detailed presentation on really a complex process and some very long documentation documents that were in the briefing book, and so it was necessarily detailed, and I appreciate your coverage of it. I am going to propose that we take a break right now, so that John can rest his voice before he has to go on, and, if we have any questions, we can kind of formulate those during the break and then come back at 3:30, and so a short break, and then we'll get started. We'll take any questions that remain on what we've heard so far, and then we'll move on to the next presentation. Thank you.

(Whereupon, a recess was taken.)

DR. SEDBERRY: Take your seats, please, so we can reconvene. I see that Rob Andrews is now in the hot seat for the next presentation, but I just wanted to make sure that we didn't have any questions of John before we move on to the next presentation.

DR. CROSSON: My question is you spoke about the Postal Service's list that they keep of households that receive mail, and a lot of the houses that you would be dealing with, say in northern North Carolina, the Outer Banks, are going to be a lot of those summer rentals, and so, presumably, the mail is not going to be delivered to that household, and so how does that go? Would the mail get forwarded then on to the owner of the house, even if they're out-of-state or they're somewhere up in the mountains? How would that work, and how would that affect the survey estimates?

MR. ANDREWS: The sample unit for the survey is the address, a specific address, and so it's a location. It's not the people at the address, and it's not the residents themselves. It's actually the address. I believe, and I will confirm this to see, so I will have a more certain answer tomorrow, but I believe we are excluding seasonal addresses from the sample, and so they're not even eligible to be included, and so they wouldn't even be in the sample frame, but there is a designator in the address sample frame that identifies a house as seasonal or permanent or group quarters or things like that, and so you can include or exclude different units as you see fit, but I will confirm that tomorrow, how exactly we're handling the seasonal units.

MR. FOSTER: Just to add to that, again, I think it's correct that the seasonal are excluded, and so, if those seasonal were being occupied by folks out of state, and they took trips, then they would be covered by the coverage adjustment coming from the intercept survey for being out-of-state residents and taking trips.

For residents, if they live far inland and they rent a house down at the coast, then we would still be covering them. They would just receive -- They would be eligible for the survey from their usual residents, and so, if they were contacted -- We're just asking about trips that occur in the state, or that return to the state, and so it wouldn't have to be while you weren't living at home, and there's nothing like that, and so they could report their trips while they were staying at the coast in a rental, but they would just receive the survey at their permanent residence.

DR. DUMAS: I'm still a little bit unclear. When the survey is mailed, is it mailed to the street address, or is it mailed to the owner of the property on the street address?

MR. ANDREWS: The greeting on the letter is -- It's addressed to State Resident, and so, if we're sampling in South Carolina, for example, or North Carolina, it would be addressed to North Carolina Resident.

DR. DUMAS: My understanding is these license databases have, a lot of times, two addresses, and there can be -- I'm sorry. The Postal Service, there can be an address -- There is the property address, and then there's the address of the owner, in the cases where the owner doesn't live at the property, for example the property is a rental, and so, when the survey is mailed, is the survey mailed to the property address, or is it mailed to the address of the owner of the property? That's the first question, and then I've got a follow-up question.

MR. ANDREWS: The sample frame is a list of property addresses, and so it's mailed to a property address, and, if it's a seasonal property, it's excluded from the sample frame, and it's not eligible to be included in the sample.

DR. DUMAS: The variable that identifies a seasonal address, is that defined as the seasonal address, an address where the owner lives there only part of the year, or is a seasonal address an address where the property is rented out to other people?

MR. ANDREWS: I don't actually know the answer to that off the top of my head. I can look into that this evening, but I think -- Well, my experience anyhow, renting properties, is that they don't have mail delivered to them.

DR. DUMAS: If the survey is delivered to the address, then whoever is living in that property at the moment is going to receive that mail, and so that could be the owner, if the owner is living there at the moment, or it could be someone who is renting the property who is just temporarily staying there, and so then it might be kind of important -- Now it becomes important what the mailing label actually says, and so does the mailing label -- Is it addressed to the owner, the owner's name, or is it addressed to the current resident, or is it addressed to the current occupant, because, if I'm renting a beach house, and say I'm not the owner, and I'm just renting the beach house for a week or two, and I'm there, and I get the survey in the mail, if it's addressed to the owner, I'm not going to open it.

If it's addressed to current resident, I still might not open it, because I think, by current resident, they mean the owner, and that's not me. If it's addressed to occupant, then I might open it, and I don't know, and this could be important in some areas where a large percentage of the housing is rental housing, like a lot of the Outer Banks and Wrightsville Beach, and maybe areas in Florida, where there are a lot of condos and houses that are rented by out-of-state people, and I don't know.

This also becomes important if the out-of-state people are not picked up in the -- If they fish from shore, or they fish from a private vessel that is kept at a private dock that is attached to that property, then they would not be picked up in the APAIS survey that is surveying at public access points, but they would also not be picked up the FES survey that is surveying state residents, and so it's going to the address, but they're an out-of-state person who is not answering that survey. In places where lots of people are renting or seasonally visiting, there could be a lot of people whose fishing effort is not being picked up, I think, and is that correct?

MR. ANDREWS: I will start with the mailing itself. It is addressed to State Resident, and so North Carolina Resident. That's what is on the envelope, and that has the property address. I believe that the rental properties, the beach houses, are not included on the sample frame, and I'm going to try and confirm that this evening, and so they are not eligible for sampling. I don't, at least for short-term, like weekly rentals, I don't believe that people are generally checking the mailbox, and I don't even know if -- I don't think mail is getting delivered to those places, and I just don't think that happens, but, again, I will confirm that. Your second question, if somebody has a boat, they would not be getting intercepted at a boat dock on a private residence, that is correct, and so all of those corrections come from the public access points.

MR. FOSTER: To that, could those people be somehow surveyed, and, again, assuming that the FES does not send questionnaires to the seasonal, which I think is correct too, could they actually

be surveyed, no, because they're fishing from private access, and it's not on the APAIS frame, and they are staying at an address that is likely not on the FES frame, but they would still, and this was sort of to Chris's point earlier, they would still be covered by the adjustment factor coming from APAIS, in terms of their effort, but they are being represented -- The private access trips are being represented by the rate that's coming from the public access, where we can survey with APAIS.

It's possible that that rate does not exactly match what we would have from the private access, but, again, the public access generally has a very high fraction of in-state trips to begin with, the large majority, very large majority, and I think it's fair to say that, in the private access, with the exception potentially of some of these areas with seasonal rentals, is likely very high in-state resident, and so the bottom line there is that the effort is being represented in the total effort estimate, but there could be some accuracy issues there, but, again, there is a bound on the bias, because the proportion from the public access is already -- It has a very high, large majority in-state resident proportion. I rambled on, and that might not have been clear, but I'm happy to take follow-ups.

MR. ANDREWS: If they are a state resident, and so say, for example, somebody from Raleigh was vacationing in the Outer Banks, their address in Raleigh, their home address, would be eligible for the survey, and so, if they get the survey delivered to them at their home address, and they took trips while they were in the Outer Banks, we would expect them to record their trips, those trips, in their survey for that wave.

DR. DUMAS: Right, but the out-of-state people who rented their house during the time when they were not there would not be included.

MR. ANDREWS: If they were out-of-state residents, that is correct. We would not ask them about -- For example, somebody from Maryland who went to the Outer Banks, we could sample them from Maryland, but we would only ask them about their trips in Maryland, and we wouldn't ask them about their trips in North Carolina.

DR. DUMAS: The APAIS survey, that surveys everyone, both in-state residents and out-of-state residents who are using that public access where the survey occurs?

MR. ANDREWS: That's correct. The APAIS samples all trips that are occurring at a site, regardless of where people live, and, in fact, that's how we get the out-of-state correction factor.

DR. DUMAS: That could be one minor issue, and so, even if out-of-state people are under-sampled, then that still doesn't address the issue of why might some of the FES estimates of effort be so much larger than the MRIP estimates before, and so that still doesn't address the problem at all, and so you've still got that problem.

MR. ANDREWS: Or we have a much better estimate now. I mean, we don't really see that as a problem. We see that as a benefit, but, if we were under-sampling out-of-state residents in the APAIS, the estimates would be -- We would be underestimating fishing effort.

DR. DUMAS: Right, and I don't mean that because estimates are large that it's a problem, but I mean the problem of understanding why the estimates are different between FES and --

MR. ANDREWS: Hopefully we'll clear some of that up not in this presentation, but in the next one.

MS. SAULS: I am still having trouble understanding, for Florida, where people respond to the mail survey and they just say that they took a trip anywhere in Florida, and then you calculate an out-of-state adjustment, based on the statewide APAIS survey, and so it's one proportion that gets applied to the total effort estimate for Florida, and it's weighted by the APAIS survey, but that's one proportion, but then the next calculation is the area fished, and at what point -- How are you splitting out Atlantic and Gulf?

MR. FOSTER: I didn't go into it, because it's very similar to the other estimates, but, specifically for Florida, for both shore and private boat, there are the proportions of total trips for the state by coast, weighted proportions of total trips for the state by coast, are calculated from the APAIS as well, and so there's another proportion just specific to Florida coming from the APAIS that is used to make that partition, and so you could look at it as either -- I would have to go back and double check, and I think we do it in two separate steps, but you could either think of it as taking the area fished proportions and crossing it by coast, and so, instead of having just three proportions, you've got six, and you've got west coast inland, state, EEZ, and you've got east coast inland, state, EEZ, and you got those six proportions, and you take the total for the state, and then you partition it out into those six cells by those proportions, which I actually think is how we do it, or you could think of it the other way, as you've just got the east/west split, and do that one, and you take the east and the west and you split that out again into the three areas, but I think we do it all in one step with the six cells, and so it is still using the APAIS data to make that split.

DR. REICHERT: Is recall bias at all part of the consideration, and at what step is that being considered in FES?

MR. ANDREWS: Recall error is certainly something that we're concerned about, and it's something we have tried to address that I will talk about in this presentation.

DR. REICHERT: Thanks.

DR. LI: I have three questions and one comment. The comment is I do not see FES estimates larger than the telephone survey as a problem. However, I feel that what is a problem is that FES estimates so far does not match the experience and observations from the states, the coastal states. Then the three questions are related to the post-stratification adjustment in the calculation, and I am wondering what kind of -- You are saying there is adjustments to match the demographic features in the United States census, and so what kind of features are you matching? Like, for example, is it the proportion of households with more than one kid?

MR. FOSTER: I will let Rob speak to the details.

DR. LI: I am thinking that some features may be closely related to fishing activity, but some features may be not, and so I am wondering what kind of features are you trying to match?

MR. FOSTER: Right. I will let Rob speak to that.

MR. ANDREWS: The source for the control values is actually the American Community Survey, and it's not the decennial census, and so they're updated regularly, annually, and so we have annual control totals, and the ones specifically that we were looking at are the percent of households with kids, and so under eighteen, the percent of households with seniors sixty-five or older, the type of telephone service, and so whether they have a landline telephone or not, and that's actually from -- That is not from the American Community Survey. That's from the National Health Interview Survey, which I will actually talk about in a couple of presentations.

Then the other one is whether they rent or own their house, and we chose those ones because they are all predictors of fishing activity, and so we see differences between people who are different with respect to those characteristics. We are attempting -- By making these adjustments, we are making our sample representative with respect to those demographic characteristics.

DR. LI: I appreciate that. That is great. That is what I am looking for, because I am worried that you are matching with some factors that are not related to fishing activity at all. Then my second question is so you are matching with -- Like every year's estimate you will match with that year's census, right?

MR. ANDREWS: That's correct, yes, and so every wave of 2017 we matched to the 2017 American Community Survey.

DR. LI: Okay. That's great, and the last question is, if you have multiple factors to match, do then you just multiply those weights together? Like, for example, if you have like three factors and you come up with three weights for each factor and then you multiply them together as a total weight for the post-stratification weights?

MR. ANDREWS: When you refer to factors, no, it's actually -- It's an iterative process, and we actually only make the adjustments at the margins, and so, for example, if you have two factors, rent versus own and percent of households with seniors, that would give you four boxes, right, and we don't have information that would allow us to make those adjustments for every box, and so we make them at the margins, and so we would do it for one factor and then the other factor and then the other factor, and then you take I think the average distance, and it's an iterative process, where it goes through at multiple iterations until the average distance from a control value for all of them sort of a converges to a minimal value.

DR. LI: Okay. That's great. Thank you.

MR. FOSTER: I was just going to say that it's a standard technique known as raking, or iterative proportional fitting, and, again, essentially, the ratio there that I showed -- At each step, you would recalculate the new estimate and calculate that ratio then with the control total, and so, again, at each step, you are adjusting the sample weight, and, again, at some point in the iterative process, again assuming that there's enough sample size and things like that, the margins will all converge, so that it doesn't have to keep making the iterative adjustments.

DR. SEDBERRY: Okay. I think we are ready to move on to Rob's presentation. I just wanted to remind the SSC note takers for this session, which is the introductory presentations, that our note takers are Rob, Scott, Luiz, and Carolyn, and so carry on until we get to the next section. With that, I think --

DR. BARBIERI: Mr. Chairman, may I ask one quick question before we move on to the next one?

DR. SEDBERRY: Yes.

DR. BARBIERI: Thank you. John, thank you for this very good presentation and all the time explaining a bunch of things and answering a bunch of questions. It's very helpful, but, in addressing some of these last points, you mentioned sample size. Can you give me an idea, give us an idea, of how you guys estimate minimum sample size? Is there anything that is done in this process to come up with an estimate of what would be required to maintain the levels of precision that you would like to have in terms of the estimates?

MR. FOSTER: Luiz, thanks very much for that, and, before we -- I will probably ask Rob to answer most of this, but I just wanted to make sure. Are you speaking specifically about FES sample sizes or also APAIS sample sizes?

DR. BARBIERI: I am thinking about both, just because this is something that we're going to be talking about, not necessarily right now during this presentation or this specific topic, but one of the issues that we're going to be talking about is how the estimates may be impacted by sometimes the sample size, the different stocks, depending on how often they show up in the sampling, and how that could be impacting what goes into the stock assessment. I am just trying to get, at his time right now, because we're talking about these issues, to get to those points, and so both the APAIS and the FES would be helpful.

MR. FOSTER: Okay. Thanks, Luiz. I think it would probably be simpler to talk about the FES sample sizes first. Rob, do you want to go through that?

MR. ANDREWS: Sure. We generally use -- Well, not generally. From the historical data, we have -- We are able to measure sort of, or estimate, the variability within the population, and so whatever we're measuring. We use historical data to look at that variability, and then it's just sort of a standard formula that you plug that variability into to estimate sample sizes required for various levels of precision.

Right now, at the sample sizes that we use for the FES, we're able to achieve relative standard error of approximately 20 percent. Now, whether that's a target or whether that's what we can afford, that is kind of debatable. I mean, we sample at a level that's really practical and affordable for us to sample that, but I think, really, to answer your question, because we have historical data, we're able to just plug those into a standard formula and come up with an estimate of a sample size requirements to come up with certain levels of precision, assuming that variability is pretty constant from year-to-year.

MR. FOSTER: Before speaking a bit about the APAIS, just one other -- A couple of other things that we do look at. As Rob described, the Neyman allocation is a primary method used to allocate the FES sample, but we do have sort of two other considerations that we look at. One is there is a fairly standard calculation that can be done to look at sort of estimating, for a given sample size and a given estimate of prevalence, and so, in this case, what is the proportion of households that will report fishing in a state and wave combination, let's say, and so, for a given sample size and

for a given fishing household prevalence, you can calculate the probabilities of -- You can calculate the probability of obtaining at least one report of fishing.

This is sort of kind of a detection limit sort of calculation, kind of, essentially, what's the probability -- Given that there is fishing, what's the probability that we can detect it at various sample sizes, and so we try to -- In addition to just doing the Neyman allocation, sort of looking at minimum sample sizes, or lower bounds, and we take into account what's the probability of detecting fishing, given that fishing is occurring, and make sure that we have, again, sample sizes that give us a very reasonable probability of detecting fishing, and so like 80, 90, 95 percent, something like that.

Another consideration is the sample weights, and so, because these are simple random samples within the strata, we can estimate what the sample weights are going to be, or we can calculate them, essentially, before any adjustments for non-response and the post-stratification raking and that sort of thing, and so, if we have cells where, with a given sample size allocation, we're going to see extreme weights, we also try to make adjustments, as we can, to moderate those weights some, and so, again, it's trying to give us some mitigation or reduction in the potential for sort of an outlier situation.

Now, again, we're limited in our abilities to make those kinds of adjustments, because we don't want to deviate too far away from the Neyman allocation, because that is standard practice for minimizing the variances for a fixed sample size, but those are two additional considerations that, if we have the ability to make some adjustments in the sample sizes, we do that. Again, improve our chances of detecting fishing, given that fishing has happened, and minimizing or moderating extremely large weights, sample weights, to minimize the potential for outliers, and that also helps with variance as well, to try to minimize differences between the weights. That's all I had to add for FES.

On the APAIS side, of course, that's sort of a more difficult issue, because, given the general nature of the survey, it's difficult to -- We can't optimize for everything, because trying to optimize for some small number of species is likely to -- That would improve the precision -- Shifting sample, coming up with sample allocations that would improve the estimates for certain species would likely degrade the estimates in terms of precision for other species, and so inshore versus offshore, pelagics versus bottom fish, potentially.

To the extent that these trips are spatially separated, either because of site or distance from the coast or nearness to an inlet or navigable water or have you, we're likely to -- We can't optimize for everything, and so that makes it difficult to come up with sort of a best sample allocation, given limited resources and the inability to just increase sample sizes across the board.

Now, what we have tried to do though is find ways to be more efficient with the existing sample, and so, for example, in a couple of states, Florida, as well as Alabama in the Gulf, we have created an additional site group stratum, the offshore site group, and these are sites identified primarily by the state partners that will reliably produce or have a reasonable probability of producing trips targeting offshore species, whatever those species -- Priority offshore and not everything, usually, but key species in a given region of the state or across the entire coast of the state.

By creating that additional stratum, we're able to take existing sample and target it directly into those sites, in a more focused way, and that increases then the likelihood of having those trips with those particular characteristics in the data, but, again, not biasing it towards those trips. It is being reflected in the sample weighting, so that those trips are still being represented, but we're just getting more observations of them in the data, but, again, with a fixed budget, we're limited in how much we can tailor sampling to do that, because the samples going into those strata have to come from somewhere, and they're coming from -- It's essentially taking a pie the same size and cutting some additional slices out of it, and so it's not a cure-all for everything. At some point, either we will have to continue shifting sample from other strata, to the degradation of those strata, or we would have to have some sort of targeted increases in funding to increase those sample sizes.

That's all sample size, and the one other thing that I would mention, and, actually, Dave, I don't know if you would like to speak a bit about the rare-event project, but I was just going to say that there are ways that -- We have a project currently ongoing to look at are there better ways we can use the existing data, ideally fully design-based, without model assumptions, but perhaps, or likely, with some approaches that do make some fairly simple model assumptions that could improve the estimates for some species that, really, we just may not be able, feasibly able, to improve the precision on much, from an increasing sample size consideration. They're either just too rare or they're just not -- They are not concentrated enough in a predictable way in specific sites or regions of the state or time of year, things like that. I don't know if that specifically addressed what you were looking for, Luiz, but I'm happy to answer another follow-up if you have any.

DR. BARBIERI: Yes, that's exactly what I was looking for, John. Thank you so much, and that covers the range of topics that eventually we'll be discussing, and so I wanted to get that part taken care of. Thank you.

DR. SEDBERRY: Thanks, Luiz.

DR. SERCHUK: I wanted to follow-up a little bit on Luiz's intervention, but in a different way, and that is how do you handle situations where the sampling is interrupted because of an external event, either a weather event, like a hurricane, or a decision not to sample, because there is red tide there and last week nobody was fishing, and so we're going to distribute our effort, our APAIS effort, differently, and those sort of situations that produce either low sample sizes or estimates that may not be reflective of conditions, and I'm just wondering, have you come across that in the past, and, if so, how have you handled it?

MR. FOSTER: We have had situations like that that have come up, and I think it's fair to say that they are usually handled on sort of a case-by-case basis, depending on the extent of the disruption, and so, for example, in the case of Puerto Rico, with the 2017 hurricanes, that led to the sort of complete suspension of sampling for both the APAIS intercept survey as well as the FES, and we are still working out how best to resume sampling, because the mail service, the infrastructure, was so badly disrupted from the storm that it's really, to this point, still not quite to a place where we can implement the FES reliably.

Similarly, with the APAIS, there was a lot of sort of destruction, catastrophic destruction, in some cases, to select fishing access sites, and while, of course, sites have been restored and new sites have been established and things like that, it's just -- An event that's that disruptive takes a long time to kind of get things back online, but there are -- That's an extreme example, kind of at one

end. We do have other examples. Rob, do you want to talk about sort of the shifting in the FES timing? Didn't we do that once for shifting mailing schedules?

MR. ANDREWS: We did. I actually don't recall if it was last year or the year before, when there was some bad storms in the Southeast, where we just basically made the -- It interrupted mail for several days, and mail wasn't being delivered, and so we, obviously, couldn't deliver the survey, and so we made a decision to, I believe, delay the mailing to I think it was North Carolina, South Carolina, and Georgia, perhaps, and so we delayed the mailings by one week.

It's a very ad hoc process, and it's kind of hard to develop standard protocols to address these types of things, because the magnitude and the duration for which infrastructure might be impacted by these events is kind of hard to predict, and so we really kind of handle them on a case-by-case basis and try to utilize as much local information as we can to make decisions about when to resume normal data collection, and, conveniently, I suppose, the data collection contractor for the FES is located in Research Triangle Park, North Carolina, and so they were sort of right in the middle of the storm, and so they had a pretty good idea of that things were like, in terms of mail delivery, and so that was sort of convenient, I suppose, in that case, but I think the bottom line is we just sort of handle it on a case-by-case basis and balance trying to maintain normal operations with, for one thing, the ability to deliver mail and collect data and also having some sensitivity to the fact that people's lives are impacted by these events, and we don't necessarily want to start asking them about their hobbies and their recreational pastimes when they're dealing with much bigger issues.

MR. SERCHUK: We had a tremendously bad red tide season about a year ago, certainly on the west coast of Florida, and some of it went around to the east coast of Florida, and, presumably, that interrupted fishing activities, to a large extent. Would you still maintain the same intercept surveys, or would you cut back, because you recognize that fishing had really diminished, and how would you make those decisions?

MR. FOSTER: Fred, we would be relying on our state partners and their input, in particular their input into the fishing pressures in the site register and as to whether they feel that events have reduced the pressure at certain sites, whether that pressure is just eliminated, or whether it's moving to other parts of the coast, and so the state agency partners entering the fishing pressures, updating the fishing pressures, for the individual sites, even potentially temporarily retiring sites that may be heavily impacted by something, the flexibility is there to do that.

Now, that still though is, essentially, on a month-to-month basis, and so once that -- If something is likely to impact several months of fishing, that can help handle that situation for the out months. For the current month of sampling, we would likely simply suspend or cancel assignments that had already been selected at sites that might have been impacted, again depending on what the nature of the disruption is.

Then, for future months that had not already been drawn, the sample hadn't already been drawn, the pressures at the individual sites can be updated to reflect the conditions, just as they are in general, but the changes might be more extreme than what we would see in a typical update from one month to the next.

DR. SEDBERRY: Wilson, I saw you had your hand raised, but I'm really getting anxious about our timing here. If we can -- If it's something that the details might be provided in subsequent presentations, I'm asking if you can wait, but, if it's relevant to this, go ahead.

DR. LANEY: It is relative to this, Mr. Chairman, and it's sort of -- I think Fred's questions about hurricanes and red tide touched on it, but I was going to ask Rob, in particular -- Rob, you said that you assumed variability is constant from year to year, and, given Fred's questions about hurricanes and red tide, those are not particularly uncommon events here in the Southeast, and so my question is whether or not, when an event like that occurs, and John has touched on it, to a certain extent, you couldn't make some sort of an adjustment to account for the fact that you know fishing, as an activity, is going to decrease, particularly severely, during events like that, and so is that something that -- Could you flag years in which you know the variability is not constant? I mean, it's going to clearly be, I presume, a lot greater in a year like I think you're referring to 2018, when we had Hurricanes Florence and Michael in coastal North and South Carolina in the same year.

MR. ANDREWS: The question about using historical data, that was strictly to come up with our sample sizes and our allocations, and it was not producing estimates. It was using historical data to try and determine or estimate the best way to allocate our sample among strata, and so among coastal and non-coastal strata in the license match and the unmatched strata, to come up with the most precise estimate.

In the current process, we're using I believe four years' worth of data. We're using all the data that we have collected since the FES has been implemented to come up with that allocation, and it makes the assumption that there is some baseline normality, I suppose, in fishing activity, but it's just to come up with our sampling allocation and estimate sampling requirements to achieve some precision goal.

It has nothing to do with the actual estimation, and so it's the only information really we have, and it's the information we have available to us to come up with that sampling allocation, and the sampling allocation is an estimate of how to distribute the sample to come up with the most precise estimate, and so, if we deviate from that, or if we don't get it exactly right, or if fishing activity changes from year to year, our sampling allocation is not going to result in a biased estimate or an inaccurate -- It's not going to result in a biased estimate. It might be a less-precise estimate than the optimum allocation would have given us.

I think using multiple years' worth of data to come up with a sampling allocation, in an effort to sort of predict what fishing activity will be like the next year for allocation purposes, is probably a more accurate approach than like trying to make some adjustment. You can't predict when the hurricanes are going to come, and so use the historical information, because, over a long time series, that's going to be a pretty good predictor of fishing activity, but, again, this is all for the sampling allocation exercise, and it's not impacting the estimates at all.

MR. FOSTER: Right, and, along those lines then, we would expect, with disruptive weather events, or any sort of disruptive event that would negatively impact fishing activity, we would expect to see that, in terms of the data, in terms of the numbers of trips that people report, and so they should be -- We would expect them to be reporting fewer trips, or more households reporting no trips, as the result of the disruptive event.

DR. SEDBERRY: Did you have a follow-on to that?

DR. LANEY: Yes, and just one quick follow-up, John. I guess part of my question was motivated by a lot of the concerns that I have heard expressed by recreational and commercial colleagues who have said that, during certain periods of the year, some of the estimates, especially for the shore mode, seemed extremely high to them, based on some of these weather events, which they felt should have produced estimates that were much lower than what would up coming out of the system.

MR. FOSTER: I think I would just ask if we could hold that discussion until Rob gives the presentation that walks through the differences between the CHTS and the FES, in terms of the estimates.

DR. SEDBERRY: Okay. Rob, I think we're ready if you are.

MR. ANDREWS: All right. Thank you. Between John and me, we only have four more presentations of background to get through, and so I think the take-home message is be careful what you ask for, because I don't think any of us really like to talk that much, but we can, and we can talk a lot.

The next two presentations, I'm going to be talking specifically about the Fishing Effort Survey, and I promise the next presentation is going to talk about biases and differences between the mail survey and the telephone survey estimates, but, before I do that, this presentation is going to focus on the development of the survey, and so the process we went through to design the survey, and the transition to the FES from the Coastal Household Telephone Survey.

The FES was the result of extensive testing and analysis, and it was a multiyear process, and it included several pilot studies and peer reviews, and we have executed a pretty extensive transition process to facilitate moving from the Coastal Household Telephone Survey to the FES. As you all know, prior to the FES, fishing effort was estimated by the Coastal Household Telephone Survey, or CHTS, and the CHTS was implemented in 1979, and it's a list-assisted random-digit-dial telephone survey that is administered in coastal states.

For many years, random-digit dialing was a pretty standard approach for sampling households and collecting information, and it was used by many federal statistical agencies and universities to collect information. Over the past fifteen years, however, landline RDD surveys have become increasingly inefficient for sampling households, and this is primarily due to changes in telephone usage.

John mentioned the National Research Council review in his presentation, and a lot of what I'm going to be talking about John referenced at some level, and so I think I can get through most of this pretty quickly. That 2006 NRC review highlighted the limitations of the Coastal Household Telephone Survey and made several recommendations on how we might improve our data collection.

Specifically, they noted the fact that the survey was inefficient, and John mentioned this. Saltwater fishing is a rare activity, and finding anglers from the general population is challenging. The

response rates have decreased over the last several years, which increases the risk for non-response bias, and then the CHTS is susceptible to coverage bias, due to the limitation to landline telephone numbers, and this is a growing problem, due to the proliferation of cellphones, and this phenomenon is not limited to the Coastal Household Telephone Survey. It's a nationwide phenomenon that impacted every landline random-digit-dialing survey that's been administered, and so it's not just us that had some of these problems.

To address the limitations, the National Research Council recommended that we develop a comprehensive universal sampling frame of licensed saltwater anglers, and, in the absence of such a comprehensive sample frame, that we develop dual-frame procedures that included sampling from license databases, and this is primarily because sampling from saltwater fishing licenses would be more efficient than sampling from just the household population.

About the same time, or shortly after, the National Research Council review was completed, and the Magnuson-Stevens Act was reauthorized, which mandated the establishment of a national saltwater angler registry as well as the development of an improved recreational fisheries data collection program.

Given this direction, NOAA Fisheries established MRIP and initiated a series of pilot studies to identify a more accurate and efficient way to estimate fishing effort, and the goals of this initiative were to design an improved method for estimating fishing effort, attempt to maximize efficiency by incorporating fishing license information, and, specifically, working with the states to collect that license information, maximize response rates, and so we minimize the potential for non-response bias, and then evaluate non-sampling errors and bias. All of this work, which happened over several years, culminated in the FES design that we currently administer.

Initially, given the direction from the National Research Council and the reauthorized Magnuson Act, we focused on angler license surveys. In 2007, we developed and tested the Angler License Directory Telephone Survey, and we actually started planning for this before the NRC report came out. This methodology was tested in the Gulf states, as well as North Carolina, and it sampled directly from a list of licensed saltwater anglers.

Otherwise, the survey was pretty much identical to the Coastal Household Telephone Survey. We sampled a licensed angler, and we asked for that angler by name when they picked up the telephone, and then we, essentially, administered the CHTS questionnaire. We asked screening questions, to try and determine if anyone in the household fished, and then we collected effort information from anybody in the household who fished and not just the licensed angler.

This methodology was much more efficient than the Coastal Household Telephone Survey, and a much higher percentage of households reported fishing than we would see in the CHTS, and so estimated prevalence was much higher, but we knew there were gaps in coverage of the sample frame, and so pretty much, at that time, all the states had exemptions to their license coverage, and so kids under sixteen are not required to have a fishing license, and I believe that's the case currently.

In many cases, seniors were exempted, or there were exemptions for folks with disabilities or folks in the military, and a lot of people just fish without a license, and we believe this is something that continues to this day, and so we knew there were gaps in coverage that would need to be addressed

in any future data collection program. Then, finally, there were minimal improvements over the CHTS, in terms of response rates. This was also a telephone survey, and response rates to telephone surveys have just been plummeting, and that was true in this design as well.

To address the concerns with coverage, we -- Because the Coastal Household Telephone Survey was ongoing during this time, we attempted to integrate the license frame survey, the ALDS, with the CHTS in a dual-frame design, and, when you look at the schematic, there is two sample frames, the CHTS sample frame, and that's the circle on the left, and then the ALDS sample frame, which is the circle on the right.

The union of the two frames defines three different populations. There is people who could only be sampled by the Coastal Household Telephone Survey, and so these are people who lived in coastal counties, but didn't have a fishing license, or for whom we did not have a telephone number, and there were people who could only be sampled from the Angler License Directory Telephone Survey, and so this could be people who didn't have a landline telephone who lived in a coastal state, and it could be people who lived in non-coastal counties that weren't covered by the CHTS.

Then there is this portion of the population that could be sampled from both sample frames, and so these are people with a landline telephone who lived in coastal counties, and then there's actually a fourth domain that is not shown on here, and it's people who couldn't be sampled from either of these frames, and so people without a license and without a landline telephone number, and we still didn't have a way to contact those people.

In the dual-frame design, you really need to know who is in which of these domains, and so you need to know who is in the overlap, because they have a higher selection probability than people who could only be sampled from one frame, and you need to know those selection probabilities, so you can estimate the sample weights to produce your estimates.

The way we did that, in this dual-frame telephone survey, is we asked CHTS respondents if anyone in the household had a fishing license, and we had some real concerns about that, because, when you're asking people about fishing, and then you ask them if they have a fishing license, you're probably not going to get the most accurate information if they don't have a fishing license, and we knew that this design had greater coverage than either the ALDS or the CHTS, and it was two sample frames, and so it had to have greater coverage than either of the individual sample frames.

There were still some gaps though, and those gaps were growing, because the coverage of the CHTS was declining, and we still feel like there is people, even now, who are fishing without a fishing license. As I noted, the approach for defining the dual-frame domains is unreliable, and we didn't think that people were going to provide us with accurate information about whether or not they had a fishing license, and then both of the surveys achieved pretty low response rates, and so we didn't really think that telephone surveys were a long-term solution to some of the problems that we were experiencing with the CHTS.

Given these growing challenges, we began to consider mail survey designs, and, during this period, address frames became widely available from commercial vendors, survey companies or just sample providers, and address-based sampling became a cost-effective alternative to random-digit dialing.

There is many benefits of address-based sampling. As John mentioned, these sample frames are nearly 100 percent complete, which is great. They are also regularly updated by the Postal Service, and so these vendors could send their list of addresses out to the Postal Service, and the Postal Service would add new addresses, and they would correct errors in the databases, and this happens as frequently, I believe, as every two weeks. Beyond that, it's monthly or every two months, depending on how much the vendor is willing to pay, essentially.

They also provide a tremendous amount of flexibility, and so we went with the mail survey approach, but you're not limited to a mail survey approach with address-based sampling. You can actually select addresses and do face-to-face interviewing, and you can select addresses and send a letter inviting people to participate in a web survey, and you can even try and do telephone surveys with address-based sampling, by matching addresses to commercial lists of telephone numbers, and so they provide a tremendous amount of flexibility, and they also provide -- They are a rich source of supplemental information, and so they don't just provide you with an address. Many of these commercial vendors are appending information to the address frame, and so, at a minimum, you generally get -- You will get census information appended to your sample frame, and so you will know the census block or the census tract or whatever the different geographic categories that the census breaks their information into.

You can also get economic information added to it, and you can get race information appended to it, and you can get income appended to it, and, the more detailed information you get, the less accurate that information becomes, but it's a very rich source of supplemental information, and you can use that information for stratification or using different incentives to try and encourage response.

In 2009, we designed and tested a dual-frame mail survey, and this was initially tested in North Carolina in 2009, and then we expanded it to Louisiana in 2010. It was an overlapping dual-frame design, and so we sampled directly from the state license databases, and we conducted independent sampling from the address frames, and then we matched the samples by address, to try and identify the overlap, and so we were no longer relying upon angler responses to questions about whether or not they had a license. We were actually doing address matching to try and define the different domains defined by the union of the two sample frames.

The survey was a self-administered mail survey, and the license sampling was conducted in a single phase, and so we selected addresses, or selected license holders, from the sample frame, and we sent a fishing questionnaire to them and asked them to report the number of saltwater trips they took during the prior two months.

The address-based sampling was conducted in two phases, and so we sampled addresses, and we sent a screening questionnaire, to try and identify anglers, and then, once we had identified anglers, we sent a follow-up questionnaire, which was identical to the license survey questionnaire. It just asked about the fishing activity during the wave.

Because it included an addressed-based sample, it provided nearly complete coverage of the population, and so that was a bonus, and, somewhat surprisingly, it got much higher response rates than either the CHTS or the ALDS. In many cases, it was two to three-times higher than the

telephone survey response rates, and so we were kind of surprised that they were as high as they were.

There were a lot of benefits to this design, but there were still some challenges, and so, for one thing, while we thought the approach for defining the different domains, defining determining the selection probabilities was more effective than asking anglers if they were licensed, the address matching was still not 100 percent accurate. There were some bad addresses on the license frame, for example, and so it was still a little prone to error, which could result in biased estimates if those errors were large, because we couldn't accurately determine selection probabilities.

Then the two-phase design from the address-based sample was also costly and time consuming, and so one of the questions earlier about the FES design was what the mailing schedule was, and so, with the FES, the current survey, we send a single-phase questionnaire out about a week before the end of the wave. Because we were doing a screening survey first, we couldn't do it that close to the end of the wave. We had to do it several weeks ahead of time, because we first had to identify anglers before sending the follow-up survey, and so it took a while, and it was costly, because we had to do a very large mailing to come up with sample sizes large enough to do the fishing survey component of it.

One other point I want to note about many of these pilot studies is we had embedded experiments, and so we were not just testing the overall design. In that particular one, the two-phase mail survey design, when we expanded it to Louisiana, we included some studies to look at, to compare, first-class mailing to priority mail, and we looked at different incentive levels, to see what the impact of that was on response rates and cost, and we tested adding a fourth mailing, and that was pretty common throughout this whole testing period. There were several embedded experiments.

One of the things I mentioned was that the address-based sampling provides a lot of flexibility, and so one of the things that we tested in 2012 was a dual-frame mixed-mode survey, and this wasn't really a -- A mixed mode survey traditionally offers people either a choice of how they want to respond, and you could either ask them to do a telephone survey or a mail survey or a mail survey or a web survey, and this wasn't really a traditional mixed-mode survey. It was actually a test that compared mail data collection and telephone data collection.

The sampling design for this was identical to the dual-frame mail survey, and the address-based screening was completed via mail, but the topical sample, the actual questionnaire where we asked about fishing activity, we allocated into mail and phone treatment, and so half the sample went to the mail -- Once we identified anglers, either sampling directly from the license frame or through a screening-address-based sample, we allocated that sample into telephone and mail treatments, and this provided us a direct comparison between the telephone and mail for collecting the fishing data.

The results continued to demonstrate the feasibility of the mail survey design. The mail survey had a much higher response rate than the telephone survey. It was, I believe, eight to nine percentage points higher than -- The mail was eight to nine percentage points higher than the telephone survey, and the mail survey response rates eclipsed the telephone survey response rates after about three weeks of data collection, and so I think people tend to think that mail surveys take a lot of time to administer, but we got response rates that started exceeding telephone survey

response rates just after three weeks, and so we knew we could get data back pretty quickly. People who respond generally respond pretty quickly.

This demonstrated that we could produce estimates, stable estimates, within the current estimation schedule, the estimation schedule of the Coastal Household Telephone Survey, and we also found that there were pretty minimal differences between the mail and the phone estimates, particularly for the address-based sample, and, ultimately, we concluded that mail surveys are a feasible alternative to the Coastal Household Telephone Survey.

Then, finally, we developed the revised design that addressed some of the weaknesses identified in the prior pilot studies. This, ultimately, was the FES design, and so the big difference between the FES design and the dual-frame mail surveys is that we changed how the license databases were used. In the dual-frame mail survey, we sampled directly from the license databases and then tried to identify that overlap with the address-based sample. In the FES design, we used the license databases to stratify the address sample, and John talked a little bit about this, and we matched the license database to the address frame, which defines license matched and license unmatched strata, and then we sampled directly from those strata independently.

We also collect data in a single phase, and so the prior versions of the mail survey asked a little more detailed information about specific fishing trips. The FES is a very simple questionnaire. We just ask people how many times, how many days, they went shore fishing during the prior two months and how many days they went boat fishing during the prior two months, and so it's a very simple questionnaire, which allowed us to collect the data in a single phase. We didn't have to do a screener questionnaire first, followed up by a fishing survey. As I noted, we tested incentives in multiple versions of the questionnaire, and this pilot study, again, demonstrated the benefits and the feasibility of the FES design.

Before I move on to the sort of final recommendations in the review process for the FES, I am going to talk specifically about non-response. The pilot study demonstrated that mail surveys result in much higher response rates than the Coastal Household Telephone Survey, and so there was a lower risk for non-response bias in the mail survey than the Coastal Household Telephone Survey, but we still had concerns about non-response bias, and, specifically, we were concerned that avid anglers, or people who fished, were more likely to respond to the survey than people who don't fish, and prefer refer to this as avidity bias, and so avid anglers are more likely to respond than non-anglers, or less-avid anglers.

For the FES, we evaluated and addressed non-response in all stages of the survey estimation, and so from survey planning and design through weighting and estimation, some of the things that John talked about, and, finally, through administration of pilot studies to specifically try and evaluate non-response bias.

For example, the FES was specifically designed to appeal to a broad audience, and we included a two-dollar cash incentive. The survey is not branded as a fishing survey, and so it's a household survey, and we want fishing households and non-fishing households to respond, and so we don't just ask about fishing questions, or fishing activity. We ask questions about how people access the weather, and they're all very NOAA-centric questions, because we are NOAA Fisheries, and we didn't feel like we could ask questions about grocery shopping or visits to the dentist or things like that, and so we just wanted to ask some -- It's a weather and outdoor activities survey, and so

we ask some questions about the weather, and we ask about visiting public beaches, and we ask about freshwater fishing.

When we were testing this, we compared this weather and outdoor activities survey to a fishing-specific instrument, and what we found is that the more general survey, the weather and outdoor activities survey, resulted in higher response rates from both anglers and non-anglers than an actual fishing survey.

We still know, even though we make the survey appeal to a broad audience, we still know that households with licensed anglers are more likely to respond to the survey, and so our matched stratum has a higher response rates than our unmatched strata, and they're also more likely to report fishing activity, and so, if we don't do something about this, anglers will be overrepresented in the same, in the effective sample, people who are responding, and we will overestimate fishing effort, but, as John noted, because we can't identify households with licensed anglers and distinguish those from households without licensed anglers, we're able to make weighting adjustments to our data, to ensure that the samples are representative of all households with respect to license status, and, as John also noted, we perform a similar adjustment with boat registrations, and so we can make sure that our sample is representative, with respect to fishing license information and boat registration, which minimizes bias and also results in more precise estimates.

Then, finally, when we initially tested the FES design in 2012 and 2013, we conducted a non-response follow-up study, and this is a pretty simple, straightforward study. Every wave that we administered the survey, we took a sub-sample of 320 non-responding households, and we sent another -- We made another mailing that we sent via FedEx, and we included an additional five-dollar cash incentive, and so we became a little more aggressive in trying to solicit a response.

What we found is that the people who responded to the non-response study were not significantly different from the initial FES respondents, and so this is a pretty good indication that the non-response weighting adjustments that we use are effective and that non-response is not a significant source of bias in the FES. The response rate we got for the non-response follow-up study, sending via FedEx and including the five-dollar incentive, was about 40 percent, and so it was a pretty good response rate for a non-response follow-up study.

By the numbers, FES development, the design was administered over the course of eight years, and we conducted pilot studies in nine studies, and there was a total of five different pilot studies, and we conducted a total of three independent reviews of various methods, and one thing I wanted to note, and I mentioned this earlier, is, once we implemented the FES, we didn't say we're done and we're just going to proceed with this method and we're not going to make any changes moving forward. I think we kind of learned, from the Coastal Household Telephone Survey, that not changing methods over the course of thirty years probably wasn't a very good idea.

In addition to all of these pilot studies, we have probably implemented another half-dozen since 2015, when we implemented the FES for benchmarking, and so we're constantly trying to improve the methodology and improve response rates and improve our sampling design and improve the questionnaire to try and solicit some more accurate information.

The final recommendations from all of this work is that the FES design is a feasible alternative to or for collecting recreational fishing data, and the survey results in higher response rates and gives

respondents more time to complete the survey, which we believe leads to more accurate responses to questions about fishing activity. The survey can generate stable estimates within the CHTS estimation schedule, and so, just because it was a mail survey, we weren't going to delay our estimation schedule. Then, because we were incorporating license information in the design, we believe it's more efficient than simple random sampling, and we can sample different strata at different rates, which results in more precise estimates and a more efficient design.

We believe the design is less susceptible to bias, resulting from non-response and non-coverage, and it's not that we believe that. It is less susceptible, because it has much greater coverage than the CHTS, and, as I'm going to show you eventually, that is a pretty big deal, and it had much higher response rates than the CHTS. Address-based sampling provides more representative samples than landline random-digit dialing, and the overall conclusion was that the FES design is a superior approach for monitoring recreational fishing effort.

Don't just take our word for it. The results from these pilot studies have been reviewed by our internal staff, external consultants that we've worked with regularly, independent peer reviewers, and a committee convened by the National Academy of Science. In terms of the peer review process -- By the way, all of this, the report, the peer review comments from the FES, are all available on our website, and I believe there was a direct link to it in some of the materials that were sent out for this workshop.

Staff from NOAA Fisheries Office of Science and Technology provided a substantive review of the design and the results. Three peer reviewers, independent peer reviewers, were selected by the American Statistical Association Survey Research Methods Section to review the design, and then we recruited five members, five external expert consultants, to provide reviews, and so these are people that we have worked with to help design methods for MRIP, but they weren't involved directly in the development of the FES. The reviewers provided comments on the methods, results, and conclusions described in the final project report.

Generally -- Well, the peer reviewers concurred with the project team's overall findings and conclusions, and the project report was subsequently reviewed and approved by the MRIP Operations Team and the Executive Steering Committee, and the design was certified by MRIP as a statistically-valid design.

The design was also reviewed as part of the National Academy of Science's ten-year progress report on MRIP, and the review committee met on four occasions in Washington, D.C.; here in Charleston, South Carolina; New Orleans; and Irvine, California. At each meeting, the committee heard from state and federal employees, as well as regional stakeholders, and then the committee also received written input from stakeholders during the study process, and that review concluded that the methodologies associated with the current FES, including the address-based sampling mail survey design, are major improvements from the original Coastal Household Telephone Survey.

There is a reason that you keep seeing these cover pages from these National Academy reviews. I mean, these are kind of big deals, these reviews. They were expensive, very time-consuming processes, and they were completely independent reviews, and we put a lot of stake in the recommendations that they made, and so John showed pretty much these same recommendations earlier.

In addition to their conclusions about the FES, they made several recommendations on things that we should continue to look at, and so they suggested that we continue to evaluate the cognitive properties of a two-month recall period, and this got to somebody's question earlier about recall error with the two-month recall period, and I will talk a little bit more about that.

They said we should consider evaluating a prospective data collection methodology and continue to research survey panels, and we should consider conducting targeted annual non-response studies. We should evaluate the use of electronic data collection, evaluate variables on the sample frame to support either stratification or variance estimation, and then we should evaluate other variance estimation methods.

I am just going to talk about a couple of these, things that we have worked on or continue to work on. In 2015, we conducted a pilot study to evaluate the relationship between recall error and the length of the recall period, and recall error occurs when respondents are -- When they can't accurately remember what they did, basically, and recall error is generally classified into two different types, telescoping error, which occurs when a respondent misplaces an event in time, usually placing it more recently than it actually occurred, and then omission error, and this is when just people forget events that happened.

This particular study compared the FES, which asks about fishing trips that occurred during the previous two months combined, and so, for Wave 4, for example, we send a questionnaire at the end of August, and we ask about fishing activity that occurred in July and August combined, and we don't distinguish between July and August, and so we compared the FES to one treatment, where we asked about the two months of the wave separately, and so we asked first about July, and then we asked about August, and then another treatment that asked about just fishing activity during the most recent one, and so we would only ask about August, and then the next month we would ask about September, and the next month we would ask about October.

Then we compared -- To look at recall, we compared -- I apologize that this figure did not turn out better, and I, for some reason, can't get my mouse to go over it anyhow, but we compared the FES estimates with the two-month recall period to results from these two experimental treatments, based only upon the most recent month, and so, even when we asked about the two months separately, we produced estimates just from the most recent month, and so the month with the one-month recall period, and then we added up the estimates from all those months, so they were comparable to the FES, and we found no significant differences between the experimental estimates and the FES estimates.

It doesn't appear that the two-month recall period is any more susceptible to recall error than a one-month recall period. In fact, we think the one-month recall period might be susceptible to telescoping error, because those estimates, while they weren't significantly different from FES estimates, they were systematically higher than FES estimates, and we don't think this is because people are forgetting about trips in the two-month recall period, because we only see it when we ask just about one month.

If we ask about two months separately, there was no systematic difference between the FES and the experimental estimates, and so we think what is happening is that respondents want to tell us about their fishing activity no matter when it happened, and so, with a shorter recall period, there is a lower probability that they actually fished during the wave, and so they are telescoping trips

from prior months into the one-month reporting period. The bottom line from this experiment is that we saw no evidence of recall error in the FES relative to a one-month recall period.

Another recommendation was that MRIP consider evaluating prospective data collection methodology and continue to research survey panels, and this is also something that we've explored, and we actually tested a longitudinal mail survey, a longitudinal panel survey design, to estimate fishing participation, even before the National Academy's was completed.

This study compared estimates from the FES design, which is a cross-sectional design, to estimates from a longitudinal design that we contacted the same respondents three times over the course of the year, and we were primarily trying to find an alternative method for estimating participation, and so just the number of people that participate in fishing.

Results from the project demonstrated that effort and participation estimates derived from a cross-sectional design with a twelve-month recall period, and so we're asking about fishing activity that occurred during the prior twelve months, were not significantly different from estimates produced by a longitudinal panel design in which respondents knew they would be asked to report fishing activity and they had a shorter recall period. Again, we found no significant differences between this particular panel design and the FES cross-sectional design for estimates of participation.

DR. SEDBERRY: Rob, can we interrupt with a question?

MR. ANDREWS: Yes, please.

DR. REICHERT: Forgive my ignorance, but what is a -- Remind me what is a longitudinal design?

MR. ANDREWS: A longitudinal design is a design that measures change over time, and so --

DR. REICHERT: Thanks.

MR. ANDREWS: Yes, and so we selected a sample of households, and we sent them a questionnaire, and we asked them if they fished, and then we asked them if we could re-contact them a few more times over the course of a year to collect similar information.

Then the final recommendation that I'm going to address is that the National Academy suggests that we continue to evaluate the -- That we evaluate the use of electronic data collection, and so, just this past year, we initiated a pilot study of a web push design for the FES that encouraged participants to respond to the FES through a computer or a mobile device before we offered them a paper alternative, and so we sent them a letter with an incentive, directing them to an online survey, and we encouraged them to respond to the online survey. Then, after a period of time, if they didn't respond, we sent them a paper questionnaire.

These study results are currently being evaluated, and I anticipate that a report will come out describing this study by the end of the year, and then the decision whether or not we implement a web push design -- We'll evaluate response rates, respondent representativeness, reported fishing activity, timeliness, and cost.

That is it for the FES design, and I'm going to switch gears a little bit and talk quickly about the transition process, and so a cross-disciplinary transition team has guided the transition from the Coastal Household Telephone Survey to the FES design, and we recognized that this was not going to be an easy transition, and so we started trying to plan for this early on, working with our partners, working with our stakeholders, to prepare everyone for what was coming.

The goals of the transition planning was to appropriately integrate revised estimates into stock assessments and management actions, create a process that we could repeat, as necessary, for implementing new or improved scientific methods, and the FES is not the only change that's happening with MRIP, and so, any time there is a change in methods, we want to be prepared for that and work with our partners to prepare them for that. We wanted to build stakeholder support, understanding, and engagement in implementing the new survey, and then, finally, advance the mission of NOAA Fisheries.

The transition planning process includes a benchmarking period, and so this is when new designs and legacy designs are conducted side-by-side and differences in estimates produced by each are measured. Then, during this period, statistics produced by the legacy design are still considered the best available science for use in science and management, and so, until the transition process is completed, we're encouraging folks to continue using the legacy designs.

Benchmarking is followed by calibration model development, where differences in statistics produced by each new survey are evaluated and possibly sources of bias that may explain these differences are explored. Research is conducted to assess how bias in the legacy survey may have changed over time, and calibration models are developed and evaluated for use in correcting historical catch estimates.

Once the calibration model has been approved, we re-estimate historical catch statistics, and so the approved calibration model is used to generate revised time series of historical catch statistics, and then the revised time series is made available to fisheries scientists and managers. Revised estimates are then incorporated into the assessment process. Revised catch statistics are incorporated into the assessments as soon as possible, and stocks with substantial mortality from recreational fishing parameters, or fishing pressure, are identified as key stocks, and those are prioritized.

Then, finally, the revised estimates are incorporated into the management actions, and that's kind of where we are right now, those last two points, with the FES, incorporating the estimates into the assessments and then into the management process, and so, for the FES -- Again, that transition process is sort of universal. We want to execute that process any time we're ready to implement a revised survey method.

For the FES, the timeline, the transition team was established in 2014, and that includes members of state and federal partners, scientists, stock assessors, and managers. We published a policy directive in 2015, which lays out the general transition process, and the transition plan for the FES transition was developed by the transition team in 2015, and that's when the benchmarking period began, and so we conducted the FES and the Coastal Household Telephone Survey concurrently for three years, beginning in 2015.

In 2016, NOAA Fisheries Science Centers explored stock assessment scenarios, and so this was looking at sort of the impact of potential changes in fishing effort on stock assessments, and so different ranges of changes and what impact that would have on the assessment process. We began working with independent consultants to evaluate differences in the CHTS and the FES estimates, and we explored possible sources of bias, and that really marked the beginning of developing a calibration model.

In 2017, the calibration model was peer reviewed, and it was unanimously endorsed and approved, and you're going to hear a lot more about this, I believe, on Wednesday. The peer review workshop, the members were -- The panel was convened by the Center of Independent Experts, and the workshop included members of the Mid and South Atlantic SSCs, and it was accessible via webinar. Also, in 2017, we revised historical estimates, and the CHTS was discontinued, and we were all very sad.

In 2018, we began incorporating -- Not we, but folks began incorporating new estimates into stock assessments, and then, through 2019, we're working with states and marine fisheries commissions and fisheries management councils and SSCs to respond to questions and concerns about the reasons that the surveys produced different estimates, and that's why we're here today. Are there any questions?

DR. LANEY: Rob, when you all are doing the surveys, do you ask anglers if they are keeping a journal and if they write down information? Given, if they do so, then their recall bias should be very much more limited than otherwise, and, to the extent that you are interviewing anglers who are avid enough to write things down, assuming they are writing them down correctly, then whatever that proportion is would tend to either increase or decrease the reliability of the data, I would think.

MR. ANDREWS: We do not ask if they record their information in a journal.

DR. GRIMES: You didn't mention this, but I think I remember reading in the report that the review panel produced that one thing they were concerned about was running the two methods side-by-side that you used to calibrate the two things. You got very different results, and that was an issue that concerned them. Otherwise, they were very enthusiastic about it, in general, and so did you guys reconcile that in some way or explain it?

MR. ANDREWS: I am going to tell you about that next. I mean, we've been evaluating the differences for years, as long as we've been conducting these pilot studies, and there are very big differences, and I will tell you about those next.

DR. SEDBERRY: Any other questions?

UNIDENTIFIED: Not a question, but just a comment. I just wanted to point out that the transition team that Rob described -- Actually, several members of that transition team are in this room today. I don't know if they would be willing to raise their hands, but --

DR. SEDBERRY: Do we have any questions from the webinar?

DR. ERRIGO: If you're on the webinar and you have a question, just speak up, because you should be unmuted. Then, when we hear you, we'll turn your mic on.

DR. SEDBERRY: Any other questions?

UNIDENTIFIED: On the recall work, the different periods of time, were they fishing directed pilot surveys, or were they also like the activity-type survey that you're looking at?

MR. ANDREWS: The design was exactly the same as the FES. The only difference was the reference period.

DR. SEDBERRY: Okay. Thanks, Rob. We have a decision to make here. It's ten after five, roughly, and we're scheduled to recess at 5:30, but we're one presentation behind, and I don't know if we can squeeze that presentation in between now and 5:30. Kind of judging --

DR. ERRIGO: I was thinking that maybe we can stay until 6:00.

MR. ANDREWS: I can give it as fast as you want. I can easily give it in twenty miles. No, I think that's probably -- I am guessing there's going to be some questions.

DR. SEDBERRY: I think so too, and this actually addresses the first term of reference. In addition to being some more background material, we're starting to get into the terms of reference, and so we don't want to rush through it, but I don't want to get too far behind, and so I'm open to suggestions, but I'm thinking that we should probably recess and maybe start at 8:00 tomorrow morning, rather than 8:30. Does that sound good to everybody?

Before we recess though, I wanted to take some time each day to kind of summarize where we were and see if we needed to make any recommendations or come up with any consensus statements, but, since we're still pretty much in the background material now, I don't think we're at that stage, and so I think we can go ahead and recess now, unless anybody has any additional questions on the stuff that's already been presented or any additional discussion points they want to make before we break for the evening.

DR. CROSSON: Just look at the terms of reference, if you haven't, tonight. I mean, that's just absolutely vital, because I am really hoping that tomorrow is a productive day moving forward on this.

DR. SEDBERRY: Right, and the terms of reference are multi-part, and so it looks like there is only two terms of reference, but there is really -- There's really a lot more to it than that, and it's very convoluted, and so, yes, take a look at them, and we want to make sure that we address them tomorrow in the presentations and in our discussions. I don't hear any objections to adjourning for the day, and so 8:00 tomorrow.

(Whereupon, the meeting recessed on August 19, 2019.)

- - -

August 20, 2019

TUESDAY MORNING SESSION

- - -

The Scientific and Statistical Committee of the South Atlantic Fishery Management Council reconvened an MRIP Workshop at the Town and Country Inn, Charleston, South Carolina, on Tuesday, August 20, 2019, and was called to order by Dr. George Sedberry.

DR. SEDBERRY: Welcome back to day two of the SSC MRIP Data Workshop extravaganza. As a reminder, we have our regular SSC folks at the meeting, and then we have a lot of experts and guests with us at this meeting, and so please identify yourself when you speak into the mic, because we haven't learned everybody's voice yet, and we're having a hard time figuring out who is talking, particularly for those folks that are on the webinar, that aren't in the room, and so, if you could, please identify yourself when you speak.

We're also a little bit behind schedule. We were supposed to get in three presentations yesterday afternoon, and we just got to the first two background presentations, and then the presentation we'll start with this morning, and the next few, will deal with the terms of reference, and so the first term of reference is Number 1, which was to look at survey estimate discrepancies, to review and describe the sources of disparity between the telephone and FES estimates of recreational effort for SAFMC-managed stocks, considering the impacts of the effort survey change and Southeast Fisheries Science Center post-processing.

This morning, I would like -- We're going to have a presentation with some background information to look at Sub-TOR 1A, which is describe, for a set of South Atlantic Council-managed species currently in the SEDAR process, how the sources of disparity between the telephone and FES surveys affect the FES catch estimate time series, with attention to trends, uncertainty, and potential outliers.

We'll jump right into that, and, for the SSC, this would be like note-taking Agenda Item Number 2, and so I have Jeff, Church, Anne, and Chris designated to take notes and provide a summary for this next session dealing with the Terms of Reference Number 1.

If there are no questions about any of that, or the organization of where we are, we can jump right into it, and so the first presentation this morning is -- I believe it's Presentation 3, and that's going to take an extensive look at the differences between the telephone and the FES surveys, and Rob is going to be presenting that this morning, and so go ahead.

MR. ANDREWS: Thank you, and good morning, everyone. Just one quick follow-up from yesterday is we had some discussion about what addresses were and were not included in the FES sample frame, whether seasonal addresses are included, and I confirmed this morning that they are not included in the sample frame, and so the seasonal addresses are not included in the sampling.

Yesterday, I discussed the process that we went through to come up with the FES design. This morning, I'm going to spend a little more time actually discussing the estimates and describing sources of bias that we believe contribute to the differences between the FES and the CHTS estimates.

DR. SEDBERRY: Rob, I hate to interrupt you. Sorry, but we're having -- Some of us are having a hard time figuring out which presentation this is. Do we have a PDF of this? Is it Presentation 3? Because it looks different.

DR. ERRIGO: Yes, but it should have been the very last Presentation 3 that I sent you.

DR. SEDBERRY: Okay.

DR. ERRIGO: It was the original Number 4.

DR. SEDBERRY: It was the original Number 4, the new Number 3, is the one that's on the screen now and the one that Rob is speaking from.

DR. ERRIGO: I will re-send it, just in case.

DR. SEDBERRY: Okay. Thanks, Mike. Okay. Go ahead, Rob.

MR. ANDREWS: Okay. As I mentioned yesterday, and I believe John also mentioned in his presentations, the FES was designed to address known limitations of the CHTS, and, as a result, the survey designs are very different. For example, the CHTS sampled landline telephone numbers, while the FES samples residential mailing addresses, and so the coverage of the CHTS declined pretty dramatically over the last fifteen years, while FES provides nearly complete coverage of the target population.

The CHTS asked the initial respondent a series of questions about household-level fishing activity, and so the CHTS assumed that the initial respondent was aware of the fishing activity of all household members, while the FES, the mail survey, provides respondents time to consider the request, determine who should respond, and potentially consult with others in the household.

The CHTS contacts households with no prior notice and expects an immediate response to questions about fishing activity. It's essentially a cold call. People were unaware that the call was coming, and they are expected to provide details about fishing activity for all members of the household with no advance warning, while the FES gives respondents more time to provide complete answers to the questions about fishing activity, and it also includes some memory cues to help enhance recall, such as providing a calendar that shows the dates that we're actually interested in.

The CHTS requires trip-level reporting, and so we asked respondents to recall and provide details about each angler's individual fishing trips that occurred during the reference period, and you can imagine, for people who took more than just a few trips, this becomes pretty burdensome, and so many respondents actually opted out after the survey, and they essentially refused mid-interview, after describing details for one or two fishing trips, and so, consequently, a great deal of the trip-level information from the CHTS was imputed from the information that they actually did provide.

In contrast, the FES requires summary reports of the number of shore and boat trips for each household member, and it's a much less burdensome questionnaire, and it takes much less time to complete, and, as a result, imputation rates are very low for the FES. Finally, as we discussed

quite a bit yesterday, the CHTS suffered from declining rates of coverage and response, while the FES was designed to maximize coverage and response rates, and so the FES, which was much less susceptible to bias resulting from non-coverage or non-response.

These differences between CHTS and the FES designs resulted in large differences in survey estimates, and that's pretty much why we're here today. The FES was a wholesale redesign, and it utilized a different sample frame, a different questionnaire, different survey mode, and, as I mentioned again, we specifically developed the FES to address growing limitations of the CHTS, and so we really didn't expect the estimates to be the same. It was a completely different survey, and we actually probably would have been concerned if the estimates were the same. We believe that the differences between the two surveys are largely the result of non-sampling errors in the CHTS, which have increased dramatically in recent years.

To date, when folks have been making comparisons between the FES and the CHTS, much of the focus has been on the FES estimates and whether or not they are feasible, but, really, to fully explore the differences between the two surveys, similar attention must be paid to the CHTS estimates, and so here you can see -- This is an index time series of CHTS estimates over the last twenty years or so, and, through the mid-2000s, the CHTS estimates tracked very closely with population, and so the blue line here is increase in population, relative increase in population, and the red line is relative change in the CHTS since 2000.

This increase continued through about the mid-2000s, and it was followed by a collapse in fishing effort that continued steadily through 2017, which was the last year that the survey was conducted, and so one of the questions we wanted to address is was the decline in fishing effort in the last ten years measured by the CHTS realistic, or was that an artifact of increasing survey errors.

Some have suggested that the decline in effort is a result of the recession that occurred during the mid-2000s, and that period is highlighted in gray on the figure on the screen, and so that was from between about 2007 and 2010, but, when we look at other indicators of fishing activity, it suggests that the effects of the economic downturn may not have continued through the end of the CHTS time series, and I apologize that this is a bit of a mess of spaghetti on the figure, but some of the other time series we looked at -- The green line, the green time series, is charter boat fishing effort, and so that's estimated from an independent survey. It's estimated from the for-hire survey, which includes sampling of for-hire vessels and then a telephone survey of those vessels to collect information about fishing activity, and so that also showed a decline through the economic downturn, but it has since recovered to levels that were even higher than fishing activity prior to the recession.

Some of the other ones on here -- The blue line, which is really kind of hard to see, is registered boats, and this is all for the South Atlantic region, and so registered boats increased slightly from 2000 to about 2007, and it took a bit of a dip, but it has since recovered, and rod-and-reel imports is the yellow line, or the gold line, and that tracked very closely with the CHTS effort through the mid-2000s, and it showed a very similar decline as the CHTS, but it then has since recovered to levels that are much higher.

Then some of the other ones that we're looking at that show similar patterns, to different degrees, are gasoline sales, which is the bottom line here, I believe, or, actually, I think that's one of the

ones in the middle, and then outboard engine sales, which I believe is the one that shows a pretty big dip during the beginning of the recession, but it, again, has since recovered.

All of these, we believe, are indicators that should track very closely with recreational fishing activity. They all declined during the recession, but they have since recovered, and so believe that the continued decline in CHTS estimates after the recession is largely the result of growing non-sampling errors, and it's not really a function of actual fishing activity.

The first type of non-sampling error that I'm going to discuss is coverage error. Coverage error occurs when the survey sample frame excludes some fraction of the target population and the covered and uncovered populations are different with respect to what is being measured.

The CHTS was a landline random-digit-dial telephone survey, and so the target population included all households within coastal counties, but the sample frame only included households with landline telephone numbers, and so the CHTS would have been susceptible to coverage error if a large fraction of households in the target population did not have a landline telephone and households with and without landline telephones were different with respect to recreational fishing activity. Just the fact that some households are not covered does not necessarily mean that there is coverage error. The covered and the uncovered portions of the population need to be different with respect to what is being measured.

Beginning in 2003, the National Center for Health Statistics added a question about telephone use to the National Health Interview Survey, or the NHIS, and this survey is a face-to-face survey that's administered by the Census Bureau on behalf of the Centers for Disease Control and Prevention. Since 2003, when they began asking the question about telephone usage, the number of Americans living in wireless-only households has steadily increased.

When they first asked the question, less than 5 percent of households reported being wireless only. Since 2003, that has increased dramatically, and, in June of 2018, the survey estimated that 55 percent of American homes are wireless only and do not have landline telephones at all, and so the four series that are plotted on this graph are four different regions of the country that the survey recognizes, and so the three that track very closely together are the Midwest, the South, and the West, and those are all about 60 percent wireless now, and so about 60 percent of all households are wireless only in those regions, and that would include the South Atlantic, and so the South Atlantic is included in the South region. Then the Northeast shows a similar trend, but is lagging a little bit behind, and so that's about 40 percent wireless, as of 2018.

Even among the households with landline, 42 percent of those who said they had a landline identified themselves as wireless mostly, and so these are households that have a landline, but almost all of their recalls are received on wireless telephones, and so these are, effectively, non-landline households as well. These are people who have a landline, but they never answer them. They might not even remember that they have them.

In other words, about 70 percent of U.S. households either don't have a landline or they're unlikely to answer a landline telephone, and, so, over the next several slides, I am going to refer repeatedly to the landline population, and, unless otherwise indicated, I am generally referring to the portion of the population that excludes wireless only and wireless mostly. We observed a similar trend to the National Health Interview Survey in telephone usage among anglers.

Between 2000 and 2012, the MRIP Access Point Angler Intercept Survey asked anglers whether they had a working landline telephone at home. During this period, the proportion of anglers that reported being wireless only steadily increased, and so 2000 was over, and 2012 was the last year that we actually asked it, and that's designated by the vertical bar. The points to the right of the vertical bar are modeled based upon the time series prior to 2012, but, during this period, the proportion of anglers that reported being wireless only steadily increased. The intercept survey did not distinguish between wireless only and wireless mostly, and so this particular slide just refers exclusively to the wireless-only portion of the population.

By 2012, between 35 percent and 60 percent of intercepted fishing trips in the South Atlantic were taken by anglers in the wireless-only households, and so each of these series represents a state in the South Atlantic. The top two, the orange and the gold, are South Carolina and Florida, where about 60 percent of trips were taken by anglers that were wireless only, that reported being wireless only, and then the blue and the gray, the bottom two, that track pretty closely together, are North Carolina and Georgia.

These results clearly show that the landline population, and, consequently, the coverage of the CHTS decreased dramatically over the past decade, to less than -- A minimum of 50 percent, and possibly less than 30 percent by the last years of the survey and this is not just based upon our data. This is primarily based upon the National Household Interview Survey, and this is -- The NHIS is sort of the standard benchmark that is measured by U.S. statistical agencies to monitor the wireless-only population.

This magnitude of under-coverage will result in extreme bias if the covered and uncovered populations are different with respect to fishing activity. If we're only covering half the population, or less than half the population, if the covered and uncovered portions of the population are different, there is going to be a large bias, and so the big question is, is the landline population, and so those households that would have been covered by the CHTS, representative of the overall population? Are they different than the people that are covered?

The answer to this question is absolutely not. Here, you see age distributions also from the National Health Interview Survey, and the panel on the left is the age distribution from the full population, and so it includes the wireless population and the landline population, and the panel on the right is the age distribution from the landline portion of the population, and this is for each year from 2013 to 2017.

In 2017, half of the landline population was estimated to be sixty-five or older, and so these are age distributions, and so the top two portions of the bar on the far right-hand side are sixty-five and older, and so the green bar is sixty-five to seventy-nine, and the top bar, the blue, is eighty-plus, and so this survey estimated that over half of the population in landline households were over sixty-five. Compared to the full population, the survey estimated that about 25 percent of the overall population is over sixty-five, and so, according to this survey, the landline population is much older than the full population.

The NHIS also estimates a variety of health and fitness parameters, and so, here, you can see the proportion of adults who reported not being able to walk one-quarter mile, or find the task very difficult. The green series, which is on top, is the landline population, and the blue series, which

is on the bottom, is the full population, and so the landline population -- By 2017, about 25 percent of that population reported not being able to walk a quarter mile, or find the task very difficult, and that compares to less than 15 percent from the full population. It looks like the rate at which this metric is increasing is larger for the landline population than the full population, although they are both increasing as the population ages.

Here, you see the proportion of adults who can't stoop, bend, or kneel, or find the task very difficult, and, again, the landline population has a much higher prevalence of this than the full population, and then, finally, the proportion of adults who can't stand for two hours, or find the task very difficult. Again, over 25 percent of the landline population report having difficulty with this, compared to about 15 percent for the full population, and so the landline population is much older than the overall population, and it generally exhibits characteristics associated with poor health, or health that is not as good as the full population, and, again, this is information that's from the National Health Interview Survey.

We can also evaluate basic demographic characteristics of the FES samples for both the full sample and the portion of the sample that identifies as landline mostly or landline only, and we talked a little bit about some of these characteristics yesterday, and these are the demographic characteristics we used to adjust some of our weighting, and so this figure compares demographic characteristics estimated from the full FES sample, and that is the orange bars, the landline portion of the sample, which is the blue bars, and the American Community Survey, which, as we mentioned yesterday, is administered by the Census Bureau, and it's considered to be the gold standard for population data.

Compared to the American Community Survey, the FES landline sample, which, again, is in blue, severely underestimates households with children, and so that's the first series of bars here, and the landline sample reported about 2 percent of households with children, and this compares to 21 percent from the American Community Survey, and so the landline sample, the FES, underestimates, or underrepresents, households with children.

The landline sample, the landline FES sample, also overestimates, or overrepresents, households with seniors. 86 percent of households in the landline sample reported at least one individual that was age-sixty-five or older, and that compares with 43 percent estimated from the American Community Survey, and so the landline sample overrepresents households with seniors. Households with seniors that are eighty-plus, 38 percent of the landline sample reported having a senior residing at the address in the landline sample, compared to about 10 percent from the American Community Survey, and then single-female households --

MR. FOSTER: Rob, sorry to interrupt you, but I think the orange series is the FES full sample and not the census, and the green series is the American Community Survey census values, correct?

MR. ANDREWS: That is correct, yes. Sorry. Right now, I should be comparing the blue bars to the green bars, instead of the orange bars. Fortunately, they are pretty similar, and so the FES landline sample underrepresents households with children, and it overrepresents households with seniors, both sixty-five-plus and eighty-plus, and it overrepresents households with single females, and so single-female households, relative to the American Community Survey.

In contrast, estimates derived from the full FES sample, which includes both the landline and the wireless-only households, are much closer to census values, and so that's -- Again, the FES full sample is orange, and the census values are in green. They are much closer to the census values, and, as we mentioned yesterday, differences between the FES and the census values are reduced during estimation, through adjustments to sample weight, and so we use these census values, the green values, to adjust the weights from the FES full sample, and so the population, the FES sample, is representative of the U.S. population with respect to these parameters.

The bottom line of all of this is that the landline samples are not representative of the U.S. population with respect to these demographic characteristics, and these are the households that would be covered by the CHTS.

Now, again, this suggests that landline samples, or the CHTS samples, would have been severely biased with respect to these demographic parameters. Now, that doesn't necessarily mean that the CHTS estimates of fishing effort were biased. The primary measure of interest in the CHTS is fishing activity, and so CHTS estimates will not be biased unless fishing activity correlates with these demographic characteristics, and so we were able to take a look at that, using the FES data, and so here we see fishing prevalence, and so this is the proportion of households that fished for each of these demographic characteristics.

The blue bars, in each of these cases, are households with a specific attribute, and the green bars are households without the attribute, and so the first set of bars are households with children, and so the blue bar represents households with children, and the green bar represents households without children, and we see that fishing prevalence is quite a bit higher in households with children, and so it's about 13 percent of households with children reported fishing, while only about 10 percent of households without children reported fishing, and the dots on this graph are the ratio between those prevalence estimates, and so the difference -- They are different by a factor of a little less than 1.5. Again, these are estimates for the South Atlantic states.

The second set of bars are households with seniors, and so the blue bars are households with seniors, and the green bars are households without seniors. Households without seniors are about 1.5-times -- They report fishing activity at a rate of about 1.5-times that of households without seniors, and so the prevalence is a little over 12 percent, compared to a little less than 8 percent.

Single-female households are much less likely than non-single-female households to report fishing activity, and so that's about a little over 12 percent, compared to about 5 percent of households with single females, only a single female, and then landline households are much less likely to report fishing than non-landline households, and so the blue bar here is landline households, and the green bar is non-landline households, and those are different by about a factor of 2.5.

Clearly, fishing activity is correlated with these demographic attributes, and this is a very clear indication that landline samples, or estimates derived from landline samples, are severely biased with respect to fishing activity.

We can evaluate the magnitude of this bias by comparing effort estimates derived from the landline portion of the FES sample to those derived from the full FES sample, and so these are actual population estimates. The sample weights for both of these estimates are adjusted to reflect the full population for both sets of estimates, and so the blue bar is the FES landline sample, and so

these are people who reported having a landline, but we re-weighted that sample to reflect the entire population, and then the green bar is the full FES sample, and so these are actual estimates of fishing prevalence, or the proportion of households that fished, and this is for overall prevalence, boat fishing prevalence, and so the proportion of households that reported boat fishing during the wave and the proportion of households that reported shore fishing during the wave.

Full sample FES estimates of fishing prevalence are over two-and-a-half-times larger for overall prevalence, and so that's this first comparison here, and 3.2-times larger for boat prevalence, and so that's about a little over 6 percent of households in the full sample report boat fishing compared to about 2 percent for the landline sample and then 2.6-times larger for shore fishing prevalence, and so almost 8 percent of the full sample population reported shore fishing during the wave, compared to about 3 percent for the landline sample.

FES prevalence estimates derived from the landline sample, these estimates in the blue bars, these are actually pretty similar to the estimates we observed for the CHTS during the most recent years of the survey, and this is very strong evidence that the FES landline estimates, as well as the CHTS estimates, underestimate fishing activity by over a factor of two, because the population represented in these samples, older residents, more single-female households, household members in poor health, are unlikely to fish, and so the people covered by landline samples just don't fish very much. Just a quick summary of the coverage error --

DR. SEDBERRY: Can we interrupt with a question?

MR. ANDREWS: I'm actually -- I have two more slides after this one, and then I'm going to switch to kind of a different topic, and so if we can get through just maybe the next couple of slides and then --

DR. SEDBERRY: Great.

MR. ANDREWS: Just a real quick summary of coverage error. Landline households have become increasingly different from the rest of the United States. These households report older residents, fewer children, smaller household sizes, more households comprised of single females, poor health, and less leisure time and physical activity, and these demographic differences really matter. Households covered by the CHTS, they just don't fish. By the time the survey ended in 2017, the population was so different from the overall population that the estimates just were not valid estimates of fishing activity.

Large differences in household characteristics between those who were and were not covered by the CHTS sample frame resulted in significant coverage bias, which we believe is responsible for a large portion of the differences between the survey estimates. These results suggest that this collapse in fishing effort after the mid-2000s was largely a function of eroding coverage error in the CHTS, and so, as the landline population became increasingly different from the overall population, estimates derived from landline samples became increasingly more biased.

Had effort continued to increase, and so remember it increased pretty steadily with population through the mid-2000s, and, had that increase continued through the entire time series, the CHTS estimates at the end of the time series would have been much closer to the FES estimates, and so

we believe that this comparison that everybody has been looking at in 2017, looking at these really low CHTS estimates, and these big FES estimates, just is not a valid comparison.

The 2018 FES -- This time series shows the relative change in the CHTS time series, and then the green dot here at the top is the full FES sample estimate. The orange dot here is the estimate of total fishing effort based upon just the landline sample, and so this estimate, the 2018 FES estimate derived from the landline sample, actually was lower than the 2017 CHTS estimate, which we think might suggest that coverage would have continued to erode, had the CHTS continued.

Just, quickly, to kind of wrap up coverage error, the FES estimates of total fishing effort derived from the landline samples weighted to the full population are similar to CHTS estimates, and so this figure shows estimates of angler trips for private boat fishing and shore fishing, and this is for the South Atlantic states, and the orange bar is the FES full sample estimate, and the gray bar is the FES landline estimate, and the blue bar is the CHTS estimate.

When we produce an estimate from FES data that's just based upon the landline sample, we get an estimate that's actually pretty close to the CHTS estimate. In fact, for private boat fishing, the estimates from the landline portion of the FES sample was actually lower than the CHTS estimates we've been getting, and so we believe this explains a large portion of the differences between the CHTS and the FES estimates, particularly for private boat.

For shore fishing, coverage error we believe explains a large portion of the difference between the FES and the CHTS estimate, and so, again, the orange bar is the full sample FES estimate, and the gray bar is the landline estimate, and the blue bar is the CHTS estimate, and so this explains about 80 percent of the difference between the full sample FES estimate and the CHTS estimate, but, even after adjusting for coverage error in the CHTS, the FES estimate is still about two-times larger than the CHTS estimate, and so clearly other sources of error must be contributing to this remaining difference, and it seems like it has a differential impact for shore and boat fishing, and so these other sources of error seem to have a bigger impact on shore fishing than private boat fishing. Now I'm going to switch gears a little bit, and so, if there are questions, I would be happy to take some.

DR. REICHERT: Thank you. Were other attributes looked at, for instance income, or were they correlated with some of the attributes that you guys analyzed?

MR. ANDREWS: We actually don't collect any income data in the FES. We haven't looked at any of that yet. We do get census data, census geographic data, appended to the FES data, which allows us to look at some sort of larger-scale economic data that we can try and correlate with some of this, but we haven't done any of that yet.

DR. LI: Thank you for this presentation. It has kind of cleared up half of my concerns about what has caused these differences between the two survey estimates. Here, my understanding is -- So the telephone survey is targeting the landline population, and the landline population no longer represents the true population. However, back to before the mid-2000s, we see the landline population still tracking the true population, right, which means that the landline population back before the mid-2000s still represents the true population, right?

Given that, I would expect to see the effort estimate, fishing effort estimates, before the mid-2000s, from those two surveys, they should look similar, given they are tracking the same true population, but, after the mid-2000s, I agree that the landline population no longer represents the true population, and that's why we see the different estimates from those two surveys, and that's the first question.

The second question is, when you showed those demographic differences, those bar graphs, what kind of -- What years of data did you use? Again, if you used the data before the mid-2000s, would we still see those different bars in the proportions in the demographic features?

MR. ANDREWS: Are you referring to these ones or the ones from the FES?

DR. LI: Like this. For example, if you use the data before the mid-2000s, the landline population versus the full --

MR. ANDREWS: We didn't conduct the FES in the mid-2000s.

DR. LI: Like the census.

MR. ANDREWS: Right. I would expect, the further back in time we go, that landline samples would be more similar to the American Community Survey estimates.

DR. LI: That goes back to my first question. If that's true, which means before the mid-2000s, we are sampling the same population. I mean, the telephone survey is sampling the true population, and so we would expect to see the effort, fishing effort estimates, before the mid-2000s should be similar between the two surveys.

MR. ANDREWS: If the two surveys were conducted side-by-side during the mid-2000s, we would have expected the estimates to be closer together. At least according to this analysis, we would have expected that, yes.

DR. LI: I am starting to think that we are sampling different populations, and the estimates after the mid-2000s are not comparable, because we are sampling different populations.

MR. ANDREWS: Yes, exactly. We believe that those estimates are not at all comparable, because they are sampling two very different populations.

DR. LI: But, at least before the mid-2000s, they are comparable.

MR. ANDREWS: In terms of the survey coverage, yes, they would have been sampling a population that is much more similar, absolutely.

MS. FRANCO: Can we go back to page 20, or forward to page 20, I guess, from there? When you were talking about this, you said that the CHTS was underestimating by a factor of two. Should we interpret that as the difference between FES and CHTS should be around a magnitude of two, because what we're seeing is magnitudes of two, sometimes eight, sometimes ten, difference, and so I just wanted some clarification on the factor of two.

MR. ANDREWS: The impact of coverage error on the CHTS we believe resulted in an underestimate by a factor of two to three, and that's just talking about the coverage error, and I'm going to talk about some other sources, potential sources, of bias moving forward.

MS. FRANCO: Okay, and so we should not interpret that as -- Okay. It does not explain all of the magnitude, is what you're saying.

MR. ANDREWS: That is correct, and keep in mind that all of these figures are looking at the South Atlantic as a whole, and so there is going to be state-by-state differences as well, and I believe -- They're not actually included in the presentation, but, at the end of the presentation, I included some state-specific estimates as well, so you can take a look at that for Georgia, and, if we have time, we can -- I can perhaps present some of that, but they are in the presentation that was sent around, the state-specific estimates. I guess the bottom line is we believe the coverage error is responsible for a large portion of the differences between the FES and the CHTS estimates. We think that is the predominant source of the differences between the two surveys.

MS. FRANCO: The population coverage and not just the wireless, the difference between wireless and landline.

MR. ANDREWS: It's driven primarily by the eroding coverage of the CHTS, which was because of the growth of the wireless-only population, but, yes, that was sort of the driver behind the decreasing coverage of the CHTS.

DR. SEDBERRY: Any other questions for Rob before he moves on?

DR. DUMAS: Getting back to the seasonal households again, just for a second, let's assume that seasonal households fish less than non-seasonal households, and so seasonal households were included in CHTS, but seasonal households were not included in FES, and so, if the seasonal households are not included in FES, then the FES would not be including households that fish less, and so the FES estimate of fishing would be higher, because it was not including households with a low level of fishing, namely the seasonal households.

I don't know if that effect is significant, because it kind of depends on the proportion of seasonal households as a proportion of all the households, and it also depends on how much -- The difference between fishing of the seasonal household versus fishing of a regular household. I don't know if that effect is significant, or if it's anywhere close to the effect that you just outlined, which is clearly significant, what you just mentioned, the difference in the coverage based on landline versus cellphone.

MR. ANDREWS: In the CHTS, we also attempted to remove seasonal houses from the sample. Those numbers could have been included in the sample, but we have -- In the screening portion of the survey, we specifically ask if this is a permanent residence, where they spend more than six months of the year, and, if they weren't, we terminated the interview, and they were classified as ineligible, and so we attempted to account for that.

DR. DUMAS: Thanks.

DR. SEDBERRY: Any other questions?

MS. SAULS: Can you explain how you identify a seasonal household and what that definition is?

MR. ANDREWS: We identify it -- Well, there's a flag in the address-based sample frame. The Post Office includes a designation of a seasonal address in their list of addresses. I assume that that is provided by the property owner, but I'm not entirely sure about that, and I can try and find the answer to that. We don't define it. It's defined by the Postal Service.

MS. SAULS: I am just thinking about the residents who come down to the southern states seasonally, for the winter months, like a two-week rental, or is it a six-month rental, and how is that defined?

MR. ANDREWS: I don't know the answer to that right now.

DR. DUMAS: If both the CHTS and the FES were trying to estimate the fishing behavior of non-seasonal households, then, when the sample is expanded up to the full population, it should be expanded up to only the population of non-seasonal households.

MR. ANDREWS: It's expanded up by census estimates of occupied housing units.

MS. SAULS: I was curious about the American Community Survey, and so I found the questionnaire yesterday, and it actually asks about people who have been living in the house for the past two months, and it really doesn't seem too concerned between a resident of another state versus a resident of the state they're being surveyed in, and so I don't really -- I mean, I don't know enough about the census to understand how that is being expanded, but is a census estimate for a state actual legal residents of a state, like they have voting registration, or a driver's license in that state, or is it just people who are living in the state at the time that they were counted?

MR. ANDREWS: I am not entirely sure what their definition is.

DR. SEDBERRY: Rob, I think what you were saying earlier is that the Postal Service's database would screen out some addresses before we even got to the question that Bev is asking, I think.

MR. SEDBERRY: Yes, they would not be included in the sample frame, but then we do, even after we select our sample, send out the questionnaire and get them back, and we post-stratify our weights to population control totals of the number of occupied housing units in a state, and that information comes from the Census Bureau or the American Community Survey estimates.

I am not entirely sure how they define what a resident is, or an occupied housing unit is, off the top of my head. State estimates are mutually exclusive, and so you don't have somebody who occupies a housing unit in Florida also being counted as occupying a housing unit in Ohio, for example, and so they do account for that. I just don't know, off the top of my head, what that process is.

DR. SEDBERRY: Any other questions before we move on? Thanks, Rob. I think you're ready to go.

MR. ANDREWS: Those are very good questions. I don't want to lose sight of the fact that the information I just presented explains a lot of the difference between the CHTS and the FES estimates. The differences in the number of occupied housing units, that's all -- It's very important, and they are very good questions, but I don't want these results to get lost in that discussion. I mean, we're accounting for over half of the difference, much more than half. For boat fishing, we believe that we're accounting for almost the entire difference between the FES and the CHTS estimates, because of this coverage error, and so I would like to keep that sort of in the front of your mind as we move forward. That's the big portion of the difference that I'm going to talk about.

For shore fishing, there was still -- Even when we account for the declining coverage of the CHTS, there was still some differences between the FES and the CHTS estimates, and so the next type of error that I'm going to talk about is the type of measurement error that we refer to as a gatekeeper effect.

In the CHTS, we ask the initial person who answers the telephone a series of screening questions to identify fishing households, and these questions are how many people in this household go fishing, how many people in your household, including children and adults, have been recreational saltwater fishing in the past twelve months, anywhere in the U.S. or in a U.S. territory, and then, thinking just about the past two months, how many people living in your household have been recreational saltwater fishing in the past two months in the U.S. or a U.S. territory, and so these are the questions that are asked of the person who picks up the phone.

We don't ask to speak to an angler, and we don't ask to speak to a specific member of the household. It's whoever picks up the phone is asked these questions. If that person provides a negative response to any of these screening questions, then the interview is terminated, and the household is classified as a non-fishing household, and so it becomes a zero.

In looking over the CHTS data, we find that the screening respondents are overwhelmingly female, and so about two-thirds of the people who pick up the telephone are females, and this is common in all landline telephone surveys. A female is much more likely to answer the phone than a male, and females are much less likely to report household fishing than male screening respondents.

Based upon this observation, we hypothesize that the screening respondents are inappropriately screening households out of the CHTS interview, because they don't fish, and they really are not aware of the fishing activity of other household members, and so this type of screening error, which we call the gatekeeper effect, would result in underestimates of fishing effort.

We designed an experiment in 2012 to test for the gatekeeper effect, and, in this study, samples were selected from lists of licensed saltwater anglers in North Carolina and randomly allocated into two treatments, and so there was a control treatment, where we asked for the licensed angler by name prior to administering the CHTS screening questions, and so we sampled a licensed angler, asked for that individual by name, and then we asked the CHTS household screening questions, and we collected information for all household members and not just the licensed angler.

Then the second treatment was a CHTS-like treatment, in which we administered the screening questions to whomever answered the telephone, and so both treatments sampled from a list of

licensed saltwater anglers, and so they were independent of samples from the same population, and so, theoretically, they should have come up with similar results.

The purpose of the experiment was to determine if the identity of the screening respondents impacted reports of fishing activity, and what we found is that reported fishing activity was considerably higher when we asked to speak to the licensed angler by name prior to administering the household screening questions, and so this figure shows prevalence ratios, and it's the prevalence when we asked for the licensed angler by name divided by the prevalence when we just administered the screening questions to whomever picked up the telephone.

Then, on the left-hand side, the dots on the left-hand side are shore fishing, and, on the right, they're for boat fishing, and we did this for four waves, and so each of those points represents a wave, and then the horizontal line is the overall mean for shore fishing on top and boat fishing on the bottom.

This, we thought, was pretty strong evidence that the CHTS respondents, and so the person who answers the phone, are underreporting household fishing activity, and the impact of the gatekeeper effect was larger for shore fishing than for boat fishing. For shore fishing, household prevalence is 1.26 times higher overall when we administered the screening questions to the licensed angler, and, for boat fishing, the ratio is about 1.1.

These differential effects between fishing modes seem logical, considering the nature of the fishing activities, or at least they seem logical to me, which I don't know if that's logical or not, but, if you think about boat fishing in general, it requires a much greater investment in time and money than shore fishing, and so boat fishing is likely to be more memorable to all members of the household and not just the angler, him or herself.

MS. LANGE: I find it hard to believe that people living in a home together wouldn't know if their spouse or child had gone fishing, number one. Number two, you are putting all of the misreporting on the supposed non-fisher, as opposed to -- How do you account for me being a fisherman and not wanting to admit that I haven't been fishing this month, or that I haven't caught anything when I did go fishing? Is there any way to look at that bias, as opposed to just the fact that whoever answers the phone may not know that their other half has been fishing at all? I am just curious on that, how that's accounted for, and it's something that applies to any survey, but I am just curious.

MR. ANDREWS: No, you're absolutely right. This makes the assumption that the person who is not the licensed angler is underreporting fishing activity, but it is possible that the licensed angler is overreporting fishing activity, and the results could be the same for that. This doesn't account for that. We don't know whether people actually did fish or not, and so we don't know that the licensed individual is overreporting fishing activity or the person who picks up the telephone is underreporting fishing activity, if that's what you're asking.

MS. LANGE: But your adjustments are strictly in the add more effort, as opposed to leave it, and it seems as though your bias is always increasing, as opposed to accepting that maybe the fisher is overreporting.

MR. ANDREWS: Correct. We are just looking at potential sources of differences between the two surveys, and so, whether we're adjusting FES estimates down to account for bias or adjusting

CHTS estimates up to account for the opposite bias, the estimates are still moving in the same direction. Can you follow-up with your question? I just might not be understanding your question.

MS. LANGE: Again, it applied probably to MRFSS as well, but, if you're only looking at -- If you're assuming that the phone calls -- If you're assuming that everything was underestimated and that nobody in my house fished, okay, and then you talk to the fisher, and he says, oh, I went fifteen times, when he only went once, are you assuming that his effort estimate is the one that you should be using in estimating the overall effort?

MR. ANDREWS: This was just an experiment to determine if the screening respondent mattered, and we're not making -- We didn't make any adjustments to estimates based upon this. This was just an experimental design where we screened households. We used two different screening approaches, and we wanted to see if there was any difference in reported fishing activity between those two different screening approaches. That's all we were doing here, was just an experiment, and this was not -- We are not making adjustments to our historical estimates based upon this or not.

We just wanted to see if we got different answers when we asked the questions two different ways, or we asked the questions to two different people, and we found that we did. We got very different answers, and so the person that -- The person we asked the screening questions to matters. Now, whether the licensed angler is overreporting fishing activity, or the person who picks up the phone is underestimating fishing activity, we don't really know true fishing activity, but we are making the assumption that the person who picks up the phone is underestimating fishing activity. Again, we didn't make any adjustments based upon that. It was just an experiment to see if the screening approach made a difference.

DR. LANEY: Rob, your use of the term "true fishing activity" causes me to reiterate a conversation that I guess some of us were having earlier about whether or not it would be possible to perhaps resolve some of the uncertainty around this whole discussion by actually designing an experiment to count who is really out there and who is really fishing and then compare that to the estimate that you generated for the same area for the same period of time, and I understand, Beverly, that you guys may be trying something along those lines in Florida, and so would that be a possible resolution to a lot of the uncertainty around this, if we could design something like that?

MR. ANDREWS: Yes, of course. If you could design an experiment to measure true fishing activity, then absolutely. You would know what the right answer was, but I think -- This is something we've been thinking about for a long time, but it's -- Given the nature of the coastline in the Atlantic and the Gulf of Mexico, that's not a simple thing to do, I guess is the bottom line. Coming up with an experiment to do that would be very challenging, but, yes, certainly, if we could come up with the right answer, we would know the right answer, and then we could base everything on that.

DR. LANEY: For follow-up, we were also talking about the fact that the state of the art, in terms of remote sensing for aerial imagery, continues to just get better and better and better, and so I'm just wondering if there's a possibility that, for counting -- We count snow geese and tundra swans now from aerial imagery, and I'm wondering if we couldn't count boats, or we count trailers, again in a certain area for a certain period of time, and maybe get a better handle on what the true fishing rate really is.

MR. ANDREWS: I think using advanced technologies like that has a lot of possibilities for data collection. We are certainly not averse to exploring other ways to try and get a handle on fishing activity, absolutely.

DR. SEDBERRY: Scott, did you have a question about the experiment?

DR. CROSSON: Yes, I do. Actually, this is something that I meant to get to yesterday. We have been talking about the Gulf of Mexico and the South Atlantic, but I'm curious. A lot of these increases that we have seen in the Southeast with recreational landings through FES, has that been a similar issue in other areas that have MRIP, such as the Northeast? Also, MRIP is Hawaii, right? Don't they do MRIP in Hawaii? Have you seen any of these increases there? That's several questions, and I'm sorry.

MR. ANDREWS: It's the same answer for all of them. Yes, we've seen similar increases across-the-board. The FES is systematically higher than the CHTS everywhere we conducted it. The magnitude of the differences might vary a little bit from region to region, and I can't speak off the top of my head to how the different regions compare, but the FES estimates are higher across-the-board.

MR. DETLOR: Just in response to Wilson, we have looked at using aircraft surveys for a number of different survey instruments within NOAA, and it's really cost prohibitive. There is also significant concerns when it comes to confidentiality and using basically a civilian agency taking a look at the populous.

DR. LANEY: For aircraft, yes, I can understand that, but there is, I think, an increasingly accessible and relatively low cost aerial imagery from some of these new little-bitty satellites that they're putting up that have real high resolution imagery in them, and so it's worth looking into, I think.

DR. SEDBERRY: Okay. I think now we're ready to move on.

MR. ANDREWS: Thank you. This slide showed the overall differences between the two treatments we were looking at in this gatekeeper experiment. We believe that this represents a minimum impact of the gatekeeper effect. The sample frame we used for this study, again, it was a list of licensed anglers in North Carolina, and it included cellphone numbers, and so it wasn't all landline numbers, and so there was a higher likelihood that, when you called one of these numbers, that you would actually get the licensed angler on the line than you would find in a landline survey, and so, if you're sampling landlines, you would be less likely -- It would be less likely that a licensed angler would pick up the phone than if you're sampling from a license frame or a list frame that includes cellphone numbers, which are attached to an individual and not a household.

We did match the license frame to commercial databases of landline numbers before we did the experiment, and so this allowed us to distinguish between known landline numbers and unknown numbers, and so, if a number on the license frame matched, we were reasonably sure that it was a landline, but, if it didn't match, it didn't mean that it was not a landline, and so these commercial databases of -- I mean, it's essentially the white pages, and they're not complete, and so there's a lot of people who have landlines that aren't listed, and so you can't match them.

When we look at the differences between households that do and do not match to a landline directory, we see that the difference between the experimental treatments were quite a bit larger when we limit the analysis to known landline numbers, and so the known landline numbers here are the green points for shore fishing and the blue points for boat fishing, and then the orange and the yellow are ones that we don't know whether they are landlines or not.

For the landline sample, prevalence was about 1.4-times higher for shore fishing when we asked for the licensed angler by name, and about 1.16-times higher for boat fishing, and, again, these are when we administer the screening questions to the licensed angler. Then, for the other portion of the sample, the prevalence ratios between treatments were pretty close to one, and, again, that's this point and this point, and so that's the portion of the sample that could not be matched to a known landline number.

To summarize the gatekeeper effect, females are more likely than males to answer a landline telephone, and females are much less likely to report household fishing activity than males. The results from the gatekeeper experiment confirmed a gatekeeper effect. The screening respondent does matter. You get different answers when you ask the screening questions to different people. The gatekeeper effect is larger for shore fishing than boat fishing, and we believe the gatekeeper effect results in an underestimate of fishing effort by as much as 30 percent, and that's approximately, just based upon the results of the study.

When you consider both the under-coverage the CHTS sample frame and the gatekeeper effect, we believe that explains about 95 percent of the difference between the FES and the CHTS estimates of shore fishing effort, and so this is just shore fishing effort. The orange bar there is the full sample FES estimate, and the gray bar is the FES landline sample estimate. The blue bar, or the purplish bar, is the FES landline estimate when we account for the gatekeeper effect, and then the lighter-blue bar on the right is the CHTS estimate.

When we adjust the FES estimates of shore fishing effort to be similar with the CHTS, with respect to coverage and the gatekeeper effect, the estimates are within about 20 percent of the 2017 CHTS estimate, and so, if we simulate coverage error in the FES, that reduces the estimate by about two-thirds, and so we go from the orange bar to the gray bar, and then simulating a gatekeeper effect further reduces the estimate by about an additional third, and so that gets us from the gray bar down to the purple bar, and we believe that explains a large portion of the difference between the CHTS and the FES for both shore fishing and boat fishing. I am going to switch gears a little bit again, and so, if there's more questions, this might be a good time.

DR. SEDBERRY: Any questions?

DR. NESSLAGE: Maybe you said this, but why such underreporting of shore effort versus boat? Do we have any idea why humans are behaving that way? It doesn't seem to make sense to me. Is it an income thing? I mean, do we have any idea who these people are that are -- These women are who are underreporting shore effort and why that might be? I don't understand human behavior that well.

MR. ANDREWS: I don't either. We have talked a lot about this, and we think one of the things that's going on is that shore fishing is just not as memorable as boat fishing. It's a less salient

event, and it can be walking down to the dock and throwing a line in for fifteen minutes, and it can be that you're out on the dock getting your boat ready to go out, and you see a big snook swimming around the dock, and so you throw a line at it, and that's a shore fishing trip, as far as the FES is concerned. It doesn't have to be an all-day event. It can be five minutes, or it can be ten minutes, or it can be all day. It's all the same, in terms of the fishing effort estimates.

We think one of the things that is happening is that it's people don't remember these sort of shorter-term, less-impactful events, like shore fishing, whereas, if you go out on a boat fishing, it's a little bit of a bigger deal. You've got to pay for gas, and you've got to haul a bunch of stuff down to the boat, and you've taken more time off of work or whatever, or your day, to go out and do it. It's just a more memorable event, and so that's one thing that we think is happening.

Another thing in the CHTS is one of the things we noticed, when we look at it, the CHTS estimates, households either report shore fishing or they report boat fishing. A very low percentage of households report both, which doesn't seem to make sense to me. In the FES, it's about -- It's split pretty evenly, and so about a third of the households report boat fishing, and about a third of the households report shore fishing, and about a third of the households report both types of fishing.

We think -- One of the big differences between the FES and the CHTS is that the CHTS is a cold call. We're asking these questions of whomever picks up the telephone, and they don't have a whole lot of time to think about it. In the FES, it's a mail survey. Somebody opens it up, and it sits on the kitchen counter for a while, maybe, and somebody walks by and sees it and thinks it looks interesting, and so they fill it out, or somebody opens it, and they're sitting around the dinner table, and they say, oh, I got this really interesting survey and two-dollars, and so I want to have a really serious discussion about our fishing activity, and so you can actually talk about it. There's time to talk about it. You don't have to send it back right away. You have a little more time to think about it, and so we think that provides a better recall experience for the respondent than the telephone survey, and so that's one thing that we think is going on, that there's just more time to think about maybe the trips that are a little less memorable.

The other thing is I mentioned that it was very common in the CHTS for people to tell us about one or two of their fishing trips, and then they're like, I'm not going to do this ten more times, and I'm not going to tell you the fishing mode, where I fished, whether it was a private-access trip. I'm not going to go through that whole thing again, and they just either say I'm done, I quit, or they say all my trips were exactly the same, and so there were opportunities for them to exit the interview before they told us about all of their fishing activity.

If people are more likely to remember a private boat fishing -- What we do in that situation is, if they only tell us about two trips, but they took ten trips, we imputed all the rest of their trips to be like the ones they did tell us about, and it was a hot-deck amputation approach, and the level of imputation was imputing from the information they did tell us about.

If they're more likely to remember a boat-fishing trip than a shore-fishing trip, they're going to be more likely to tell us first about their boat-fishing trip and then their shore-fishing trip. They're going to get tired after telling us their boat fishing trip, and they're not even going to get to their shore fishing trip, and so we're imputing -- More likely, we think, to impute boat-fishing trips than shore-fishing trips.

I think those are a couple of our hypotheses, and we haven't tested it, and we haven't figured out a way to -- Again, we don't know the truth, and so we just kind of suspect that shore fishing is an activity that people aren't going to remember as much as boat fishing activity, and so one of the comments earlier was you couldn't imagine not knowing what your family was doing, but I have a kid at home right now, and he's twelve years old, and he's staying home by himself.

I know where he is, because I can track his cell phone, but I don't know what video game he's playing right now, and so, if somebody called me and said how many times has anyone in your household played Fortnite in the past two months, I would guess that he probably played Fortnite, but I wouldn't know how many times, and so I'm pretty sure he's not fishing right now, but, if we had a dock -- If I lived on the water and we had a dock, I'm pretty sure he would spend a lot of time fishing, and so it's very plausible to me that the person who answers the phone really doesn't have time to think about what their family is doing, and they certainly don't have time to ask all their household members what they're doing, and they just might not know what they're doing.

DR. SHAROV: Thanks for a very, very detailed and highly-informative presentation. I think lots of us got a very clear understanding of principal sources of bias in CHTS design, and everybody is convinced that there are lots of sources for bias in it, and that was a dynamic process. It's pretty convincing, but, in comparing -- We still are comparing two theories. Therefore, FES, even though we probably would agree that the design is likely to or is doing a better job to reduce the biases, nonetheless, it's not a validation, and it's another survey, and so I think what many of the different questions alluded before is that what is required is an independent verification beyond the self-reporting system, be it at a smaller scale or a medium scale, but independent verification of the estimate of effort, either the private boat or the charter or the shore mode.

Maybe I am asking this question ahead of time, but are there any experiments of that sort that have been conducted in the past, or are they planned for the future, or maybe you have a couple to describe in the next half-hour, and I wonder, but I think that's a very important question. Thank you.

MR. ANDREWS: Thanks, Alexei. We talked a little bit about this just a few minutes ago, and it would be great if we could validate these estimates, and we are constantly discussing ways to do this. We have the research and evaluation team within our program, and we have a team of consultants we work with and discuss things with on a regular basis, and this is a topic that we often discuss.

We are planning a pilot study, hopefully for next year, that will be similar to the FES in design, in terms of sampling, but it's going to be a little different questionnaire, and it's going to focus on boat fishing activities. We're going to attempt to collect some additional information about types of boat fishing activities and locations of boat fishing activities, and so one of the things we want to do is try and get an idea of fishing activity that might -- I'm going to talk a little bit about this in the next couple of slides, and not this project specifically, but this notion of hidden fishing trips, and so fishing activity that just isn't obvious to most people, and so this is fishing activity of people sitting on their own private dock fishing or kayak fishing or paddleboard fishing, sort of the non-traditional boat fishing activities.

We want to get a handle on the magnitude of that, and I think that type of fishing activity is probably a little larger than a lot of people expect, and so that's one thing we're going to do. We

don't have anything else specifically lined up now. I mean, it's very challenging to come up with these validation studies in a cost-effective way that we can do on an ongoing basis, or even a small-scale basis, but we're constantly discussing this and trying to come up with ways to do that.

DR. SEDBERRY: Thanks, Rob. Kathy, did you have a question to this point?

MS. KNOWLTON: I do. Thank you. When we're in the field for the APAIS, and we intercept anglers, they have to be fishing. For us to derive catch information from their fishing trip, they have to be fishing for at least a minimum of half an hour, and we don't have a definition of a fishing trip when we receive the results from the FES questionnaire, and that's correct, right, and so we just simply ask them how many times did you go shore fishing and how many times did you go fishing from a private boat, and so have you guys thought about maybe any kind of pilot studies that can get at the difference between -- Not so much that's a difference in the definition of a trip, but it's a difference in the way that we access information from anglers between the two surveys, and so there is a portion of the anglers that, when we intercept them dockside, we don't ask them for their catch information, presumably because the trip was too short, and you assume it's an efficiency issue.

I can tell you, in Georgia, we run into anglers, fairly routinely, that say, well, I just threw it in for a couple of minutes, and the tide is running too hard, and I decided not to go, or I decided to go somewhere else, or take a break until the tide changes, and we don't know what that person ends up doing, and so your example of waiting to do something else at the dock or running down for a few minutes before dinner is ready, or you see the snook, it's a great example of there is a difference between what we're measuring in terms of catch that is then expanded by the total number of effort trips.

Obviously, the impact of that would be so much greater for shore, and I agree that people remember getting ready for a boat trip. They remember the excitement, especially depending on if they don't engage in that activity very often, like they're invited to go with friends for a boating trip twice a year, and they are probably going to remember that, and so I was just wondering if you guys have any plans or have discussions about whether or not the impact -- How to measure the impact of those short-duration shore trips. Is that something you can add to the hidden trips that would help elucidate this? I mean, everybody is just struggling, I think, much more with the shore estimate than the PR estimate and that increase, and so I'm just curious about that. Thank you very much.

MR. ANDREWS: Thanks, Kathy. We don't have anything specifically planned to evaluate that. I think that's something certainly worth considering. The challenge with a mail survey, one of the challenges with a mail survey, is, the more constraints you put on what you're really trying get, the less accurate information you're going to get, and so, if you say tell us about the fishing trips you took on Tuesday afternoons that were between 6:00 p.m. and 6:30 p.m., you really start to get some inaccurate information, and so it would be very challenging to get at that in the current mail survey, but certainly we could conduct a follow-up shore fishing survey just to characterize people's general shore fishing activities, to try and figure out what a shore trip really means and what it is, and I think that's a very good suggestion, and so thank you.

DR. SEDBERRY: Okay. We are running a little behind schedule, and so, as strange as it may sound, I'm going to propose that we take a break here, so that the organizers can figure out how

we can move things along a little bit, and so let's take a fifteen-minute break, ten-minute break, and come back at 9:30, and we'll move on to the plausibility section.

(Whereupon, a recess was taken.)

DR. SEDBERRY: We are going to try and move the agenda along a little bit. We're going to finish up Rob's presentation here, and then John is going to come back and talk about some calibrations, and then maybe we can get to Kyle's presentation that will involve one of the species and stock assessments of interest, and so, if we can get through these presentations, and then the idea is what are we going to do with all of this information, and we need to get to that point, and so I don't want to stifle any questions or discussion, but I really would like to move through this stuff as quickly as possible, so we can actually get to some consensus statements and some conclusions that will help move this whole process forward, and so, with that, Rob, I think you're on your plausibility.

MR. ANDREWS: Thank you. Okay. The past couple of sections and the presentation yesterday, we concluded that the FES design is less susceptible to bias than the CHTS design, and this isn't just our conclusion, and this isn't just an MRIP conclusion. This has been endorsed by MRIP consultants and independent peer reviews, as well as the National Academy of Sciences, but people still have questions about the estimates, and people have pointed them out today, that they wonder if the estimates are actually realistic, and particularly for shore fishing.

Part of the challenge of defending the FES estimates is that the perception of what is true is based upon the historical CHTS, which I think we have demonstrated pretty clearly were highly biased, and particularly at the end of the time series, and so that comparison isn't a valid comparison. Our basis for comparing FES estimates, or evaluating FES estimates, shouldn't be based upon CHTS estimates, and so that's a big challenge.

Despite the fact that the FES estimates are considerably larger than CHTS estimates, fishing is still a rare event among the overall population, and so these numbers from the FES are much larger than they were from the previous telephone survey, but it's still a pretty uncommon activity among the overall population, and so, in July and August of 2018, which is a high-activity fishing wave in most locations, and this was last year, about 3 percent of Georgia residents reported fishing, and the average angler took three trips.

About 5 percent of North Carolina residents reported fishing, and the average angler took five trips. About 5 percent of South Carolina residents reported fishing, and they took an average of six trips, and about 10 percent of Florida residents reported fishing, and the average angler took seven trips, and so we're talking about less than 10 percent of the population participating in fishing, and so it's still a pretty uncommon thing, even though the estimates went up with the FES, and this is during a high-activity wave in some high-activity states.

DR. SEDBERRY: Rob, we have a question.

DR. CROSSON: This isn't just for you, Rob, but one of the biggest concerns that we keep getting is that, or at least it's phrased this way in the letter that's Attachment 1 in the briefing book, is that something like -- Georgia said that, based off of these numbers, the average Georgia angler took something like fifteen trips last year, the average licensed Georgia angler took about fifteen trips,

and Florida came up with a number in their letter that they have, and it said something like they came up with eighteen or twenty trips for the average angler over the previous year, and the states found those numbers implausible.

I was wondering if you or your team had a chance to look at those letters and if the math is incorrect, if the states are calculating this in a way that's not accurate, or do you all have any thoughts on that, because I think that's one of the biggest issues that this committee faces, is that we are the SSC, and we're supposed to be advising the council, and one of the biggest parts of the council, of course, are the states that make up a significant portion of the representation, and they're really concerned about some of these numbers that are coming out of here, and so, whatever this committee does, in terms of accepting the FES, which we probably may well do, but we have to provide some numbers to the council, and I think that the states -- Given these concerns, this is something that keeps holding up this committee, and so I don't know if this is a good time to get into that, but, if you have any thoughts on that, I would like to hear it.

MR. ANDREWS: I am going to attempt to address some of those questions over these next several slides, and so I suggest maybe we hold that discussion until after I get through these, and, if I haven't answered any of those questions, or if there are still outstanding questions, maybe we can address them then. I am aware of those letters. I don't have numbers in my head. I'm not entirely sure where they come from, or what they were, what the source of those were, but we can have a discussion, I think, about that after I get through these next few slides.

Another challenge about defending the FES estimate is that the perceptions of fishing activity are based largely upon what people see when they're on the water, and this was actually alluded to yesterday in the public comment period, I think. If an angler isn't surrounded by boats or other anglers, if they don't see a bunch of other people out there, then they assume there is not that much fishing activity.

Given the size of the ocean and the number of fishing access points, the characteristics of fishing access points, it's possible that a large portion of fishing activity isn't even observable, or noticeable, to most other anglers. These, and I'm calling them hidden fishing trips, may actually be the predominant type of fishing in some areas, and so it's really obvious, when you're out on a reef or something, and there is a hundred boats around you, you think there could be a lot of fishing activity, but a lot of fishing activity isn't quite as obvious. A lot of people don't have big boats, but they still like to fish, and so they take advantage of whatever opportunities they have, and it just might not be observable to most other anglers.

Here is a map of Florida, and this shows all the Access Point Angler Intercept sites in Florida, and it looks like there is pretty good coverage of the coast. There's about 1,500 access points in Florida, which is a lot compared to any other state, and these are generally publicly-accessible fishing locations, and many of them are high-use sites, where you expect to see a lot of fishing activity, but Florida has thousands of miles of coastline, and so these, in most areas, are not really densely populated, in most areas. There's a lot of access sites, but they are still, relative to the amount of coastline in Florida, it really might not be that many.

I am first going to focus in on a couple of specific areas within Florida, to demonstrate the potential magnitude of these hidden fishing trips that I mentioned, and so I'm going to focus in on this little box right here. This is Vero Beach, and you can see it enlarged over here on the right. There is a

little bigger view, and so, this area, it's about five square miles, and it includes five Access Point Intercept Sites from the MRIP dockside survey, and you probably can't see them very well, but there is three of them up here around this bridge on the north end of town and then another out here on the coast and then another out here, which I think probably should be pushed a little bit over to the right, but there is five Access Intercept Sites in this area of Vero Beach, which is actually quite a bit for a small area.

The highlighted areas are residential land parcels that are within fifty meters of saltwater, and so the center point of each of those parcels is within fifty meters of saltwater, and so these are essentially waterfront residential properties. They are all residential properties. Within this map, and so, again, it's only five square miles, there are more than 1,300 residential waterfront properties.

If you zoom in a little closer, you can see that pretty much every one of these properties has a dock, and many of them actually have a boat tied up to the dock, and I'm guessing a lot of these docks also have some chairs on them, and are probably used for shore fishing, whether it's just a five-minute shore fishing trip or kids going out for a day or whatever, and so we believe that these are a source of hidden fishing trips that wouldn't be obvious to many other anglers, and, yes, you might go up and down the canal and see a couple of people, but there's a whole lot of these things, and so the potential magnitude of these private access, or hidden fishing trip, is pretty high.

Here is one more example from Florida. I could do this all day. I had a lot of fun creating these, just zooming up and down the coast of Florida. We're going to go a little further south, and so this area down here, and I'm not actually sure how close this is to what this is, but I just drew a box somewhere on the coast there, but the zoomed-in portion of the map is Delray Beach, Florida. This has a land area of four square miles, and there are no intercept sites in this location, or in this area of the map.

Again, if you light up the residential waterfront properties, there is 4,361 waterfront properties just in this four-square-mile portion of Florida, and so it's a tremendous amount of potential fishing access locations, and there is over 500,000 of these properties in Florida, and so the potential magnitude of private-access fishing and hidden fishing trips is pretty big.

Now, Florida is a bit of an exception. I mean, the whole coastline was pretty much engineered to maximize waterfront property, and so I realize that not every state is like Florida, but we can look at other areas of the coast as well, and so this is Harkers Island, North Carolina, which is just inside of Cape Lookout, I believe, and this is -- The island is about five square miles, and there is 280 waterfront properties in Harkers Island, North Carolina, and, again, I have never been to Harkers Island, North Carolina, but you zoom in on these, and you see a ton of docks. This is actually -- These are actually inland waters here, and so it's not directly on the coast, and so there's docks coming out of all of these, and so there's a whole bunch of places where you can access the waterfront.

Then the last one, this is Elizabeth City, North Carolina, and this is about ten square miles of land area, but it's only probably a mile-and-a-half or two miles of linear coastline, and, just in this short coastline, there is 329 residential waterfront properties. Again, it's not quite like Florida, but we see similar things in all of these states. People want to live on the water, and they want to live on the water because they have a boat, and they like to access the water, and they like to fish, and so

we think people use these sites for shore fishing activities, boat fishing activities, whatever, and there is a ton of them, and so the potential magnitude of hidden fishing trips is really high.

I have a couple of other examples. Don't feel left out if I didn't include your state, Georgia and South Carolina. The access to this -- We pulled all of these information off the web, and the accessibility of this type of information varies a lot by state, and so, believe me, we tried finding it for your states as well, but we weren't quite as lucky. I did notice, flying in yesterday though, there is lots of marshes flying into Charleston, and then you see docks that stretch a quarter of a mile out to a little creek, and so I think there's a lot of these in Georgia and South Carolina as well.

That is what I have for plausibility, and we can have that discussion that you brought up now. If you really want to see a picture from Georgia, I can show you that as well, but it's in the back of the presentation, I believe.

DR. SERCHUK: I am fortunate to live in a condominium building in Bonita Beach, Florida, which is on the west coast, but I'm also an early riser, and my lanai overlooks the waterfront there, the beach, and, many times when I am up early in the morning, and I'm saying between 6:00 and 8:00, I can look out my lanai, which, again, faces the beach, and there are several fishermen there, and they're not there later on, because, later on, the beach becomes populated by swimmers or people that just want to go to the beach and enjoy themselves, and so the timing of seeing these anglers is related to the multiuse nature of the beaches.

Early on, the anglers are there, because they're not in anyone's way, and they can enjoy fishing without interrupting anyone else's use of the beach, but, if they don't survey at that time, or regularly survey at that time, you will never see these anglers, and I am wondering whether that's a consideration for many of the beach sites. Because they are multiuse, that the anglers prefer to fish at a time when they are not interrupting others, whose use of the beach is not compatible with fishing.

MR. ANDREWS: Thanks, Fred. I think I can provide an answer to that question, and, John, please feel free to follow up if I don't address it thoroughly, but we do -- The intercept survey is stratified by time of day, and so we have, I believe, three or four time periods, four time periods, that the day is broken into, and they cover all twenty-four hours of the day, and so we do -- Each of those time periods of the day, there is an expected site pressure, or expected amount of fishing activity, and so, those sites that are mixed use, we might expect that there would be a higher -- In our sample frame, higher site pressures during those activities where people weren't swimming and it was more likely to encounter fishermen, and they would have a higher selection probability. We do sample twenty-four hours of the day, and we stratify our sampling by time of day, and so we should be covering those trips in our intercept survey.

MR. FOSTER: The only thing I would add to that is -- I can't remember a specific example for Florida necessarily, but I do know -- The one that I can remember is Massachusetts, when we looked at -- We run checks to look at the distributions of the estimated fishing activity, or pressures, by site and within month and across the entire year.

We did see, sometimes, sort of an unexpected pattern, where there would be higher pressure assigned to a site, say for the morning and the evening time intervals, and a lower pressure in what we called the peak interval, where we would typically expect that higher level of activity, and,

generally, when those cases occurred, they were at shore fishing sites, generally for natural shoreline beaches, in particular, and so that does happen, and that is the case, and I just only can remember some examples from Massachusetts, but, yes, that is how we would handle it, and, again, it's our state partners, state agency partners, that are populating those fishing pressures within the site register.

DR. CROSSON: This is plausible, because, at least on the opposite side of the island from Fred, on the side that's on the Indian River in Florida, or any of these coastline areas, you fish at night, and you don't fish during the day. Fishing at night is far better. You put a light out there, and the bait fish come up, and then all the snook and speckled seatrout start hanging around, and it's very easy to haul in fish, and so you're definitely not going to see that going around on a boat. It's happening in the dark, and any of us from North Carolina know that it's the same thing for red drum and a lot of the other species.

I don't have a problem with some of the big increases, I guess, maybe with the shoreline fishing with FES. I guess my concerns are the increases on the boat -- I guess there is lots of different codes, but the boat-based fishing. I mean, is there -- I guess Florida mentions it in their letter, and they're talking about the Gulf coast, but they said that the number of private boat fishing trips increased drastically with FES as well, and is that something that -- Is that accurate, and is it something that, if so, there is any kind of explanation that we could get, I guess?

MR. ANDREWS: I think it's important to note that we are not suggesting that private boat effort increased. The baseline went up across-the-board, and so effort didn't actually increase when we switched from the CHTS to the FES. It wasn't truly an increase, but we just feel like we now have a better way to measure fishing effort, and so we have a more accurate estimate of fishing effort, and I kind of hope that the prior fifty slides would have addressed some of our rationale for why those numbers are higher.

I mean, we believe, and we're very confident, that the CHTS estimates were very biased and were not an accurate representation of fishing activity and that the larger estimates from the FES are more accurate. I guess I need to rethink some of those slides, maybe, but, again, it's not because we think more people were fishing this year than last year. We think there were more people fishing in all of those years and that CHTS just wasn't covering them, because it did not provide a representative sample of residents.

MR. WILSON: I've got two things. The first thing is the site register, where the state port agents will try to predict what the average pressure would be at a site, based on a series of categorical slides, and, when we were doing the pilot, some of the lower-pressure sites, the availability of our staff and then having those at a very low level, was still producing sample draws that exceeded our staff capacity, and so there was some manipulations on some of the lower end of those categories, but John is absolutely right.

When we were going through the pilot -- Even now, we try to screen through and clean those up, but we were very poor at estimating what pressure was going to be at those sites, because what the count showed is sometimes we were less, and sometimes we were more, and very rarely were we dialed-in, and we had a discussion early on that it may be better to have less categories, and say we have like a low, medium, and high, and maybe we would hit those pressures, because that weighting -- The reciprocal of that site register weight, is that applied to the estimate?

MR. FOSTER: The selection -- The short answer, I believe, is yes, but there is a lot of technical detail that needs to be unpacked to answer that, and so the values that I presented, as far as those site measures of size or size weights, those values themselves are not the actual weight values. The calculation involves the actual probability proportional to size sampling process, sample selection process, as well as how many actual samples are being selected, and so there's a bunch of different components that go into the actual first stage weight, but there's a lot more detail than what I presented, and it's available in the reference document, the MRIP sample design and estimation document, that gets into the really detailed level calculations for the sample weights, but one thing I will say, just to, again, reiterate it, is the methodologies that we're using are not -- Two things.

One, they are just standard practices from sample theory, standard sampling estimation procedures, and it's not anything that we sort of specifically developed that isn't tested or unproven. Everything we're doing is standard, and then the other thing I would say is that none of it is relying on sort of modeling or model-based assumptions, in terms of how we calculate those sample weights. It's all following standard -- It falls directly from sampling theory.

One last point I guess I would say is, while the better the measures of size are, the better that those pressures at the sites are estimated, the more efficient it makes our sampling, but, if the pressures are not exactly right, that doesn't equate to biases in the estimates. The measures of size don't have to be exactly right to get unbiased estimates. We just have to know what those measures of size are that were used in drawing the sample, and we have to use those and calculate the sample weights.

As long as we do that, that is the requirement for getting unbiased sampling, again, in standard design-based sampling, and so, again, measures of size, those estimates of fishing pressure, they don't have to be exactly right to get unbiased estimates. We just have to know what they are, because they are used, in part, to calculate the probabilities of selecting units, and it's those probabilities that we have to use. As long as we can calculate the probabilities correctly of a sample unit being drawn, that will give us the correct weight that we need for unbiased estimates.

MR. WILSON: So the previous fifty-four slides state -- I was paying attention, and it's like the one that hit more like a hammer was the old survey just wasn't the covering the population, right, and so, if you track the old CHTS up to the point where it was covering the population, from the 1980s to the late 1990s, fishing effort was increasing the whole time, but then, in 2000, and I don't know if this is an artifact with the census data or what, but, across-the-board, there was a huge increase in 2000 of fishing effort, and so then it started increasing more to the time we had the economic turndown, and I don't know if this is a revisionist history type deal, because I remember seeing some of them CHTS estimates, where, at that time, I thought we should see some reductions in effort, because of people were losing expendable income, but these American Community Survey graphics you had -- You were showing some going down because of those, but our public information office, at the time -- Every year, we do landings updates and how was fishing last year, and it's like, well, we had a stock market crash and a bubble burst, but, yet, we're going on with business.

There was a series of articles written across the country that, during these economic downturns, people were using fishing as a way to spend family activity, and, actually, fishing trips were going

up during the economic downturn, because that was a cheaper activity for families to take, but I don't know if those articles were written because they were using the CHTS-based estimates showing fishing trips were actually going up, and so I don't know if it's an erroneous conclusion back in 2008.

Then I was looking at the thing about -- I think private boat -- I don't disagree, and I have never disagreed, with fishing effort that we were underestimating with CHTS, and I know that we were, and the twofold increase for private boat, just because these graphs that you showed, and people that have these residences on the coast that would not have been included in that Coastal Household Telephone Survey, but they're coming down to these locations that have docks, and so I understand that completely. I understand it on the shore mode, but I think there is a mismatch, particularly on shore mode, with the off frame angler adjustment.

It's like the private boat, I think, is about a -- It's a very small adjustment, but I think the shore mode has a little bit larger, maybe 1.6-times adjustment, for North Carolina, and so -- But that doesn't explain the six-fold increase, and so, in looking, it's like -- The letter that came from North Carolina, it didn't doubt the numbers so much as it was asking to maybe explain these increases, which is what you've done.

The only thing that we had to look at was when we were working on the National Angler Registry requirement, and we had exempted anglers on ocean fishing piers, and so, at that time, we enacted a daily reporting effort on ocean fishing piers, and so one of the things that I was looking forward to, when these new estimates came out, was to be able to take the manmade structure and compare it to what we get from our fishing pier daily effort from the pier operators, but, because estimates aren't made at that level anymore, we had to use the domain analysis for just those fishing piers, and they do incorporate the weightings at the manmade and jetty -- The Mode F.

They do include the weightings at the Mode F, and so you can do the domain analysis on those, but, for North Carolina, there was like a tenfold increase from what the pier operators tell us that they sell, as far as pier passes every day -- It's a lot lower than what I can come up with an estimate for the domain estimates from samples just on those piers. That is the one thing that I'm just trying to figure out, and I know it's big, but the sixfold increase is quite large, and so that was the biggest -- When we're not representing the sample, and that helps me a lot.

MR. ANDREWS: I think we would very much like to work with you with that pier information and talk more about that, because I think we would be very interested in seeing the numbers you get from the piers.

MR. WILSON: When we did the agreement, I was sending it every month, and so it doesn't get used for anything, and so I started sending it -- I send it about three or four times a year, but somebody is downloading it, because, after I submit it, they will clear it off for the next one.

MR. FOSTER: Can I just respond quickly? I just wanted to mention, Chris, that the domain analysis -- We don't have a way to partition the FES effort estimates into private access versus public access, and what we really would want is something more -- Private versus public access is really just a proxy for what we actually need, which is trips at sites on the APAIS frame and trips at sites that are not on the APAIS frame, and that can be a very important distinction, in some places, depending on how people interpret private versus public access.

Anyway, so, the APAIS data that are collected, those have to represent -- Those are representing both the weights that are in the public use datasets -- Those records are having to represent both the public access trips and the private access trips, and so, when you do a domain analysis where you essentially identify all those trips that are taking place at those specific piers, they are representing not just the trips at those piers, but also the rest of the manmade trips, fishing shore trips, on manmade structure, and that will include a lot of private access trips, and so that's -- The short answer is we just can't get a comparison with the data that we currently have, I think, to do what you're trying to do.

MR. WILSON: Thank you, and it's like the hidden trips -- I was calling them shadow people, because we're trying to explain to our staff what could be happening, and so at night, if you go across a bridge in eastern North Carolina, and there's people out there. We don't send our staff on bridges to interview people fishing, because it's dangerous, and so, every one of them, it's like you see people fishing along the marsh and stuff that we don't sample those people.

Then, as far as the ability to match them to put them in your strata, I know the pilot showed anywhere -- Depending on which state it was, it was like 30 to 70 percent non-compliance with license requirements, and so a case of, well, maybe there is a loss in efficiency for the license match strata, because there's such a high non-compliance with licenses, and so, with marine patrol, I asked if they had a number, and they don't have a systematic or a sampling type. How many people they interact with, how many people they have licenses, it's like 15 percent non-compliance, but it's not a number that -- I don't think we included that in the letter, because it's just a number, but I think, if there's a way that we can learn more about the private access component, or the magnitude on this shore mode, it would help further explain a lot of stuff.

DR. SEDBERRY: Okay. Are we ready to continue on? I would like to remind people, when they have a question or a comment, please identify yourselves, for those that are online and for the transcript as well. Thanks.

DR. VAN VOORHEES: I just wanted to make a comment that goes back to Scott's points about the letters earlier. I just wanted to say that some of the calculations in the letters that we got were based on the number of license holders as being an indication of the total number of people fishing in the state, and I just thought that it would be helpful, perhaps, to share with everybody what we do know about the amount of trips that get reported by households that don't match with license holder addresses. Rob, maybe you could say something about that?

MR. ANDREWS: I can't say much beyond what Chris just said. I don't have the numbers in front of me, but we can produce an estimate of fishing effort for households that do not match to a license list and look at the relative contribution of those trips to the total fishing effort, with a few very large caveats. First of all, our license frame is compiled about a month before data collection starts, and so, for example, for Wave 4, we provide our license list to the data collection contractor at the beginning of August. The data collection doesn't start until the end of August, and so the license list doesn't include anybody who would have gotten a license in the second month of the wave, and so keep that in mind.

We estimate -- There is a lot of variability along the coastline, and the licensing requirements, the cost, things vary quite a bit, and so I can only provide you some general information with what I

have in hand, and so we estimate anywhere from -- Chris said 30 percent to 50 percent of fishing effort, or 70 percent, may be unlicensed fishing activity, and I think, the best-case scenario, we estimate that about 50 percent of fishing activity is by licensed anglers. Not licensed anglers, but by households that we have matched to a license database, and it can be as low as probably 20 percent, in some states, and it varies by fishing mode.

For shore mode, it's a much lower percentage of effort by licensed households, or households that match the license database, and it's a little better, quite a bit better, for private boat fishing, and so, I guess, generally, I would say anywhere between -- We estimate that anywhere between 50 and 80 percent of fishing trips are taken by households that we don't match to a licensed angler.

DR. CROSSON: There are a lot of exceptions, and I was wondering about that as well. There are an awful lot of exceptions in Florida, and so, if I go fishing with my father up in the Port St. Lucie area, he doesn't need a fishing license, because he's over the age of sixty-five, and he fishes a lot.

If I'm fishing off the dock with him, I don't need a fishing license either. If you're fishing off of a manmade structure in Florida, and you're a Florida resident, you also don't need a fishing license, and so you're not going to be included in that database. I would expect that there are going to be a lot of biases with that, and so I don't know how Florida accounts for that when they deal with these numbers, but those are probably significant, especially considering there is a lot of retirees in Florida.

MS. SAULS: You're not matching to the out-of-state license holders, which I don't know the numbers off the top of my head, but that's a significant portion of license sales in Florida. When you say under-coverage, does that include people who are in the license database, but -- I don't understand.

MR. ANDREWS: I didn't say under-coverage. This isn't a coverage issue. It's a stratification issue, and so we are matching Florida addresses to Florida resident license holders, and we just using that to stratify the address sample in Florida. For that portion of the sample, about -- In Florida, I think it's about 50 percent of fishing activity comes from the unmatched stratum.

MS. SAULS: I just wonder if some of those households are non-resident Florida fishermen.

MR. ANDREWS: If they were non-resident Florida fishermen, they wouldn't be in the address sample for Florida. They would presumably have an address somewhere else.

MS. SAULS: They own a house in Florida, and I don't know. I mean, I have a lot of questions, I guess, about how that census is matching up to the out-of-state angler adjustment and what's being included in that, or not included, and I don't -- In Florida, we, and this was addressed in the Florida letter, we are doing our own specialized reef fish mail survey of anglers who are signed up, and required to sign up, in order to legally harvest certain species.

We have worked very closely with the Office of Science and Technology on the design, review, and certification of that survey, and our effort estimates are about -- The MRIP effort estimates for reef fish trips on the west coast of Florida are about two or two-and-a-half-times higher than what we're getting through our mail survey, and so the only under-coverage we have in our mail survey

is people who are not signed up through our saltwater fishing license database to be included in our mail survey, and so do an under-coverage adjustment for that.

We have been concerned about is the difference between the FES and our mail survey due to some bias in our under-coverage, and so we're testing -- Right now, we've been doing some separate on-site surveys, where we've taken the Panhandle of Florida, where you pretty much have to go through one of the main passes in order to get out into the Gulf of Mexico to fish, and we're looking at that as a way to validate the mail survey estimate that we're currently doing in Florida, and our results are preliminary at this point, but they seem to be closely aligned with the Gulf Reef Fish Survey estimate, which is positive for us, because it tells us that we don't have a big bias in that under-coverage adjustment, and so we're encouraged by that, and hopefully we'll be presenting and publishing those results soon after this year, but I think that some of the expansions they're getting done really need to be looked at as a potential source of bias in the surveys that we're doing. I am less worried about under-coverage in the Gulf Reef Fish Survey as a license angler database survey, as opposed to a coastal household survey, now than I was before we did that work.

MR. FOSTER: A couple of things. One, we had proposed a study to try get at potential sensitivity in the license question, and I know you know that, Bev, and there was concern, legitimately, from the state partners in asking that kind of a question the way we wanted to ask it in the intercept survey, to try to tease out, coming from that particular aspect, looking at how accurately do anglers in the field, intercepted in the field, report on having a license, because, not only for the Gulf Reef Fish Survey, but also with the FES, if we collected that information in the field, we could look at calculating an FES-based effort estimate just from the license matched stratum and seeing how that compared to the way we calculate the FES estimates now.

Because of the concern about potentially negatively impacting response rates in the intercept survey, our state partners weren't comfortable pursuing that, and so we're still trying to figure out if there's a way that we can conduct that, again, to start teasing out sort of sensitivity issues in license questions, and general survey research, not just specific to fisheries, suggests that there are a number of different types of questions, particularly if they have sort of legal implications, where respondents are not necessarily comfortable answering it. Either they will refuse the question at a higher rate than other questions, or they will respond -- They are more likely to give the response that is socially desirable, or, in this case, it would be what should they be doing, whether they're actually doing it or not, because it's legal versus illegal.

That is one thing we would still like to do, coming from the license side of it. Coming from the other direction, coming from the FES side, we are, and I think Rob mentioned this earlier, we have a pilot that we are planning to implement next year, where we'll be looking at getting the area fished information from the FES, in addition to getting it from the intercept survey, the APAIS, with the thought being that, because we're only able to sample at public access sites in the APAIS, we might be -- There could potentially be differences in those area fished proportions between the public access sites and the private access sites, and so, if people are taking the kayak trips, the canoe trips, or just any sort of boat trips from their private docks, they may be more likely to be fishing inland, let's say, for example, or near shore, than they are fishing offshore, compared to the public access sites, and so, at the public access sites, there might be a higher rate of fishing offshore than at private access sites, people fishing from boats at their personal residences.

That is something that we do have planned for next year, and what that could mean is, for example, if we see that, for the FES, we get different rates of fishing by area, that's not definitive, but we will need to continue working on that, and it might be that, while the total effort estimate for say private boat fishing from the FES might not change, the distribution of effort among the areas, the different fishing areas, inland, state seas, EEZ, we could see a very different answer, potentially, from if we collect the data through the FES, which covers both public and private fishing, compared to the APAIS, which just covers the public access fishing.

That is one, and that could bring the MRIP sort of based effort estimates for west coast Florida offshore fishing more in line with what is being produced from the GRFS, and so I suspect there are issues that would move sort of -- Every survey has non-sampling errors in it, and it's just the nature of the beast, and so it's possible that the estimates could move in both directions if we could do all of those studies.

MS. FRANCO: Rob, to follow up on what you said, Georgia would be interested in seeing how it's broken down for the households that were non-matched to the license frame. That would be very interesting, and we understand that's probably not available in your presentations here.

MR. ANDREWS: Right, yes, and it's a very simple analysis to do, and I think I probably have it buried in my laptop somewhere, and so maybe even by the end of the day today I can dig that out and share it.

DR. CROSSON: For those non-matched surveys, you have both coastal and non-coastal counties. Are the ones in the coastal counties much more likely to produce fishing effort than the ones in the non-coastal?

MR. ANDREWS: Yes, absolutely. The majority of fishing effort definitely comes from the coastal counties.

DR. SEDBERRY: Okay, Rob. I think we're ready to resume.

MR. ANDREWS: I only have four more slides, and so we should be done by about four o'clock this afternoon. These are really great questions, and I am very much joking. I appreciate all the discussion, and so don't take that the wrong way. That was a joke, a bad one. I am going to really sort of switch gears here, and the next couple of slides are really more of a setup for a presentation that I believe Kyle or Erik is going to be giving on some of the assessment work, and so this is not -- It's a little different direction from what I've been talking about the last several slides.

I pointed out earlier that the trend in the CHTS estimates is inconsistent with other indicators of fishing activity, and so charter boat fishing effort, boat sales, rod-and-reel sales, a couple others, all of these indicators decrease modestly during the economic downturn, but then they quickly recover, and, in some cases, they recovered even beyond pre-recession levels.

We can use these indicators to adjust the post-recession CHTS estimates such that they reflect the similar post-recession recovery, and this is done by first indexing each series to a common starting point, and so, in this case, it's the year 2000, and a common starting value of one. Then each point within each series, including the CHTS, then represents the change relative to this index starting point, and then all series are on the same scale.

We then can calculate the average, quote, unquote, recovery by dividing the average index value from the final three years of each series, and so, in this case, 2015 to 2017, by the average indexed value from the three years just prior to the recession, and so 2006 to 2008, and then we can apply this, quote, unquote, recovery ratio to the CHTS, and so we took the ratio for each series and then multiplied it by the average CHTS indexed value from the three years immediately prior to the recession to estimate CHTS endpoints relative to each series, and so, basically, we took, for each of these series, the ratio of where they ended up to where they were right before the recession, and then we took the -- For each of these series, we took that ratio and applied it to the CHTS series right before the economic downturn to come up with where this endpoint would be for the CHTS had the estimates recovered and had they been consistent with some of these other series. Then we just sort of uniformly applied that ratio back to the recession level, and this is all described much more elegantly in I think some of the handouts you have.

Basically, it just -- It models CHTS effort to be consistent with what we would have expected, had CHTS effort been consistent with these other indicators, and so this is what it looks like for each of these series that we looked at, and so there was charter boat effort, rod-and-reel sales, boat registrations, and then maybe gasoline sales, and then this is what the average looked like, if we take the average of all four of those independent series.

The reason I'm showing you this is because this value, we believe, would be much closer to -- It's much closer to the FES effort than the actual CHTS value, which is way down here, and we think it's probably a much more accurate modeled prediction of what CHTS effort would have looked like had the survey pretty much not entirely collapsed, due to coverage error and other sources of bias, and I believe that Kyle and Erik looked at this series, to see what the impact would have been on the assessment, had this been the CHTS time series instead of the actual CHTS time series, and so I apologize for fumbling through that so badly.

I promise you this is the last slide with actual words on it, and so this is just kind of a quick summary, and I covered a lot of information, and we had a lot of great questions, but these are just some of the highlights, and so, over the past fifteen years, CHTS samples have become increasingly biased as a result of declining coverage. People don't have landline telephones, or telephones, anymore, and the people that do have landline telephones are not a representative cross-section of a population.

This bias resulted in a severe underestimate of fishing effort, and our results show that. The results from the FES show that. In fact, we can use the landline portion of the FES sample to come up with an estimate that was very close to the estimates we got for the CHTS over the last couple of years, and we think this is a really good indication that landline samples were severely underestimating fishing effort.

Screening errors in the CHTS, which we refer to as the gatekeeper effect, we believe also resulted in an underestimated fishing effort, and then, when you consider both coverage error and the gatekeeper effect, that explains nearly all of the difference between the FES and the CHTS estimates.

Despite the larger FES estimates, fishing is still a pretty rare event within the general population, and we believe that the potential magnitude of hitting fishing trips is really pretty big, and it's

probably a pretty big source of the difference between what people perceive as actual fishing effort and what we are estimating as fishing effort, and, if there's still any more questions, I would be happy to take them.

DR. SEDBERRY: Thank you, Rob. Any questions?

MR. WILSON: Everything is reflected past the downturn, economic downturn, but, going back to what I had just said earlier, and I would say that I just pulled this off a query, the economic downturn on CHTS was where we had our most effort. For North Carolina, the highest effort was like 2006, which had a little downturn, and then again in 2008, and we were at our highest level at that point. That is the only thing that is a disconnect with me on this.

MR. ANDREWS: This information is the aggregate level for the entire South Atlantic region. I think there is some question about -- I don't know, specifically, what the starting point and end point for the great recession was. 2006 to 2008 sounds like it's -- 2006 and 2007 I think was kind of on the early end of it, or maybe even before it, in some cases, and I don't know that the official start date was uniform across the country, and so I don't know that I'm really that surprised. I mean, I think this shows a peak in 2006 and 2007, and then the estimates begin to decline in about -- Or maybe they start declining in 2008 and 2009, and so I think there is probably a little wiggle room in when you define the beginning of the recession or not, and so, again, I don't have the numbers from North Carolina in front of me, and you do, but I assume they do show a downward pattern after that that continued through the time series, and are you looking at the calibrated estimates or the --

MR. WILSON: The base CHTS, and so the downturn for North Carolina occurred in 2009, but I just remember having to write press releases back in -- Because there was a bump in 2006, and everybody expected it to be lower, and it wasn't, and then there was a big thing in 2008. Everybody expected 2008 to be lower, and it wasn't, and so that's when I found out, across the country, people were taking more trips because of the recession. Because they couldn't go on regular vacations, they were staying close to home and doing fishing within their states. There's, I think, like three articles that were written at the time.

MS. SAULS: Did that apply to shore and boat fishing, because I remember when all of this was happening in Florida, and we were all talking about how there was a lot more shore fishing, but there was a lot less boat fishing, because it's expensive to have a boat.

MR. WILSON: It was more shore. It was.

DR. CROSSON: There was a drastic rise in the price of fuel in the year or two preceding the whole crash of the market in the summer of 2008, and I remember there are articles in -- I was working with you all. Remember this?

MR. WILSON: (Mr. Wilson's comment is not audible on the recording.)

DR. CROSSON: No, we did. We were trying to track the price of fuel, because it went upwards of $4.50 a gallon, and that's just at the pump, and, at the marinas, it would have been much higher than that, and I was driving back and forth across the bridge for the Atlantic Beach Causeway, and there were just no boats going out the Beaufort Inlet at that price, and there were all kinds of

articles in the paper, up and down the state, about how nobody was going out on their boats, and they were just fishing, and so I would expect that there was a huge increase in -- Or replacement, at least, maybe, with boat-based fishing.

MR. WILSON: To follow-up with that, what we saw at that time, I do recall, is I got our guide fleet -- The offshore contingent scaled back, and the guide fleet expanded, and so it roughly became, at that point in time, those smaller boats became at least a third of our fleet, and they have maintained a third of our fleet since then.

DR. SEDBERRY: Any other questions or discussion or comments?

MR. FOSTER: Scott, I think you had asked about some of the values that were in the letters, in terms of trips per angler per year, the effort estimates and things like that, and one sort of study that was done in 2014, independent of MRIP, and I will credit Jason Didden at the Mid-Atlantic Council for bringing this to our attention, but there was a report put out in 2015 for a survey done for 2014 by the Recreational Boating and Fishing Foundation and the Outdoor Foundation, again totally independent and unaffiliated with our program, where they -- There is a number of methodological differences, but they did a sort of a national study of a variety of different fishing, including saltwater fishing, and I will send this to Mike, so folks can see it, but there is some good agreement between the information that they include, or what they found for saltwater fishing in 2014, and what we're seeing now with the FES.

For example, I will just pull out one data point, if I can find it again, and, nationally, on average, they found that saltwater anglers took -- They went fishing on sixteen days, and so how they measure things are a little bit different, but, essentially, in 2014, on average, saltwater anglers fished on sixteen days, and so they roughly took sixteen trips, and so that's a national average, and so that will include people that are coming from interior states that are fishing just a few times a year on a vacation-type trip as well as resident anglers that will be taking many, many more trips, and so sixteen days as a national average, and then you compared that to what was calculated within the specific coastal states, and that's not that different.

Again, this is just one study, and it's got whatever issues it has associated with it, and I'm not saying that it's correct or the truth or anything like that, but it is at least similar to what is calculated for some of these -- Using some of the FES data, and so, again, I will send this to Mike, so that folks can take a look at it. I'm not trying to say that I endorse this in any way, but it's just another data point to sort of help inform the discussion.

DR. SEDBERRY: At this point, I would like to move on to --

DR. CROSSON: Can I ask John another question then, or Rob or whomever? I don't think you could answer this based off of my understanding of the MRIP methodology, or FES, but, in the literature that you're familiar with, is it a relatively small proportion of anglers that are driving a large portion of the trips? Is it kind of the outliers, the guys like my father that probably fish 200 days a year? Is that what is -- I am thinking that anytime I see a mean that I start getting careful about what I'm going to think about that.

MR. FOSTER: Sure, and so, when we look at the distribution of trips reported by angler, they're all sort of kind of Poisson distributions, where the largest numbers of trips reported are the 1s and

2s and 3s, and then there is a tail, and so we don't see sort of any usual pattern with like a big disjointed distribution, where there is some influential number of anglers taking, reporting at least, very large values of trips and then a break in the distribution. Again, it's a fairly smooth distribution, with the majority of anglers reporting a small number of trips, and then, yes, there is a tail, but, again, it's not an extreme distribution. It's not something that looks concerning or unusual.

DR. DUMAS: My first question is about the hidden trips, and so Rob said that, through the end of his summary slide, there could be an enormous number of hidden trips, and so my question is whether the catch per trip for those hidden trips would be similar to catch per trip for the non-hidden trips, the observed trips, and, if the catch per trip for the hidden trips is much different, it could be much lower than -- How does the potentially large number of hidden trips affect the estimates? I'm not sure.

Also, these hidden trips -- Would the extent to which the hidden trips be captured by the FES survey be different from the extent to which the hidden trips are captured by the Coastal Household Telephone Survey, the hidden trips? I wasn't clear on that, and so that's sort of the first question. How does coverage of hidden trips vary between the two surveys, and how does catch per trip for hidden trips affect the estimates?

MR. ANDREWS: Both the CHTS and the FES would cover trips that would not be encountered through the dockside survey. Was everything else equal, we would expect them to provide equal coverage of those trips. Now, they're not going to be covered, provided the same coverage, because of all the problems with the CHTS, but those trips would be included in the offsite surveys, and they would not be encountered in the dockside survey, and so we would have expected these hidden trips to be part of the CHTS estimates as well.

In terms of the -- Yes, catch rates for these hidden trips could be different from the trips that we do sample from. They could be lower, but they could also be higher. We don't know, and that could have an impact on the estimates. We make the assumption that the catch rates are the same between the trips that we do encounter in the dockside survey and those that we don't.

DR. DUMAS: Thanks. My second question is that, in the summary slide, also another bullet is you mentioned that much of the difference, or close to almost all of the difference, between the FES estimates and the Coastal Household Telephone Survey estimates could be accounted for by coverage error and screening effect error, the gatekeeper effect error, but, from my notes, the coverage effect was maybe say two-and-a-half-times -- The coverage effect could account for the FES estimates maybe being two-and-a-half-times larger than the Coastal Household Telephone Survey. The proportion of times larger varied shore mode versus boat mode, but about two-and-a-half-times larger.

Then the gatekeeper effect was an additional 25 percent to 40 percent larger, the FES estimate could be larger than the Coastal Household Telephone Survey, and so, if you combine those two together -- That depended on whether it was 25 percent or 40 percent, and it depended on whether you were looking at all households or landline-only households for the gatekeeper effect, and so, if you assume a 2.5-times effect due to coverage error and a 25 percent effect due to gatekeeper error, then you get that the FES could produce estimates 3.1-times larger than the Coastal Household Telephone Survey.

If you look at the landline only for the gatekeeper effect, then you would calculate that the FES could produce estimates 3.5-times larger than the Coastal Household Telephone Survey, and so, based on the coverage effect and the gatekeeper effect, I am looking at FES estimates being maybe three to four-times larger than the Coastal Household Telephone Survey, based on those two sources of error in the Coastal Household Telephone Survey, and so that could explain -- Those two sources of error could explain why the FES estimates are larger than the Coastal Household Telephone Survey in locations or times when the difference is about three to four-times, but, in other locations, where the difference between the FES and the Coastal Household Telephone Survey is much larger than four-times difference, then we're still looking at potentially other sources of error or other issues causing those differences that are larger than three to four-times difference. Would you agree with that assessment, or am I off somewhere in that thinking?

MR. ANDREWS: I guess my response to that would be all of this is generalized to the South Atlantic region from pilot studies that may have been conducted on a more limited scale, in the case of the gatekeeper effect, and I presented this information at the sub-region level, and so the difference between the full sample estimate and the landline sample estimate may be bigger or smaller for individual states, and I don't believe I have that information available to me right now.

Then the gatekeeper effect was based upon an experiment that we did in 2012 in North Carolina, and I generalized those results to the South Atlantic, and so there's a lot of assumptions that went into that one, but, again, this is -- The big one here is going from the FES full sample to the landline sample. What I presented was for the South Atlantic as a whole. Those ratios, or the percentage changes, based upon going from the full sample to the landline sample is going to vary from state-to-state.

DR. DUMAS: So you're saying that the size of the difference between the FES and the Coastal Household Telephone Survey, when those differences are larger than three to four-times, could be due to the fact that, in different specific locations, the gatekeeper effect could be larger than the effect that we discussed earlier, for particular locations. The gatekeeper effect could be larger, or the coverage effect could be larger, than sort of the average values that we discussed earlier.

MR. ANDREWS: That is correct, yes.

DR. DUMAS: Thanks, and so the third question is a question for John, and that is you discussed a few minutes ago that the trips per angler followed roughly a Poisson-shaped distribution, instead of say a normal distribution, normally bell-shaped distribution, and so, if trips per angler follows an asymmetric-shaped distribution, like the Poisson, then that could have implications for identifying outliers, in that we discussed yesterday that an outlier, in terms of catch per trip or trips per angler, if we're identifying outliers as those values that are more than two standard deviations away from the mean, then the portion of the distribution cut off by two standard deviations from the mean varies depending on whether your distribution is a normal distribution or a Poisson distribution?

Whether something is considered an outlier or not would depend on whether you're assuming the underlying distribution is normal or Poisson, if your trigger for calling something an outlier is whether it's beyond two standard deviations from the mean, and so that could become important

when we discuss outliers, if we discuss outliers more later on, and so, John, I would be interested in your thoughts on that.

MR. FOSTER: Sure. One thing I would say, quickly, is that, while the underlying data may have a distribution that isn't normal, in design-based estimation, the sampling theory is in the context of repeated sampling, and so, if we -- That is how expectations for the estimators are -- That is the context for calculating those and evaluating those.

In the context of repeated sampling, the distribution of the estimates would have a normal distribution, and so the distribution of the estimates. If we could take 1,000 repeat samples from the finite population that we're trying to produce an estimate for, which would be both spatial and temporal in how it's defined, we would get a normal distribution of the estimates calculated from those 1,000 repeated samples. However, the underlying data of the population might still be distributed roughly Poisson or negative binomial or what have you, but the estimation theory works so long as the distribution of the estimates themselves, in the context of repeated sampling, is normal, and we would see that.

The other thing I would say is, when I was talking about how we sort of identify outliers, that's more of a sort of a QA/QC procedure. It doesn't -- When we identify, when we flag, those estimates for review, it doesn't affect how we do the estimation, or we don't necessarily make -- We don't make automated adjustments to the estimates, but it just flags those cases for review, and, ultimately, things may be done and estimates re-run, but we don't use any sort of model-based or distribution-based adjustments. We don't make anything like that, based on whether an estimate has been flagged or not, and so you're right, in the sense that we have made a decision about two standard deviations being sort of thresholds by which we flag things for review, but that's just it. We're just flagging them for review, and we're not making some sort of a model-based adjustment to the estimates based on that two-standard-deviation threshold.

DR. DUMAS: Thanks. Yes, I agree with you that the distribution of an estimator would be normal, in a repeated sampling sense, if you have a large number of samples, and I agree with that. My question is just that we need to be careful and make sure that those assumptions of normal distribution hold for any particular estimate that we're looking at, and, in many cases, the assumptions will hold, but we just want to be careful that, in some cases, they might not. Then your second point, which was -- Now I can't remember. What was your second point?

MR. FOSTER: Just that we're using the two standard deviations to flag potential outliers for review, as opposed to doing any sort of adjusting to the estimates directly.

DR. DUMAS: Right, and so you're flagging your outliers, but then, if you determine it's an outlier, and the outlier was not based on an error -- If it's an error, then you correct the error, but, if it's not an error, and it's still an outlier, then it's still retained in the estimations, still retained in the analysis, and so that, quote, unquote, outlier data point could still have influence on the estimates, because it's not dropped from the analysis. It's still retained in the analysis, and that could be a situation where the sample size is low, the sample size is small.

MR. FOSTER: I guess one thing I should clarify is that that analysis is not done on the actual data points, the recorded survey raw data, or microdata. It's done on estimates and estimate components, and so we're not --

DR. DUMAS: Thanks.

DR. SEDBERRY: Okay. I believe we're ready to move on to -- John is up next. This is the calibration. The next presentation, again addressing TOR 1A, is an explanation of the two calibration methods that will be presented by John Foster. This is Presentation Number 4.

SURVEY ESTIMATE DISCREPANCIES

MR. FOSTER: Okay, and so this presentation covers the two major calibrations that we have made to the time series to account for both the change from the CHTS to the FES for effort estimation as well as the design change that we made to the intercept survey in going from the MRFSS intercept to the new APAIS design.

I am going to give a little bit, just a very few slides, on the sort of background for the calibrations, and then the presentation contains detailed sections for the FES and APAIS, and then, at the end, I will sort of touch on some summary, example summary, information for the effects on the time series, but we will be going through, in a separate presentation, much more in-depth on the five species that were identified in the terms of reference.

In an attempt to help us sort of get back a little closer to on track with the syllabus, or the agenda, excuse me, I am going to just give a very quick description of the FES calibration, because Jay Breidt from Colorado State will be presenting a detailed presentation on the FES calibration as well later, and so this would cover a lot of the same information, and he will be much better equipped to answer a lot of detailed questions on the FES calibration, and so, if that's okay, that's what we'll do, but you have all the slides in the presentation as well.

Getting started, I'm going to skip that slide. You know why we need to calibrate, because things changed, and we want to make it consistent across time, and so just some high-level summary information, and so, at the FES calibration, we had a three-year benchmarking period, where we were able to conduct the CHTS and the FES side-by-side in all states and all waves within those three years where the surveys were conducted, and that allowed us to build a statistical model to relate the differences between the two estimates, between the two surveys and a series of estimates.

That is kind of an ideal way to do the calibration. Now, clearly, having only three years of data for the benchmark is a limitation. We would have liked to have had considerably more years of data, but it's expensive to run these two surveys side-by-side, essentially at production levels, although the FES was a little bit reduced, in terms of sampling, compared to where it is now, but that is essentially the most that we could afford to do.

For the APAIS calibration, it's a little different, and so we didn't -- There was one pilot done in North Carolina for a full year that ran the new APAIS design against the existing MRFSS intercept design at the time, but that really wasn't sufficient data to try to build a calibration model that we would then apply to all of the other states, and even really just for North Carolina as well, because changes were being made within the year as the new design was essentially being piloted and adjusted to help make improvements to it. The APAIS calibration approach uses a raking -- It

makes adjustments to the sample weights using a fairly extensive raking procedure, or process, and I will be walking through this one in detail today.

Both of these methods were peer reviewed using -- Well, first, they were developed with support from some of the MRIP consultants that we have from Colorado State. The FES calibration was basically developed directly by our statistical consultants at Colorado State, with Dr. Jay Breidt and a student and now I think post-doc, Teng Liu, and Dr. John Opsomer worked with us on the APAIS calibration, although we did more of the development and coding of that one in-house.

They were both peer reviewed, and the FES calibration was peer reviewed in 2017, and the APAIS calibration was peer reviewed in 2018, and the reviews were sort of coordinated through the Center of Independent Experts, and they identified independent reviewers for us, and then, of course, the council SSCs were represented as well as ASMFC, and I would like to thank, again, Fred and Carolyn for participating from this council SSC in those reviews. The consensus position was recommending the use of both calibrations from the reviews.

DR. SERCHUK: I know you mentioned that you were limited to three years, and I can fully understand that. One of the recommendations that I think came out of the peer review panel, or at least one of the recommendations that I made, was to compare the estimates after the first two years of data went into the calibration versus the estimates using all three years' worth of data, to give a feeling of how much an additional year's worth of data actually changed the calibration estimates.

I don't know whether that was done, but I feel that would give some inkling, in terms of the value of doing an additional year, or if there was slight, only slight, changes between the calibration estimates using two years' worth of data versus three years' worth of data, and one might say, well, gee whiz, it would be expensive to do a fourth year, but we didn't think there would be enough change, based on looking at the two years versus three years' worth of calibration. Do you understand my comment?

MR. FOSTER: Sure, Fred. Yes, I do understand, and I will -- I think I will leave it to Jay to make sure that we revisit that, when Jay is able to present. I guess what I would say at this point is I believe we did that sort of sensitivity analysis, and I can't remember all the details, but what I would say is that the three-year window is -- Whether it was two or three years, it's a limitation of the approach, and not only just because of the number of years, but also where those years occurred, in terms of the underlying changes that were going on and the coverage of the CHTS.

It would have been very helpful, in terms of the calibration, to have had years prior to the sort of wireless impacts on the CHTS, prior to folks abandoning landline telephones for cellphones, in addition to years during the decline and then years when we actually did the benchmarking, which were after half or more of the country had already essentially abandoned landline phones, and so that is a limitation of the calibration.

One thing that we will be doing, at some point in the future, after we have more years of data collected with the FES, is to revisit -- Not that we would necessarily revise all of the estimates, but revisit the calibration again and assess how things may have changed after we have more years of mail data to feed into the model, because, essentially -- Well, I'm about to start explaining the

whole calibration. Let me just leave it at that, and we can revisit it when Jay is here to present, if that's okay.

DR. SERCHUK: Thank you.

MR. FOSTER: Okay, and so just a very, very quick summary of the FES calibration, and so the calibration is essentially done -- The model used for the FES calibration is essentially a very standard model used in small-area estimation, and it's known as the Fay-Herriot Model, and it was originally developed to provide small-area estimates for income, down to sort of county and municipal-levels, from census data. It is essentially a mixed-effects model, a linear-mixed effects model, that has both, again, both fixed effects as inputs as well as random effects, and I'm not going to -- Again, at this point, I'm going to just sort of skip over the details, and Jay can fill us in, but that is the basic framework of the model.

It also has a bit of a time series flavor to it, and it has inputs that are used to model trend, long-term trend, and it has inputs that are used to model seasonal patterns and effort, using the MRIP two-month waves as inputs, and then it's allowed to vary by state, and it also has separate model runs for private boat versus shore mode effort, and so there are individual -- The effects of the calibration can vary by state and year and wave and fishing mode, and so there is -- We can summarize sort of general effects, but, again, in any specific case, the calibration results can differ, can be larger or smaller, than the overall average, and some of that will come out when I'm presenting on the calibration effects on the five species that we'll go into more detail on.

I apologize for just barely scratching the surface of this, but I'm going to, at this point, sort of fast-forward to the APAIS calibration. This was a more -- For a number of reasons, this was a bit of a messier problem, or challenge, to calibrate. The slide here sort of has a timeline of the status of the estimates and the designs, going from the beginning of the program back in 1981 through 2013, Wave 2, when we implemented the new design, and so, from 1981 through just the very beginning of 2013, we essentially had the original MRFSS intercept -- Data collected under the original MRFSS intercept design and unweighted estimation, the original sort of estimation approach that was used by MRFSS, which made a number of simplifying assumptions, and did not account for the complex nature of the design.

In 2011 and 2012, we started -- One of the MRIP projects was to try to incorporate weighted estimation using the data collected under the old intercept design, and we found that, as we went back in time, more and more of the information we needed to try to produce sort of pseudo-weights for the data were just missing, and there was less and less information that we needed, such as the detailed information on the site pressures used in the APAIS draw, or the MRFSS intercept draw, and there was missing information on counts of angler trips during sampling that were not interviewed, were just missed, and even, as we go further back in time, there was not information on just the number of assignments even drawn, as we went further and further back.

When we looked at what we had, we felt that we could do a sort of pseudo-weighted estimation, re-estimation, from 2013, Wave 1, going back to the beginning of 2004, and so we produced an additional series, a revision, of the estimates, to account for -- To attempt to account for weighted estimation with the data that had already been collected.

In hindsight, this probably just created a disruptive interim step, but we were thinking that, if we could do this weighted estimation, it would account for a lot of the potential biases in the estimates that were there because of the sort of simplified assumptions that were used in the simple MRFSS estimation approach, and so it was a reasonable thing to do, but, again, because it didn't fully account -- Once we implemented the new design, we still had systematic differences between these re-estimated series and the estimates produced from 2013 Wave 2 forward that used the new fully-design-based intercept design. Anyway, the bottom line here is that we have a number of different pieces and time series that we are trying to standardize to what we have from 2013 Wave 2 moving forward.

Again, the general approach here is to adjust -- To attempt to do this calibration by making adjustments to the sample weights in the intercept survey data. By making adjustments to the sample weights, we can then re-calculate all of the estimates and have a calibration that has essentially been in place for all of the catch estimates, and we don't have to develop sort of a species-by-species or estimate-domain-by-estimate-domain approach to try to do the calibrations, which would be very hard to try to keep things consistent across the different species or across the different estimate series.

The way that we went about trying to adjust the sample weights was, again, using a raking approach, a raking ratio approach. Again, it is a very standard approach in survey methodology, and I will be walking through more of the details of it, but we essentially did this sort of in a sequential approach going back through time, and so that's what -- This slide is a little busy, but that's what it's trying to show, which is, starting with what I will call the new estimates, and these are from 2013 Wave 2 forward, we have the fully-revised APAIS design in place, and we have sample weights that are completely design-based, and there is no modeling assumption, and there is no missing pieces, and this is the appropriate way to do the sampling and calculate the weights and do the estimation. Then we want to, stepping back through time, make the estimates as comparable as we can by, again, adjusting the sample weights using this raking methodology.

Going back through time, the next piece is this period of time where we have these re-weighted, pseudo-re-weighted, estimates from 2004 through the very beginning of 2013, and, essentially, the -- This is sort of a simplified way of showing it, but we're essentially trying to get the -- The raking uses a ratio approach, where we have the control information in the numerator and then the estimate from the time periods that we're trying to adjust in the denominator, and what these estimates are is not the catch estimates, but they the sum of the sample weights, and so what we are essentially trying to do is -- In a number of different dimensions, we are trying to get the sums of the sample weights to line up between the different time periods.

I know that may sound like that's not the right thing to do, but I will talk about how we don't expect those to be constant over the full time period. We allow for expectations that the sample weights may be truly changing over time, as fishing effort is increasing, moving from 1981 forward, and so I will try to explain how we are able to -- We took into account the need to allow the true sample weights to change over time, but the bottom line here is -- All I'm trying to say is that we essentially did this in a sequential fashion, where we sort of chunked the full time series into roughly ten-year bins, and, starting with the current time series that didn't need any adjusting, we used that as the control information to then adjust the period of 2004 through 2013 Wave 1.

Once that was calibrated, we then used that as the control information to make adjustments to the next piece, which was from 1993 through 2003. Once that was calibrated, we then used that as the control information for the earliest portion of the time series, from 1981 to 1992, and we calibrated it, and so, again, it's sort of sequentially stepping backwards through time and doing these raking adjustments to the sample weights. Once the sample weights are adjusted, we would then re-estimate all of the catch estimates using those newly-adjusted calibrated APAIS intercept survey sample weights. I know that was hard to follow, and there will be more details in the next slides.

For the first adjustment period, 2004 to 2013, there were sort of three main variables that we wanted to use to define these raking adjustment cells, area fished, household status, again, residency status, really, whether the angler trip was made by someone that lived in a coastal or non-coastal region of the state, and then the for-hire frame status, and this is relevant only to the charter mode, but the APAIS calibration wasn't just for private boat and shore mode. It had to account for all of the trips that are covered by the intercept survey, and so that includes the for-hire modes as well.

Then the raking was also done within cells defined by the sub-state region, and so, states like Florida and North Carolina and South Carolina, the APAIS sampling is now stratified by region, sort of groups of counties, coastal counties, within the state, and then by state and two-month wave and fishing mode, and so all of this raking was just done, again, within these different cells, and it wasn't -- We didn't do it across big groups.

The reason that these variables were selected, area fished and the sort of residency status and the for-hire status, was that these are variables that are used from the intercept survey in estimation. These focused on the sort of effort pieces that we get out of the intercept survey, but we wanted to make sure that, as we made the adjustments to the sample weights, they were going to consider these variables again, because they are used directly in estimation.

If we could have done -- We evaluated different ways to try to incorporate catch information as well, but we just -- There are just not enough samples to do that, and so, for raking to work, you need to have sufficient sample sizes in the different cells, and, if you think about trying to parse out cells by species, or even species groups, there is just going to be enough data. You will end up with a lot of data gaps, a lot of holes, and the raking procedure can't really account for that, and so we even looked at just presence and absence of catch for any species, and there's just not enough data to reliably incorporate a catch variable in this process.

In addition to these variables, when we stepped back to the earlier time periods, we had to account for some additional design variables, including kind of day, and so whether it was a weekend or a weekday type, as well as just a general proxy for site activity. Is it a high-pressure site, high-activity site, or a low-pressure and low-activity site? These were added because the sampling -- Even though we didn't have a lot of detailed information, the further back in time we went, on the design, we know that these were used in the stratification, or in the sample selection, in the earlier years, and so we wanted to account for it, the fact that, even in the earliest years, weekends were sampled at a higher rate than weekdays, and so we needed to account for that in how we adjusted sample weights, calculated and adjusted sample weights.

The same with the site pressures, the activity levels at the sites. That's always been used in the intercept design, going back to the earliest years, where the higher-pressure, higher-activity sites were sampled at a higher rate than the low-activity sites, and so we wanted to account for that, again, in this procedure as well.

Again, the sort of overview of how the raking algorithm works is we have an initial sample weight for an observed angler trip, wherever that came from, and, for 2004 to the beginning of 2013, it was these pseudo-weights that we had calculated after we did the re-estimation process. For the earlier time periods, where we were not able to re-estimate, it is essentially an equal weight across all of the trips in the domain, because we were starting from essentially assuming a simple random sample, and then we're going to correct that through the raking process.

The estimates here, these N-hat, these are essentially the sums of the sample weights within the cells that are defined by the raking variables and the things that I listed here, and so area fished, household status, for-hire frame, kind of day, and site activity class, and so, within these cells, we can sum up the sample weights for all of the intercepted angler trips, and the control total is what's in the numerator. That's coming from, again, the control reference period, and so it's the newer time period, and then, in the denominator, again, it's the sum of the sample weights within the adjustment cell from the old time period, or the one that we are trying to calibrate the weights for.

We make these adjustments, going one variable at a time, through the raking process, and that results in an adjusted sample weight, and that is then used to calculate the old estimate for the next variable, and so you sequentially move through the variables, doing these calculations, re-calculating the sample weight using those in the next -- For the next step in the raking iteration, and, eventually, after you have made a number of these adjustments, they converge, so that these ratios for each of the different variables end up converging to one, because, as the weights are being adjusted, these estimates get closer and closer to the control totals that are in the numerator, and so this ratio gets closer and closer to one as you iteratively move through the process.

As we talked a little bit about, what is this actually doing, and you're matching these estimates on the margins, and so we've got sums of the sample weights within cells defined by area fished, and we've got sums of the sample weights within cells defined by household status, and the same for the for-hire and within the other cells. What we don't have -- The reason this isn't post-stratification is we don't have all of the joints for those different -- We don't have area fished by household status by for-hire frame status within all of the cells. We don't have that. We just have the marginals, and so raking works on the margins. It doesn't do the joint distributions.

Here is a little bit about -- We realize that the sums of the sample weights from 2013 forward don't necessarily need to exactly equal the sum of the sample weights from 2004 to 2013, and we wouldn't necessarily expect those time periods to have exactly the same sample weight sums, because we would hope that, over time, as fishing effort is increasing, our sample weights actually should also be -- The sums of our sample weights should also be increasing over time, because the sums of the APAIS sample weights are an index of total effort, and so they should be roughly tracking our overall effort estimates.

We don't want to just assume that -- If we made them all match across the entire time series, that would just assume that effort is constant, and we know that's wrong. It's not constant, and so the way we allowed for things to change over time was by performing linear regressions of these

sample weight sums over time and then testing if the slopes of those regression lines were significantly different from zero, and, if they were significantly different, then, instead of trying to match the sums for -- Let me back up to this slide.

Instead of trying to make the sums of the sample weights over the entire time period match from one period to the next, we would just work on the three years at the beginning and end of the adjacent time periods, and so we're just trying to anchor the last three years, and that three years was sort of an arbitrary choice, but we did evaluate kind of stability versus including more years and trying to anchor the ends of the time periods.

Again, if we detected a significant trend in the sample weights over time, we would just use the last three years in doing the raking adjustment, and so, essentially, we're just going to make the sum of the sample weights in the last three years line up, but then we'll maintain the rest of the trend through that time period, and so that allows us, again, to do the calibration, but allow for changes over time in the series, and, without going into a lot of the details, in some cases, yes, we did see significant trends, and, in other cases, we didn't, and so it had the flexibility to, variable-by-variable, either have the trend or not have the trend.

I am going to stop there, and I'm sure there is a ton of questions, and there's a lot of details here that, in the interest of time, I didn't include, but I'm happy to take questions, and, again, in the MRIP design and estimation reference document, there's a lot more detailed sort of explanation of how raking works, more detail on our application of the method, and there's also -- In the report that was produced for the peer review, it also goes through this in much, much more detail, in terms of exactly how the method was implemented. I will stop there and take questions.

DR. SEDBERRY: I would like to open this up for questions, but I would also like everyone to keep in mind that I think Jay Breidt is going to address some more details on this tomorrow, and is that right?

DR. ERRIGO: Yes, tomorrow.

MR. FOSTER: Jay will be describing the FES calibration in detail, but the APAIS is just me. Sorry.

DR. SEDBERRY: No, that's fine. The reason that I mentioned this is we're going to try and squeeze in one more presentation before lunch, and it's a relatively short one, but if we could -- If there's going to be detail provided tomorrow on the FES, maybe we don't need to ask questions about that today, and we can wait for that until tomorrow, and we couldn't have these at the same time, because of the timing of the speakers, but that just ended up being the way it is, and so, anyway, are there questions? Okay. Kyle, are you ready?

MR. FOSTER: I did have one last -- It's just a few slides.

DR. SEDBERRY: Sorry.

MR. FOSTER: There is one last section here, and, again, it's just to give sort of a very high-level summary of the effects on the time series, the catch and effort time series. This is a box plot summary of ratios of the South Atlantic annual effort estimates for shore mode and private boat

mode from 1981 to 2017, and so, again, what we're summarizing here is, for the entire South Atlantic, this is the ratio of the -- It's really the fully-calibrated effort estimate divided by the original effort estimate, and, again, at the annual level for the South Atlantic as a whole.

Again, you see, for shore mode, the mean, or rather the median, of the differences is between five and six, a little bit closer to five, but then there is a range, quite a range, from about four to over seven-times, and so the new shore mode estimates are four to roughly seven-times greater after the application of both the FES and the APAIS calibrations at the region level, and, of course, at the state level, and down state wave and all of that, you have a greater range of adjustments. For private boat mode, again, it's a -- It is a systematic increase still, after both calibrations are applied, but it's a much smaller increase, although still large, roughly between two and three-times increase on the effort. While there is variability, these results were still consistent across all the states, and consistent beyond the South Atlantic region as well.

Just some quick examples on catch, and this is just looking at the APAIS calibration, and so, again, if we don't apply -- Sort of an interim step, if we just apply the APAIS calibration, this is giving you some indication of how it affected catch estimates, and so this is for, again, 1981 to 2017, South Atlantic, annual catch, and I apologize. I believe this is total catch, but it might just be landings, and I will need to double-check that, but the basic take-away here is that the APAIS calibration, one, was -- Its effect on the catch estimates was much smaller than the FES, and it was also variable.

For some species, it might have been somewhat systematic, but, across species, it was variable, and, again, on the whole, much, much smaller in magnitude than the FES calibration. Again, I know two of these are more inshore inland species, but we will be going through the five that were identified in the terms of reference in a good bit of detail in a later presentation.

Now, when we look at the effect of including both the APAIS and the FES calibrations on the catch estimates for these species by mode, again, you can see that, essentially, the changes are really just reflecting, or primarily reflecting, the FES calibration, as we would expect, and so, for the shore mode estimates, we're having something in the neighborhood of a six-times increase, and, for private boat modes, it's an increase of on the order of two to three-times. Again, that's being driven primarily by the FES change, the FES calibration. That's it for that presentation.

DR. SEDBERRY: Any questions?

DR. LI: Again, I really would like to see a time series like before the mid-2000s, when we assume both surveys are sampling the true population, versus after, when we believe the FES is the more accurate survey.

MR. FOSTER: Yes, and that, again, is, unfortunately, a limitation of the FES calibration, is that we don't have the mail survey estimates from those earlier years, which, really, we can model something, and we could make some assumptions about it, but I don't think we have a really good way to try to reproduce that, because I think, if we did have that, we could have -- We may have very different results than what the calibration really looks like for those earlier years, but, again, it's just a central limitation of the approach that the benchmark data are only available for the three years that they are, and that relationship is informing -- Not exclusively, but it's very informative to the calibration for the whole time series, and so I'm agreeing with you that, yes, it would be

very helpful if we could have that relationship, if we could have estimates from the earlier time period. The wireless effect in the series, in the -- I guess, to keep us on track, I would just say let's, if we can, if it's okay, if we could just hold that discussion for when we have Jay with us.

DR. SEDBERRY: Any other questions?

DR. NESSLAGE: The slide that shows the effects on the time series of just the APAIS calibration, I really appreciate this. This has helped me to understand the difference. Have you done this for all of the major species, and, if so, are they anywhere -- For instance, the end of your whiskers are above the one line. In other words, are there any for which you would have -- These do seem reasonable, but are there any species for which the APAIS calibration alone creates a very large difference in the catch estimate ratios that we should be aware of?

MR. FOSTER: I will show some examples. When we go through the five species in the later presentation, there are some -- We will sort of talk about the overall trends, but then I will highlight -- I think it's between twenty to twenty-five individual point estimates that we'll sort of dig into, and, for some of those, you will see where, yes, the APAIS calibration does make a big difference, by either negatively or positively interacting with the FES calibration to result in a larger change or minimize the change that occurs because of the FES calibration, and so, because the APAIS calibration is more variable, it does -- I don't think there are many cases where it sort of completely offsets the FES calibration, but there are some where it's comparable in its overall magnitude.

I guess, to just wrap it up, there are some species, for example, where the APAIS calibration was significant, because, for some species, their catch is highly correlated with something that was included in the calibration, and so, for example, catch rates for some species might be much higher at high-activity sites compared to low-activity sites, because those sites tend to also be things like big marinas or large boat ramps or things that are correlated with boat size, or they're close to an inlet, or something about their spatial location, or catch rates might be highly correlated with weekend versus weekday or things like that. For those species, yes, the APAIS calibration could have a much larger effect than for others that are more homogenously distributed across the APAIS strata.

DR. SEDBERRY: Any other questions for Rob right now or John? Thanks, John, and, Kyle, you're up next. This is Presentation 5, again addressing TOR 1A, for those of you taking notes, and we'll switch our notetakers this afternoon, but, if you could keep on with this presentation as well, I would appreciate it. Thanks.

DR. SHERTZER: This one is a bit of a shift of focus away from survey sampling methodology and more of a sensitivity analysis or examination of what we can expect to see in stock assessments when we're using the FES data, and I am using black sea bass as a case study, but I think the trends that I will show, the general results that I will show, are generalizable to other stocks as well, and I only have a handful of slides, and so I hope you're not all terribly disappointed in the short duration.

A little bit of background, and so, at the Beaufort Lab, we did do assessments last fall using the revised landings from MRIP for black sea bass, blueline tilefish, red grouper, and vermilion snapper, and these were the assessments that the SSC reviewed last October, on a webinar, and you may or may not recall that, in those assessments for black sea bass, the revision resulted in a

change of status from not overfishing to overfishing, and, in the SSC's report from that meeting, they noted some issues that they wanted examined a little more closely, in terms of the trends, and we were asked to provide a lot more diagnostics from these four assessments, which we did, and then, shortly after the government shutdown, we had another webinar to look at those diagnostics, but, instead, we talked about the FES more in general, and questions came up and concerns came up, and, instead of talking about those assessments, we're now here at this meeting looking at FES and MRIP and the program in general.

This is sort of cycling back to black sea bass and trying to take a closer look at that assessment and the time series of landings, and so this current analysis includes the original SEDAR 56 assessment, and that was the black sea bass assessment, and then that revision, the MRIP revision, that we looked at last October, but it also includes two more time series of landings, to sort of help tease apart what features of FES the recreational landings estimates are driving the assessment results. I just had to note that this is purely exploratory, and we're not suggesting that these be used for management at all, and so don't try that.

The four time series of removals that are being used in this assessment -- I didn't make any changes to the assessment. This is simply substituting new landings time streams, as well as discard time streams. Other than that, the assessment methods are exactly the same as they were in SEDAR 56. One is the SEDAR 56 set of landings, and this was prior to the revision, prior to FES data, and then the MRIP -- The second one is the MRIP revision, and those were the ones that we looked at last October.

In this set, these are not directly the FES estimates, but, instead, they're the ones from SEDAR 56 that are adjusted according to the FES-CHTS ratio that is done on an annual basis, and the reason we didn't take the FES estimates directly is because the ones we use in the assessment are computed by the Southeast Fisheries Science Center. More specifically, they are computed by Vivian, and they include some differences from what MRIP provides, primarily in the weights, average weights, and some other adjustments that the charter boat -- I think Vivian is going to talk more about that in another presentation, but, to be consistent with what was done in SEDAR 56, we just took those landings and adjusted them by the FES to CHTS ratio for the MRIP revision.

The third one is a new one, and this is the MRIP revision average scale, and, in this one, the adjustment to the SEDAR 56 landings is not done on an annual basis, but instead as a single ratio that's applied across the entire time series, and the ratio is computed separately for landings and for discards, and the reason for doing this was to look at the scale of the landings rather than the trends in the landings that occurred, the changes in the trends that occurred with the FES calibration.

The fourth one is CHTS-adjusted, and this is the one that takes the SEDAR 56 landings, but applies an increasing adjustment at the end of the time series, and this was the one that Rob mentioned two presentations ago that tries to account for the bias in CHTS at the end of the time series due to the landline phone usage, and so, instead of it decreasing over time, it's increasing in accordance with the different time series that Rob showed on one of those spaghetti plots.

Just to sort of summarize this, in terms of an experimental design, there are sort of two features that we were looking at in the landings. One is increasing trend at the end of the time series and whether or not that was included or not, and so the two that did not include that were the SEDAR

56, those original landings, and then the MRIP revision average scalar, and we just adjusted that time series upward to account for the average FES change, and then the ones that did have that increasing trend at the end of the time series were the ones that we called the CHTS adjusted and then the MRIP revision.

Then the other feature was whether the scale was different, and so an increased scale, and the two that did not have that were the SEDAR 56 and the CHTS adjusted, and the ones that did have that were the MRIP revision that was scaled up and then the MRIP revision, which were the FES data.

Time series, what did those actually look like, it looks like the Y-axis is cut off on the screen, but everybody has it up on their laptops anyway, and so hopefully you can read the Y-axis labels. The top-left panel is showing the landings times series for these four different streams, in terms of numbers of fish, and so the SEDAR 56 is in blue, and it's sort of underneath the green line, the CHTS adjusted line, except for at the very end of the time series. They are identical until we reach near the end of the time series, but you can see it's a very slight adjustment. It's increasing from when it begins to deviate until the end of the time series up until about 1.4-times the amount at the end for the CHTS adjusted.

Then the two that have the larger scale of landings were the MRIP revision and the MRIP revision average scale, and you can see those are very similar. It's just larger landings than the other two, and then the bottom-left panel is showing the similar removals, but in terms of discards, and so those are discard mortalities, in terms of thousands of fish, and you can see a little bit more deviation at the end of the time series in the discard mortalities, but similar in the sense that they're separated into the scale, two of them that had the higher scale landing, the MRIP revision and the MRIP revision average scale, and then the two that had the smaller scale were the CHTS adjusted and the SEDAR 56.

Then the right panel is looking at the total removals, in terms of weight, and so, if we put the discards and landings together, and then look at the removals, in terms of weight, that's what the right-hand panel is showing, and you can see a similar story there as in the other two panels.

The next couple of slides are output assessment results, and the top panel here is the estimate of abundance that's being produced by the assessment, and so the SEDAR 56, the original estimates, are the lower estimates, the blue line, and, with the CHTS adjusted time series, the landings are similar, except for at the end of the time series. You can see the estimates of abundance are almost identical.

For the other two, the MRIP revision and the MRIP revision average scale, those are estimating higher abundance in the population to match those landings, those higher levels of landings, and this is really being driven by what you see in the bottom panel, which are the estimates of recruits each year, which show pretty much the same trend as the top panel of abundance, but it's the recruitment that then gets carried through the age structure to produce the total abundance, and so, essentially, with the higher landings, the assessments need to have a higher abundance, estimate a higher abundance, and able to produce that level of landings that we're seeing.

Estimates of stock status, the top panel is showing spawning stock biomass relative to spawning stock biomass at MSY, and, for pretty much -- For all of these time series, the trends and the scale are pretty much the same, and so, even though we're seeing higher estimated abundance with the

higher landings, that doesn't mean that the relative scale is going to change, and so the SSB, the denominator, is scaling along with the total abundance, and so we have a higher SSB MSY for the cases where we're estimating higher SSB, and so the ratio of the two is fairly consistent across the four time series.

Where we see deviation is more in the fishing rate, the F over FMSY ratio that is shown in the bottom panel, but, again, those are similar across most of the time series, until we get toward the end of the time series, and that's where they start to deviate, and so you can see the SEDAR 56, the initial estimates that we had from the assessment, are the lowest, F over FMSY. They're the lowest ratio at the end of the time series, but then, when we had the MRIP revision, that was the highest, and so that's that flip that we talked about at the beginning of this presentation, where it went from not overfishing to overfishing.

What's driving that flip? Well, it turns out that you can compare the CHTS adjusted, which has an increasing trend, but the smaller scale of landings, and compare that to the MRIP average scale, which has the higher landings, but not the trend at the end, and you can see that it's really being drive -- The total estimates of F over FMSY are being driven by the increasing trend at the end of the time series and not by the total scale of landings.

Just to summarize what I just said, the primary effect that we see of the FES landings on the assessment output is that the assessments need to estimate greater standing abundance in order to match that level of landings, and so there has to be more fish in the ocean to support that larger catch, and one sort of corollary of that is that it will translate into larger TACs, larger ABCs, because MSY will be higher with the higher level of landings, and so, in this black sea bass example, MSY was 36 percent higher when we used the FES estimates in the MRIP revision than it was in the SEDAR 56, and I think that's a result that we'll see across other stocks as well.

The change in the status that we saw at the end in black sea bass, and we saw it in the other stock assessments took although it may not have flipped the status, but we did see an increased level of overfishing, or at least the increased F over FMSY, and that's due to the increased slope of the removals at the end of the time series and not due to the scale. I think that's the last slide.

DR. SEDBERRY: Thanks, Kyle. Any questions?

DR. DREW: Did you look at any kind of -- Did this affect your retrospective bias at all, these differences in the landings?

DR. SHERTZER: We did not look at that at all. That's a good question, but we didn't examine that.

DR. SEDBERRY: Any other questions?

DR. SHERTZER: But I bet they did in the Northeast, and they just had reviews last week from some assessments that used the FES data. Being in the Northeast, I would be really surprised if they did not look at that.

DR. LANEY: Kyle, I'm not sure this question makes sense, but I will ask it anyway, and so, if you go back -- I mean, we have fishery-independent data that are consistent, from a time series

standpoint, with these different estimates, and so, if you go back and you look at the fishery-independent data, can you gain any insight into whether or not these estimated increases are real or not? I mean, the question really is that does get us any closer to what the true population estimate is out there?

DR. SHERTZER: I guess the truthful answer is I don't know how to answer that question. Without knowing what the true abundance out there is, it's hard to know. I don't know that we're getting that information from relative abundance from the surveys, but you're pointing at Marcel, I think. Maybe he has some MARMAP --

DR. LANEY: He's shaking his head, and so I will just ask Marcel to comment, since he's responsible for that survey.

DR. REICHERT: Thanks. I think you're right that it's a relative measure of abundance, and so it's a scaling issue that you don't know where that is relative to the true abundance, and so I don't think the fishery-independent survey, the index, will give you much information.

DR. SEDBERRY: Any other questions? All right. At this point, I think we can --

DR. REICHERT: Sorry, but I just want to make sure that everyone understands that it doesn't give you much information to answer that question. It has a lot of information.

DR. SEDBERRY: We understand. Let's recess for lunch and meet back here at 1:30.

(Whereupon, a recess was taken.)

DR. SEDBERRY: We have heard a lot about how the data are collected and how the survey is done, and we're going to hear, shortly, from Vivian Matter at the Southeast Fisheries Science Center about some post-processing of the data as well, but, before we get into that, to address the Term of Reference 1B, I wanted to kind of take a few minutes here and have the committee, the SSC, think back about what we've heard for the last couple of mornings, or yesterday afternoon and this morning, and make sure that -- See if we can't maybe summarize where we are now and see if we can maybe come up with some recommendations or a consensus statement before proceeding.

We have heard a lot about the FES versus the CHTS, and it might be useful for the SSC at this point to summarize or come up with a consensus statement about the FES versus the CHTS, and I know we have done that in the past. At our October meeting we did it, and we reiterated it again at the April meeting, but we've gotten some new information, some very great detailed information, on these surveys, and it might be useful at this point to kind of come up with a statement about it, and so I'm going to open the floor to SSC members, to see if anybody has any thoughts, or maybe even some wording along those lines.

MS. LANGE: I think it's clear that S&T has made the case that the CHTS had outlived its usefulness, because of the change primarily in the landlines and the use of landlines, and not only the fact that the overall coverage of, in general, the number of landlines out there, the relative proportion of landlines, has decreased, but, also, the difference in the population that continues to have landlines compared to the overall population, and I think that case has been very clearly

made. I'm not sure how that relates to what comes up next, but I think they've made a good case for that.

DR. SEDBERRY: I think you're right. I think that's a definite conclusion that the SSC can come to, is that the telephone survey has outlived its usefulness, for the reasons outlined and presented to us over the past day or so, and so that's a definite, that the telephone survey is no longer valid. I think we might also be able to say that the FES is an improvement over the telephone survey.

DR. REICHERT: If I remember correctly, the SSC has previously said that the new methodology is the best available science, and so I think, perhaps, it's useful to restate that, and I think that may be something that the council would appreciate, and so I am asking if any of the SSC members here would disagree. That doesn't necessarily mean that those new methodologies don't have -- That there are issues, and we are probably going to continue to discuss that, but I think that could be useful, to reiterate our earlier statement.

DR. SEDBERRY: Yes, I agree with that, and I think we can maybe even make it a stronger statement, now that -- Based on this overview we've had and the detailed differences that we've seen between the two surveys and the details we've gotten on the calibration methods, that we can kind of maybe throw those things in and say, yes, we have received the information, and we believe that it reiterates what we had previously stated about the two surveys.

DR. NESSLAGE: I guess it would help me if we could make the distinction between design and implementation. I think I have been convinced so far this week that the design, as the peer review indicated, is the best available science and that I've seen -- A lot of my questions have been answered, and thank you, and I think you guys have done a great job.

Where I think we might run into problems later is in some of the implementation, especially on an area or species-specific basis, and that's where I don't want the SSC to get too far down the rabbit hole of saying best available science carte blanche and then have to back up again and change that, and so, if it's possible, at least at this point, and I'm not saying we won't ultimately come to that same conclusion for implementation, but, at this point, if we could start with design, I would be comfortable with that, but if I may also ask a question, to assuage my fears about a few things, and is that all right, George?

The design that you guys have created for the new FES, to deal with changes with APAIS and whatnot, is very complicated, and I'm a modeler/stock assessment person, and I create complex models as well, and I was just wondering -- I'm sure you have, but, at some point, have you simulation tested your code and your modeling approach with kind of a known dataset, where you know the universe, and you know what the answer should be, and then seeing if you're able to return unbiased estimates with your model, because that would make me feel so much more comfortable, and I would be very much in support of the best available science statement if that were the case.

MR. FOSTER: I will start. There is a couple of different ways to answer the question, and so the first thing I guess I will say is the design estimation -- Well, let me cut to the chase. Yes, we have simulated. The thing that is sort of the most novel, in terms of what we do -- Everything we're doing is just standard survey design-based estimation. It doesn't rely on any sort of underlying assumptions about distributions within the data, whether they follow some sort of a -- Do the data

follow some kind of normal or Poisson distribution or anything like that. None of that has to --
We don't have to make any kind of modeling assumptions about design-based estimation to work.
That's sampling theory that goes back to the 1940s, or earlier.

The one thing that we have that is kind of the most novel, I suppose, is how our intercept survey,
the APAIS, draw process works, because it uses sort of an iterative process -- Iterative is not --
Replication-based process to implement the constraints, and so, if we were just doing a standard
unconstrained probability proportional to size draw, that is very cut-and-dried design-based
sampling, and the estimation would follow right from it.

With control selection, and our implementation of the replication-based control selection, we did
simulate that, to still ensure that we had gotten all the calculations to generate the sample weights,
the primary stage unit sample weights, correct, and, yes, we have simulated that and presented
some results at an AFS symposium on recreational data collection a few years ago, at Seattle or
Portland, I think, but, anyway, yes, we have done that.

As far as the FES goes, in terms of the sampling, there is really nothing there that is -- I mean, it's
sort of simulating to show that stratified random sampling using the sample weights returns
unbiased estimates, and that's kind of just been proven, and you can derive how it works.

The things that still are somewhat unknown, but we have done some pilot testing on, is how we
address non-response, potentially the other non-sampling errors, and those are present in any
survey, the non-sampling errors, and they are notoriously difficult to try to quantify and correct
for, and so we have tested methods to implement things like -- Things to mitigate non-response,
like use of incentives, how the questionnaire is designed, our weighting adjustment approaches to
try to minimize -- We can show that, yes, they make improvements, and they rely on standard
practices, but there has not been a lot of -- Correct me if I'm wrong, Rob, but there's not been
direct sort of simulation work for the FES, I think because it's one of those cases where, without
auxiliary independent information, we would probably just get back the assumptions that we used
to generate the simulated population that we were sampling from, but, again, the bottom line is
these surveys are following standard design-based sampling and estimation procedures, which is -
- I don't want to just say that is essentially settled science, but it's a long history of proven sampling
theory, I guess.

DR. NESSLAGE: Just to follow-up, thank you, but there's a lot of different sub-components.
Have they ever been combined in this uber survey design? Do you know what I'm saying? Like
each one of those components that you showed has been tested for years, and it's basic stats survey
design theory, but, in this particular combination, is it still providing -- Can you show that it is
providing unbiased estimates?

MR. FOSTER: We've got sort of a long-running project that we have been working on to build
sort a gigantic simulation engine that would allow us -- Now I understand your question, and I
apologize that I didn't quite get at it earlier, but that would simulate, essentially, a full population
of fishing trips clustered by households and also at access sites that would allow us to do exactly
what you're talking about.

That, again, is just something that it's in the works, but we don't have a definitive timeline for
getting it done, and so the short answer is, no, we haven't done that sort of simulation work. We

were coming at it sort of from how best can we allocate sample across the different surveys, to minimize the overall variance, or improve the overall precision, on our catch estimates, the final estimates, at the end of the process, but it could also be used to do exactly what you're describing, but the methods that we're using for each of the components are fully design-based, and so they are returning to us unbiased estimates.

Then the ways that those estimates are combined are straightforward, and then we use -- I didn't present on the variance estimates, but we use sort of standard methods for calculating the variance of products, for example, because we're multiplying the different component pieces together, and so all of those sort of follow standard procedures, but we have not done what you have suggested, is building a gigantic simulated universe that we can test everything with.

DR. SEDBERRY: Genny, and the committee, we can -- I feel like we can say, at this point, that we agree that, with all the new information and reiterated information that we have received, that the FES is the best scientific information available. However, there are concerns -- The committee has concerns with implementation, and that we would want to state what those concerns are, or we might want to wait until we've heard some additional information before we state those concerns, sometime before we depart here.

DR. NESSLAGE: Yes, and I think -- Well, I think, but it would be helpful if the SSC would consider making a recommendation that S&T prioritize a simulation study like that, if for no other reason than this type of group of folks -- That's how we test our new complicated stock assessment models, and so we're familiar with that sort of framework, that simulation testing performance-based performance evaluation, and so, if you guys can show us those sort of results, I think that would go a long way to getting everybody onboard and being like, yes, we love the new FES MRIP survey. Seriously, because, at this point, I think a lot of folks aren't completely convinced, and I think that would really bring it across the finish line for at least this group, and I will let everyone else speak, but, for me, it would really help.

MR. FOSTER: The only word of caution that I would say though is that, yes, we can do that, and I think -- Because each of the individual components are unbiased, I am fairly confident that it is going to return the result that we program into the population, essentially, that we code into the population, but I wouldn't want -- I mean, it's not necessarily -- The simulation isn't necessarily going to be able to demonstrate why the FES numbers are correct, and I think that's where everyone's primary concern -- My sense is that that's where everyone's primary concern is.

The big increase from CHTS to FES, we will have to code what -- We will have to build the simulation population, simulated population, and have some level of effort, and, if we code it at what the FES currently estimates effort to be, then we're going to get simulated results that say, yes, they're at this level. If we program them to be lower, then they're going to come out lower, and so it's going to return that result, based on how we program the population. Does that make sense? Okay. In other words, it's not an independent validation of the FES estimates. It's just a validation of our process for estimation.

DR. YANDLE: I would like to preface this with I am sympathetic to what you're saying and the concerns about implementation. However, I think we have gotten a lot of very strong information about the fundamental strengths of the science presented, and the science has been worked on for a long time on this, and I am very hesitant to, for lack of a better way to put it, starting to fall into

bring me the broomstick of the Wicked Witch of the West before we're going to be satisfied. There is always that one more thing we need before we're going to be happy, and, just thinking back over what we have been doing for the last several meetings with this, I think we're starting to get into that territory, and so I just wanted to put that little bit of caution out there.

DR. SEDBERRY: Thank you. Any other comments from the committee or questions?

DR. AHRENS: First, thank you to everyone for the presentations. They have been incredibly informative, and I thought that certainly I learned a lot from the presentations, and I'm much more comfortable with the FES. I think where my concern comes in is not the comparison between the two surveys, but it's the kind of converting the household survey effort back into time, in that time period between 2009 to the time when there was the survey overlaps, and it seems that, from the stuff that Kyle has shown, that it is those -- It's kind of the rate of change during that period that is causing the difference in the stock assessment results, and I am curious to know how comfortable the group is with that imputation and the data that's kind of used to derive those relative changes.

DR. LANEY: Well, I can't address Rob's question, but I think I had one that is somewhat related. Based on what Yan had pointed out earlier about the time series prior to the year 2000, I think, and the fact -- They say a little knowledge is a dangerous thing, and so bear with me here, since I'm new to this whole process, but my understanding of what you were saying, Yan, was that, prior to 2000, because the rate of change in the coverage was a whole lot less, and there was more synchrony between the landlines and the households and so forth and so on, did I understand correctly that perhaps those estimates from that time period were more valid then?

Then, when you get to the point beyond the year 2000, where the rate of coverage changes so dramatically, and you have to make some sort of an adjustment, does all of that mean that we could consider the time series prior to 2000 as science that doesn't need much of an adjustment? Is that what that equates to? I don't know. I am asking that question. It sort of seems, to me, to -- I guess the basic question is exactly where did you really see such a dramatic shift that you need to make a big adjustment to it?

DR. SEDBERRY: I think Mike has some comments relative to this.

DR. ERRIGO: These are actually all great questions and excellent discussion, and I think a lot of your questions will be answered when Jay Breidt gives his presentation. I sat through a presentation of the FES calibration, and I learned quite a lot, and so I think, if we hold on this until tomorrow morning, when Jay gives his presentation on the FES calibration, I think a lot more will be clear to everyone, because the FES calibration was rather -- I know they used standard methodologies, but it's rather complex to somebody who doesn't know those standard methodologies, but he will explain all the factors that were used going back in time and all of that kind of stuff and which factors were most important at which time periods, and so I think it might be best if we hold on those questions until then.

DR. SEDBERRY: Thanks, Mike. I think that's a good suggestion. Before we move on to Vivian's presentation, I wanted to just take a minute also to check with Steve Poland, who is the liaison between the SSC and the council, just to kind of get a check on where he thinks we stand, in terms of addressing the council's motion and kind of the council's charge to the SSC and this group, in terms of this meeting. I don't want to put you on the spot, but if you're -- You can think about it

and come back to us later, but, before we get -- Sometime today, but now is a good time if you're ready.

MR. POLAND: Thank you, George. First, I want to thank everybody from the MRIP Office of Science & Technology for all the detailed presentations that we've seen, and more that I know we will see, and so I do really appreciate that. From my perspective, this really diving into the background on FES, the sampling methodology, and where a lot of the biases are associated within CHTS and how they're addressed in FES, that certainly helps me, as somebody that has to convey this information the stakeholders in North Carolina and the stakeholders throughout the region.

It's certainly more stats that I have really thought about since I sat in Fred's biostats class, which he gave me an A in, and so, during all the presentations and the discussion by the SSC, I was really pleased to see the SSC members, as well as the state partners that were invited, really get into the weeds on this and really get into the details and really try to understand what was going on.

The council originally asked for the SSC to convene this workshop based off of feedback that we received from the SSC, and I went back, and I was looking at some of our discussions at prior meetings, in December and back in March, and I think March was actually the meeting where we provided direction to staff to start convening this workshop, and, really, what the council would like to see come out of this data workshop, and come out of the SSC, is just a -- Not necessarily an acknowledgement, but just provide the council with a little bit of assurance that our SSC feels like new MRIP estimates are the best available science and they are appropriate to move forward with management of our resources here in the region.

Right now, we've got I think three or four assessments that are just kind of waiting in the wings and waiting for this issue to be addressed by the SSC and for us to get catch level recommendations. I mean, we've got management of species that is waiting. We've got a red grouper rebuilding plan that we've missed the statutory deadline because we haven't received updated ABC recommendations for it, and we've got four revision assessments, which the SSC reviewed back in 2018, which kind of sent us down this path of really diving into the new MRIP catch stream and really trying to tease out where these increases in catch are coming from and if they're real or if they're some artifact of sampling design.

So far, it seems like that the sampling design is more robust, and it's providing better estimates, although they are higher, and it seems to go against what we are accustomed to in the region, and I think somebody said it earlier, and I can't remember, but it's not necessarily an increase in effort, but it's just a shift in our baseline of what we expect, and I still feel like that there is some stones that can still be unturned, and this is a continually-evolving iterative process, and I certainly am convinced that all these experts in the room, all these big heads in here, can figure out how to move forward.

Really, what we want to see from the council is the SSC to be comfortable enough with these new estimates and start the management process back. Like I said, we're kind of waiting on this input, and certainly, if there are any recommendations that come out of this workshop that the council can tackle, as far as maybe trying to get an alternative estimate of participants, to put these numbers into greater context, certainly provide us that input, but, going back and looking at the direction that we gave, we certainly want to see the Scientific and Statistical Committee identify uncertainties, or potential biases, and I feel like you've been presented that information to make

those determinations and develop recommendations on how to proceed in the short-term for using this data in stock assessments.

Not only in stock assessments, but developing ABC recommendations, and I know Mike provided a very thorough spreadsheet of all our managed species, and not only assessed species, but unassessed species and how these new estimates could potentially change catch levels for those species, based off of the ORCS process, and so I certainly want to see us move through this today and tomorrow and get to those discussions, because, at least from my perspective, I feel like those unassessed species are ones that the SSC hasn't even really had a chance to look at these new estimates, in the context of those protocols for unassessed species, and the council would really like to hear your input on that too, because we've got a few of those species, or a few of those complexes, that we've got amendments that we're working on right now, and amendments where we need to redefine optimum yield or talking about bag limit adjustments and things like that. Certainly I would like to see this committee move forward and really tackle some of those things, and you're doing a great job.

DR. SEDBERRY: Thank you, Steve. That's helpful as we move forward. I appreciate that, and, before we do continue on, I want to make sure that the SSC agrees with what we have -- With what Mike has typed up there in italics as kind of our recommendations or summary statements or consensus statements, and I really want to make sure that we capture the idea that, if the committee agrees, that the FES is a step forward.

DR. LI: Here, I would like to separate two concepts. FES itself, we all agree, is better, is a big improvement. Then now I feel our concern is the calibration part, especially calibration, back into the time period that is before the mid-2000s, where we believe the two surveys are sampling the same population, true population. However, when you back-calibrate, you've got different estimates, and so I am thinking here -- I am wondering if FES itself includes the back-calibration parts or not, and I have the feeling that FES itself is there, and then, using whatever information we've got out from FES, to back-calibrate the data, that's another thing, another project.

I am thinking here that our consensus is we consider the overall design of FES -- Here, I just have a feeling that we can take off the overall design. FES itself is the BSIA. However, we have concern for the calibration using the information from FES back through the time period, and we have concerns about that. Also, a second point is I feel -- To me, I don't have concerns for the difference after the mid-2000s, when actually the two surveys are sampling different populations. However, I have concern for the other time period, the calibration estimates for the time period before the mid-2000s.

DR. SEDBERRY: Thanks, and maybe some of those concerns will be addressed tomorrow, some of the concerns with calibration.

MR. FOSTER: I will make this brief, because I think that we will talk about it more tomorrow, but just to say, while the declining coverage was certainly a major issue with the CHTS, there were still other differences between the two designs that would have been present going back through the rest of the time series, and so while, yes, the coverage likely would have been much closer through the earlier part of the time series prior to 2000, there were still differences that would have been present, such as the gatekeeper effect and the fact that the CHTS was, as Rob described, an out-of-the-blue cold call to whomever was the gatekeeper, whoever was answering the phone,

versus the mail survey being something that sits in the house, and people have more time to think about the answers and respond to at their convenience and that sort of a thing.

The CHTS attempted to profile every single trip that was taken by members of the household, whereas the FES mail survey just asks for summary counts for each of the members of the household, and so there's a number of differences that would have -- The bottom line is there are differences that would have persisted through the entire time series, but, yes, the coverage is one of the major ones, and that would have attenuated out, essentially, in 2000, but there were other important differences that, again, would have been through the whole time series.

DR. LI: I just want to suggest that our discussion question now is not like if FES is better or not. Instead, our focus is whether the back-calibration is valid or not.

DR. SEDBERRY: Okay. Is there any additional comments on our statements there before we move on to Vivian?

DR. NESSLAGE: I'm sorry.

DR. SEDBERRY: Don't be sorry.

DR. NESSLAGE: I just want to make it -- I would like to make sure we're all on the same page. Whenever we make these statements about the best available science, or information, we make these statements, and then, for instance, we might get to the king mackerel assessment and notice a problem, and then we end up in this ten or fifteen-minute argument about it's BSIA and we have to accept the numbers exactly as they're provided to us, and then the other camp goes, but there is something wrong here, clearly, and we go back and forth.

Is there some way that -- This is why I wanted to say -- Oh, you put the word "design" in there. Mike, thank you. Okay. I am comfortable. When did you do that? Did you do that while I wasn't looking? You're sneaky. I like that. Okay. Then I'm fine with that. Then I will drop that comment.

Then the other thing is, on the first bullet, does everyone agree that it's not just the transition from landlines to mobile, but also this issue of a phone versus mail survey, because you guys -- I hadn't thought about that as much until they presented it, and now I'm really convinced that the mail survey is going to give you better recall and a whole bunch of other issues that have popped up that we might be getting better information from the mail survey than we would a phone survey, and so is it worthwhile saying that as well, to back them up?

DR. SEDBERRY: Yes, that's a really good point, and I think it is worthwhile saying that, unless there is any objection by anybody on the committee.

DR. ERRIGO: Just real quick, my reasoning was I think, here, when we flesh out the report, all of the important reasons and differences for the CHTS and the FES should be here. I put this here, and someone had mentioned it as being the critical and most important piece, especially in the later years, and so that's why I just put that there, but I think that we can go back and put in all the rest of the bullets. We can do that later.

DR. BELCHER: I was kind of thinking similarly with Genny's line of thought, is that it's kind of sticking to how like Rob had talked about coverage error and specifics underneath that and then going down the list of what's there, because we never really talked about the recall bias or the response rate relative to a phone or mail survey. Even we know that that's kind of there, that's really not what we've discussed around the table, and so I think, again, kind of sticking to the talking points that we've brought up, because that's really what we've been debating and discussing over the last couple of days.

I think sticking to those key points, and I know one of the key things that, going back and thinking about the state letters, and this was something that came up during lunch, is that how the states approached their parsing out to determine what the average number of trips were was based on the number of licensed anglers, yet it was very clear yesterday that you have licensed matched and unmatched trips, and those unmatched trips aren't getting thrown out of it, so that it's only licensed anglers that are in that list and generating that number.

It's everybody that's in that mix, and so I think that's something to take back home, is that there's reason why those trips are higher, and that's that hidden effort that we were just talking about that may not necessarily be being picked up, and so I think those are key things that we should be focusing on, rather than kind of talk about some of the stuff we know is inherently different between the two surveys, again like the mail/phone aspect of it, and we didn't talk specifically about that, and so I'm just thinking what are our specific take-homes, and it's the coverage errors, the things that we've kind of drilled into in the last two days.

DR. SEDBERRY: Any other comments or additions or questions? Okay. With that, we can move on to Term of Reference 1B, and I kind of wanted to switch the people that were taking notes, to give those folks that have been hanging in there these many hours -- For this next section of covering Term 1B, I have Fred Scharf, Alexei, Tracy, and Wilson to take notes, and, again, Mike will capture some major bullet points and any consensus statements, but just kind of general notetaking that we can use in our report to the council. Vivian is going to talk to us about the post-processing of the MRIP data that occurs at the Southeast Fisheries Science Center. This is Presentation Number 6. Mike will email it out. All right, Vivian. Thanks.

MS. MATTER: Hi, and thank you for that. Again, my name is Vivian Matter, and I work for the Southeast Fisheries Science Center down in the Miami Lab, and we prepare recreational data for the Southeast assessments, and so we were asked to come and talk to you about what are the different things that we do to the data after it leaves S&T.

We used to do a whole lot more than we have to do now, and so I have told John Foster thank you very many times, and there's a lot of processes that we no longer have to do post it leaving S&T, and so I have prepared this presentation talking about two of the things that we do, and I didn't get into the third one that we do, because it doesn't really affect the final estimates, when they are aggregated, but one of the processes that we do is domain estimation, and it's using a program written by S&T to basically break a state into smaller geographic regions, and we need to do that in the Southeast, to be able to treat Monroe County, Florida a little differently.

Sometimes it belongs in a South Atlantic assessment, and sometimes it belongs in the Gulf, and so I will say that these two that I am presenting are not the only ones, but I did want to mention the domain estimation as another post-processing thing that we do at the Center. However, once you

add up all of the smaller sub-regions, they add up to the total, and so it doesn't affect final values, and so I did not go into it in this presentation, and so, with that preface.

The first post-processing item that I'm going to get into is the for-hire survey adjustment, and so, very generally, meaning not just the South Atlantic, this survey covers the charter and headboat fishing modes, and it was implemented in the Gulf of Mexico in 2000, and that included the State of Florida, and Atlantic in 2004, and so it was essentially a change from the random digit dialing done in the Coastal Household Telephone Survey to the for-hire survey of charter captains, and this was providing more efficient and precise charter angler effort estimates than the CHTS, and so that's about as detailed as I'm going to get into that. This is a change that happened a long time ago, and the Southeast Center has been adjusting for this change since 2008.

What is new, and why I am presenting this, is because we have -- We conducted a new analysis last year, after the new estimates were released, and the reason that we're doing that is because, and John got into it in his presentation, is that the APAIS survey change did impact the for-hire survey, the intercept surveys used to calculate the proportion of for-hire vessels that are off-frame in the for-hire survey, as John mentioned, and, also, it's used to partition that area fished, and so, because of that, the changes that were done with the APAIS calibration back in time did affect the charter estimates, and so we went ahead and re-conducted the analysis and used a generalized linear model, and it says right here, with a gamma response structure and log-link.

The methods that were used to do this charter analysis are detailed in the paper that has been provided, and Kyle Dettloff, over here in the corner, is the lead author on that paper, and we just took the opportunity of re-analyzing the processes that were used, and that resulted in what we call the new charter estimation.

What I did is there are these specific species, and so we're actually going to see results of that calibration here, and we have focused on the -- I think there is five or six species that were requested in the terms of the reference, and I think five that were requested in the terms of reference and one that Mike asked me to do.

The first one is gag, and so, basically, for all of these graphs, you're going to see the landings estimates on the top and discard estimates on the bottom, and, of course, those are live discards, and so these are in number of fish, and the red line is showing the CHTS estimates, and the blue line is showing the adjusted for-hire survey estimates, and so these graphs go through 2003, because, after 2003, of course, we have the for-hire survey directly from S&T, and so, essentially, this red line that you see is what is available on the website, and the blue line is what we do at the Southeast Center, and the blue line is what is passed forward through for assessments.

I am going to go through the frames, and I will just introduce the species, and I'm going to say very little about them. If you guys have questions and you want me to stop, you can ask me to hold on, but I'm just showing you, in general, this change. This is golden tilefish. This is greater amberjack. Red porgy. King mackerel. Atlantic spadefish, and this is the extra one. I think that's the last one.

Again, Mike sent me this huge Excel spreadsheet that Steve alluded to, and so it had a lot of species in it, and so we went ahead and generated these graphs for all of the species in that spreadsheet, and it's a supplemental document in your briefing book, and so, if you're curious of how this

change affects the charter estimates for any of those species that Mike put together, you can go ahead and find it in the supplemental, and I should have prefaced this before with, of course, this is just showing the charter boat estimates, and so this calibration is affecting that mode only, and so that's the only thing that we're showing you.

When you see a lot of graphs that say old and new, it's very difficult to parse out -- There is a lot of different things that have changed, and so John has gone through and itemized the APAIS change, the fishing effort survey change, and so this is another change, and all of those are wrapped up into -- When you see a graph that says old and new, it's very difficult to say, oh, well, what is which part, and so that's why we're only presenting exactly one component of it at a time, because, once you put them all together, they confound each other. Is there any questions on the charter calibration?

DR. REICHERT: This is just a clarifying question. The variability in the Atlantic spadefish seems very small, and can you comment on that?

MS. MATTER: I can't, but Kyle can. He is asking about the variability around the estimates.

DR. REICHERT: For Atlantic spadefish.

DR. SHERTZER: The variability around the estimates is calculated at whatever particular level of aggregation directly from the raw data, and so, if you're calculating a year estimate, it would use a PROC SURVEYMEANS like approach to calculate, based on whatever raw data were present in that year. The error bars here are just small because we're dealing with a small number of fish.

MS. MATTER: Moving on to the second post-processing procedure that we do at the Center, it's the weight estimation procedure, and we also -- There is a lot of SEDAR papers in the past that, and I think Mike had put some of these up in the briefing book, that kind of give you a little bit of history as to this weight estimation procedure, but, very briefly, when I first started, we only provided estimates of recreational data in numbers of fish for assessment, and the model itself, in the assessment, would come up with a weight for those catches.

With the reauthorization of Magnuson-Stevens, we were kind of on the hook for preparing catches every two months for the ACL tracking, and this had to be done in weight, and so we were looking for a way to easily and consistently be able to provide those weight estimates on the recreational side, and the reason it was difficult on the recreational side to provide weight estimates is because we didn't always have a corresponding estimate of landings in weight from MRFSS at the time, and that was due to the fact of low sampling. They just weren't able to generate weight estimates for particular cells, and so we generally, at that time, provided estimates in numbers.

At the time, we were trying to come up with a way that we could -- Without having to redo all of these assessments every two months to get weight estimates, and we wanted a consistent methodology that could be used for management and for assessments, and so, at the time, we used the MRFSS weight estimates when they were provided, and we would fill in the holes with this weight estimation methodology that uses the sample data of those weights from the survey, and we would calculate average weights by the strata, and we have this hierarchy of species, region, year, state, mode, wave, and area fished. We implemented a thirty minimum sample size, in order

to calculate the weights, and that's where we were probably back in 2009 or 2010. This is all outlined in a paper that I'm going to mention in a minute that Mike sent out yesterday.

Going further on, we had new MRIP datasets that were available starting -- Well, they were released in 2012 and 2013, but they went back to -- That's the estimates. Last year, we got new MRIP sample data, and so, for a time, we were having -- The weight estimates from 2004 on had a different weight estimation methodology than the ones prior to 2004, and there is a lot less holes after 2004 with the new methodology, which it was a mix, and someone from S&T can probably say it a lot better than I can, but it's a mix of hot-and-cold deck imputation and length-weight relationships that really did a good job of minimizing how many times there was missing weight estimates.

However, there were still, at the end, some of them, and so the bottom line is that we needed a consistent methodology across the whole time series, in order to use the weight estimates in assessments, and so we decided to use the Southeast Fisheries Science Center weight estimation methodology across-the-board for all estimates, and so we would take the final estimates in number of fish and apply this methodology, using an average weight from the sample data of the survey, and get a landings estimate in weight. That is the methodology that we have been using, and I guess the first one was SEDAR 32 in 2013, I think, and we've been using this methodology.

What I am going to show you next is a comparison of those calculated weights from the Center versus the MRIP weight estimates, and when I am showing you the MRIP weight estimates, I am showing you the MRIP charter mode, and it includes all modes, but the charter estimates that I'm showing you have already been adjusted for the for-hire survey, because we, obviously, wanted to show apples-to-apples, and so, basically, it's the MRIP weight estimate adjusted for the for-hire survey at the Center versus the weight estimates that are generated using the sample data.

Again, we have six species that we're showing, and the red line is the weight estimates you would get from MRIP, again adjusted to FES, and so you're not going to match the website prior to 2004 in this graph, and then the blue line is the weight estimates that were generated at the Center using the thirty-fish minimum size. This is gag. South Atlantic tilefish. Greater amberjack. Red porgy. King mackerel. Spadefish.

DR. ERRIGO: This is why I asked Vivian to look at spadefish for me. It was that 2014 point, which was -- When I looked at them, I just looked at the total FES calibrated and then the original CHTS, and I was like, wow, that's a huge difference, and, when I looked at the intercepts, I was like, I can't see any reason why it's so different, because it wasn't in the MRIP estimates. It was this.

MS. MATTER: Right, and this is isolating just the difference in the weight estimation, and so both lines would have the same landings and number. This is isolating the weight estimation, and so you don't have anything else confounding it. We were working, earlier this year, before we were asked to present this, on looking at that minimum sample size, because, at the time, when we had to -- We were on the hook for 100 species every two months and weight estimates, and the thirty minimum sample size wasn't analyzed at the time, and so Kyle generated this analysis for the sensitivity of the sample size on the -- Showing the effect the sample size has on the standard error of the weight estimate, and so this is just showing the reduction factor, or the standard error, based on your sample size, and so we pulled all these species that had -- The only species that you

have been shown in this analysis are obviously ones that had thirty sample size already, but we were trying to look at, well, does it have to be thirty, basically.

What we found here is that you're really not gaining a whole lot, probably after fifteen, in terms of the reduction in the standard error, and so we're actually in the process -- I have provided a draft copy of a paper that we're preparing for an upcoming SEDAR assessment, where the Southeast Fisheries Science Center is proposing that we move that minimum sample size to fifteen, based on this analysis, and so that will be presented in SEDAR, and I think the next one is vermilion snapper in the Gulf, but it would be something that would be applied to both regions of moving to the fifteen.

The next series of slides show the addition of that fifteen, like what would be the weight estimate for the same species if we have a fifteen minimum sample size, and that's shown in the dashed-green line here, and so, in some cases, it's not really making a big difference between the blue, which is what we currently use, and we're proposing that we move to the green-dashed line, and so this is gag. Then tilefish, and there's a slight change there in 2010. This is greater amberjack and red porgy. It didn't make too much of a difference at all in this case, but it's going to really make more of a difference on the species that generally are not sampled as high. Here is spadefish. I preface that with, for example, spadefish.

That's just something that we've been working on, and we thought that we would bring this to you guys as a sneak peek of what we've been working on since you asked us to talk to you about our weight estimation methodology. Any questions?

DR. SCHARF: Vivian, for all of the species except for spadefish, you get big deviations from the MRIP estimation in the Southeast Fisheries Science Center really early in the time series. Are you able to talk a little bit about what is driving those? For the most part, they are sort of restricted to the very beginning of the time series, where you have big differences between --

MS. MATTER: Well, it's just having to -- In the early time series, in the MRFSS weight estimation, there is a two-fish rule, and so they only really need two fish at a strata level to generate a weight estimate, whereas imposing a thirty, which is the blue line, or even the green line, which is fifteen, you are getting coarser -- You're pulling together different strata that may have different-sized fish, but you're doing that in a way that you're trying to get a better estimate of weight, we think, rather than two fish.

DR. ERRIGO: There were a lot of gaps in the MRIP weight samples early in the time period, where there were a lot of fish landed in numbers that didn't have associated weights with them.

MS. MATTER: Yes, and that generally happened earlier in the time series then, especially after the MRIP weight estimation methodology from 2004 on. That did a really good job of eliminating those, but, prior to that, Mike is right.

MR. FOSTER: Just to -- It doesn't change the overall conclusion about the earlier years, because it's still true, to some degree, but just to clarify that we have since applied the weight imputation approach that was used for 2004 forward backwards through the entire time series, and so it's not relying anymore -- In the calibrated estimates, it's not relying any more on the old two-fish rule. It uses our updated methods. That said, the sampling in the earlier years still is not nearly as robust

as it is now, or more recently, and so, even with our updated methodology being applied to the earlier years, there are still more holes back in time than there are in the more recent years.

MS. MATTER: Thank you for that, John, I was pointed to the website, and it still had a -- I didn't want to say anything, but I wasn't sure about that, and so thank you for that. Any other questions?

DR. SHERTZER: It's not a criticism, but just a comment that this isn't a consequential issue for assessments, because we typically fit recreational landings in numbers and not weight, and black sea bass was maybe the only exception in the South Atlantic. It's probably more important for the management side, with the ACL tracking, which would be more important in the more recent years than in the early years.

MS. MATTER: Thank you, Kyle. That's a good perspective. In summary, to address the term of reference, we were asked to talk about how our post-processing procedures would affect the CHTS versus the FES issue, and the bottom line is that the for-hire survey adjustment does not impact the calibration between the CHTS and FES. They are completely different modes.

The weight estimation procedure, it obviously applies to all modes, and it occurs after all of the calibrations are completed, and so that's why we isolated the differences in the graphs that we showed for the weight estimation procedure, but the bottom line is that we need a consistent methodology for stock assessments, and we need a methodology that does not leave any holes in the estimates at the finest level, in order to give that to -- To use it for assessment, and so you got a sneak peek of what we're proposing moving forward, is moving to a fifteen-fish minimum.

That said, I know we've been talking with S&T for a while about matching methodologies for the weights, so our weights on the website don't differ from what we're using, and I had forgotten that John had said he was going to do the weight estimation for back to 1981, and so that's a big step that helps a lot, and this weight estimation procedure is a work in progress, and it's up to modification in the future, if we -- For example, if we decide to go back to using MRIP for all cells that have weight estimates and simply using the Southeast Fisheries Science Center weight estimation to fill in where there is a hole, and that could be a way of moving closer to aligning both weight estimates, but that's something that we'll keep working on, and that's all I have.

DR. SEDBERRY: Okay. Thank you, Vivian. Any questions? No additional questions?

UNIDENTIFIED: Just a comment to add to what Vivian was saying. I can remember a conversation with Steve Turner and Paul Rago a few years ago, and we were all going to get together on how we were doing the mean weight estimation, and so that is on the list of things to do, and so it would not just involve the Southeast Center and Science and Technology, but also the Northeast Center.

DR. SEDBERRY: Any other questions or comments? All right. Thanks again, Vivian. Mike has the terms of reference up there on the screen, on the left, and the SSC can address those questions right now, while this is fresh in our minds. Does the post-survey processing have a larger effect in certain circumstances? Does anybody on the SSC have any thoughts on this? There are no special, unusual circumstances where --

DR. REICHERT: Well, what I was thinking is, in some instances, there were some differences, and, in other instances not, and so I had a hard time finding out whether there were like consistent attributes to either species or other things that may explain why, in certain species, there was a larger difference than in other species. There can be an effect, but I'm not sure if there's a consistent reason why that is. Vivian, you have looked at those data, and is that kind of random, or --

MS. MATTER: Are you asking about the charter estimates, because remember the calibration for the charter did not affect the differences that you're seeing in FES and CHTS.

DR. ERRIGO: I think it would be the weight estimation.

MS. MATTER: Again, if you're just isolating -- If you're just looking at the weight estimation processes, I don't -- I'm not quite sure that that's what was being asked here, and, really, that's why we have narrowed down to see like -- The weight estimation does not have anything to do with the calibration, and like both lines are the same calibrated time series.

DR. REICHERT: Then I may have confused the two.

DR. ERRIGO: I think the question, or the TOR, may be unclear. There is a difference that was seen between the original MRIP estimates and the new calibrated FES estimates, and, occasionally, the difference between those two data streams was not caused by the calibration, but it was caused by the difference in the weight estimation between MRIP and the Science Center, and this is asking are there any particular circumstances where that is always the case, or it has a larger effect than in any other particular circumstances.

MS. MATTER: It just depends on the species, and it depends on how heavily sampled it is, whether moving from one strata to another is going to change the weight estimation, but, again, the way you get around that is to look just at number of fish, to actually analyze the differences based on the calibration rather than the weight estimates.

DR. SEDBERRY: I think that answers that question, unless anyone has some additional thoughts.

DR. DUMAS: Is the only type of post-processing, post-survey processing, that the Fisheries Science Center does related to weights, or do they do other type of post-survey processing work? If I remember, this was not just about weights. This was about any type of post-survey processing that could cause a difference between the CHTS and the FES.

MS. MATTER: We don't do anything to it that would affect that, because the charter does not affect FES.

DR. ERRIGO: The change in the calibration for the charter, if you looked at the total landings, you may see an effect, but, as you saw, the charter calibration is not very big, and so it doesn't actually have a huge effect in the overall landings, but it definitely doesn't have any bearing on the CHTS versus FES.

DR. SEDBERRY: The second question there, as part of the TOR, is are there any patterns in post-survey processing that might affect the disparity between the CHTS and the FES estimates? Any

patterns in post-survey processing? Mike or Fred, do you recall what the thoughts were, the steering committee's thoughts were, on that question? I am a little bewildered by it right now.

DR. ERRIGO: I think it's asking if there were -- If there are any particular patterns or biases in this post-processing that affected the difference that we saw systematically, if there was a systematic change in some way, shape, or form, which there is not. Sometimes they go up, and sometimes they go down.

DR. SEDBERRY: I didn't see anything that looked like a systematic pattern.

MS. MATTER: Again, none of the things that I showed you get at a chance from CHTS to FES.

DR. ERRIGO: All it is is a perceived change when you're looking at the total landings, if you're looking at the total fully-calibrated landings from the Science Center, which have the weight estimation, and then the full CHTS.

MS. MATTER: Right, which is why I went into the detail that I did, and I showed what was actually done, but, really, the ideal way to look at that is not in weight, and it's not all modes, if that's the question. Old and new should just be private and shore and number of fish, and then you can get at what you're after. When you add the modes and you add the weights, then you're adding more, but that's why I broke it down and I showed the differences for each of those two steps, and so, if you have questions, all of those fifty species are graphed in the supplemental data, so you can get at any of those differences of that, but that's what I would recommend, to really look at the change that you're really here looking at.

DR. SCHARF: I think -- This is starting to -- Well, it's not starting to, but it's just getting hazier all the time, but, if I remember when we were discussing this particular TOR, in some of the phone conversations that we had during the course of the summer, the post-processing of the data at the Center that we were concerned about, and the impact that it might have, it wasn't related to the for-hire survey.

It was related to -- When Erik first presented those updated assessments to us back in October, there was some adjustments that were made that we didn't really discuss in a lot of detail in October that were done at the Center outside of the calibration procedure, if I remember right There were some ratios that you applied, and there were some subjectivity, I think, to how those were applied, and I may not be remembering this correctly, but I think that was what was driving this TOR in our discussion, or am I misremembering that, too?

DR. ERRIGO: If that's how you remembered it, I think we weren't all on the same page, unfortunately. Sorry.

DR. SCHARF: Maybe Yan or --

DR. ERRIGO: Because those ratios that Erik used weren't post-processing. Those were to take the landings from SEDAR 56 and turn them into FES-calibrated estimates.

DR. SCHARF: Okay.

DR. SCHUELLER: Given what Vivian said, I don't like what we have written up here, because it implies that this has an effect between CHTS and FES, and we have already said that it does not, and so we just need to make some statement about this does not impact this and not that we don't see a general pattern, because, to me, I would read that and say, oh, well, they looked at it, and it must impact it. No, we need to be clear that there is no impact. We need to put Vivian's statement in here, in order to help our future selves remember what is going on. I don't mean under 1 there, Mike, but I mean under 2. Are there any patterns in post-survey, and then there's a consensus statement that doesn't make -- It will just serve to confuse us in the future, and I would delete it.

DR. DUMAS: I agree, and that's why I asked my question about is there anything else that the Science Center does, in terms of post-processing, that could cause a difference between the CHTS and the FES, because apparently the weight post-processing does not affect that, correct?

MS. MATTER: Well, if you're looking at the estimates in weight, it will be different, because of the two different methodologies, but you just analyze the data in number of fish, and that eliminates that disparity.

DR. ERRIGO: This came up because I was looking at outliers to look into for the SSC, and I found one that was a huge discrepancy between what I had as the FES calibrated estimates from the Science Center and the old CHTS estimates, but, when I went into the MRIP intercepts, I couldn't find how it was -- Why it was so big, and then, when I calculated out, from the MRIP intercepts, what the landings were, they didn't come anywhere near where the landings estimate was, and I realized that it was because of the weight estimation process, and so it was a perceived difference. It has nothing to do with the difference between FES and CHTS, but it was a perceived difference in the two data streams, and so it caused an outlier, or an extreme difference in one data point, that had nothing to do with the calibration.

DR. SCHUELLER: So can our statement be that the work, or post-processing, the Southeast Fisheries Science Center does not impact the numbers of fish? I mean, we need to make just a statement like that that is very specific, and so, if it impacts weight, potentially, in some way, that's fine, but, if it doesn't impact numbers, and that's what is mostly going in the assessment, we need to say that.

DR. SEDBERRY: Does anybody on the SSC object to that? Okay. That's pretty much what Vivian said before, in answer to Chris's question.

MS. MATTER: Can I just request that the statement that it doesn't affect the CHTS versus FES change to effort estimates or as it applies to landings?

DR. SEDBERRY: Is that good? Is everybody happy with that? All right.

DR. SCHUELLER: The second statement, that there may be perceived differences in the -- I mean, there are some differences in the weight estimates, and so can we just remove perceived? There may be differences in the weight estimates, but there is no pattern or systematic bias to these differences is fine with me.

DR. DREW: It sounds like there is kind of confusion about the fact that we're talking about three different time series. There is the CHTS estimates of landings in weight and landings, et cetera,

and then there is the FES estimates of landings in weight and numbers, and then there is the Southeast Fisheries Science Center processed time series that goes into the assessments, and it seems like you're conflating those last two when you're trying to talk about some of these differences, and so there definitely -- There is differences in the weight that goes into the assessment from the FES, but it's not -- There is also a difference from that to the CHTS, but that big difference is driven by the big difference between the CHTS and the FES and not between the post-processing that's being done, and so do you care about the difference between the calibrated and uncalibrated numbers, or do you care about the difference between the uncalibrated and the final post-processed Southeast Fisheries Science Center, as opposed to MRIP, estimates?

DR. VAN VOORHEES: I'm not an SSC member, but I think Katie hit on something that's been troubling me in this conversation, and it's a semantics issue, because I hear people talk about the FES estimates, and sometimes they're talking about the FES estimates that contribute to an effort estimate, and sometimes they are talking about the FES estimates as they contribute to a catch estimate, and so I think it would help if you used effort estimates or catch estimates in your description of what you're comparing, and so maybe just referring to the FES-based effort estimates and the CHTS-based effort estimates, and then, if you're looking at catch estimates, you can talk about the FES-based catch estimates and the CHTS-based catch estimates. The APAIS is obviously involved in the catch estimates, but it also is involved in the effort estimates in both cases, and so that may help clear this up a little bit.

DR. SEDBERRY: I am going to propose that we take a short break here, and then, when we come back, we're going to -- That is Presentation 7 coming up on, again, TOR 1C to look at the species that are listed in the TOR. We are going to look at the disparity in the estimates for those species. Let's take a break and come back at quarter after three.

(Whereupon, a recess was taken.)

DR. SEDBERRY: Before we hear from John again with his presentation, Mike had a few words that he wanted to say.

DR. ERRIGO: I just wanted to maybe help clarify the discussion we just had from the previous TOR. When I was putting together the ginormous spreadsheet, I came across -- I was going through and looking for outliers, and I was just eyeballing them, and I was going to look into the intercepts, to see what caused them, or if I can find any reason why this particular estimate was an outlier.

There were some in the assessed species, but, in the unassessed species, I found some, and so here is Atlantic spadefish, and there was this one or maybe two years which were much higher in what is the FES calibrated -- This is the Science Center FES-calibrated landings in weight, and then the blue is the CHTS, Science Center CHTS, and so I was like, oh, there's an outlier, and, when I went into the intercepts, I couldn't find why it was an outlier, and then, in fact, the expanded weight estimate -- If you look over here, the gray is the MRIP FES calibrated without the Science Center weight estimation, and it's significantly lower than this.

Now, I said, well, maybe I should have the Science Center take a look to see how prevalent this is, and is it systematic, because the differences we're seeing may not be from the FES calibration, and they may be from this, and I don't know. I hadn't gone through every single species at this

point. The reason why I thought it was important, and it's nothing to do with the assessments, and it's mostly for the unassessed species, because we create ABCs for the unassessed species using landings, and so, if these huge differences are being caused by these weight estimates, then we should know that going in, so that we know that, oh, this isn't the FES calibration and this is just how the weight estimation came out, and there's no reason to pick apart the FES when it had nothing to do with this difference.

It turns out, however, that it didn't really have any kind of systematic effect. It was higher and lower in some places, and it was very random, and it didn't affect very many estimates at all, which is great, and so that's what we found out. That was the whole impetus for the previous TOR, and I hope that clarifies it for everybody.

DR. SEDBERRY: It does for me, and so those things are neither prevalent nor systematic. They are just outliers. Our next agenda item is addressing TOR 1C, critical factors most likely to contribute to CHTS and FES disparities for species managed by the South Atlantic Council, and I believe we're looking at Presentation 7, is the presentation that will be given, and the related document is Attachment 10. John is going to present this for the species that are actually listed in the terms of reference, and so take it away, John.

MR. FOSTER: Okay. Thanks very much. This is the last presentation I'm giving, and so I think I remember most of what is in this one, and we'll sort of explore it together. I will start with just kind of a general description of the sources of potential outliers, and then, as we go through the five species listed there, this is going to focus mostly on the sources of changes between the original estimates and then those after application of the two calibrations.

Just kind of generally walking through how can outliers occur in the data, or in the estimates, I should say, and so here, again, is our high-level calculation of catch, and, essentially, there is three basic sources where there are commonsense, intuitive sources of outliers. The data themselves, in terms of counts of trips, or counts of fish, can contribute, and then there are several design aspects that historically have been a source of potential outliers, and then, finally, the calibrations can contribute to big changes as well, and we'll sort of see a lot of examples of those for the five species.

Just a little more detail, and so, coming from the intercept survey, APAIS in particular, of course, we've got a number of different counts of fish for each angler trip, and we have observed landings, and these are the ones that the samplers actually see and physically count, and we have what we call the B1 landings, and these are reported landings, reported by the anglers, but not directly observed by the samplers, and they can also include dead discards, and then we have what we call the B2, which are still reported by the anglers, but these are released alive fish.

Generally, those are some of the primary drivers for, quote, unquote, outliers in the catch estimates, but we also have, as I have talked about previously, some trip or effort characteristics that we use from the intercept survey too, and these are things like the area fished information, whether the angler is an in-state or out-of-state resident, or whether the charter trip was taken aboard a vessel that was on or off of our for-hire survey effort survey frame.

Then, in terms of the effort surveys, again, we're getting counts of trips by angler. From the FES, again, it's for private boat and shore modes, and, from the FHS, for-hire survey, it's charter boat

and headboat, and we also get the area fished information for those two modes directly from the for-hire telephone survey. It's not used from the intercept survey, as it is for the private and shore modes.

In terms of the design, of course, sample sizes can be an issue. We do have cases where, just sort of for an entire estimation cell, we can have a small sample size from the particularly -- Usually from the intercept survey, and so, for a low-activity wave, Waves 1 or 2, or 6 sometimes, we may have just a small number of intercepts all together, and whatever data happen to be in those intercepts can end up being expanded by the effort estimate, which usually is a smaller effort estimate, but, if you have a very small number of intercepts, it could still lead to an unusually large estimate for a low-activity wave for those species.

In addition, we have sort of truly rare events, things that aren't necessarily clustered in space or time and they just very infrequently occur, and so, some of the more offshore deepwater species, we just rarely see them in the data, so that, when we do intercept a trip with catch of some of these species, just seeing one fish will lead to what appears to be a spike in the estimates for that particular species, where, most of the time, there is either no estimate or very low estimates, and, again, it's just because we would have to sample at a much higher rate than we can afford anyway to sort of stabilize those estimates.

Then there are things that aren't necessarily uncommon when they occur, like pulse events, but they are very tightly clustered in space or time, and some of these we can try to adjust for, if we know that they are coming, like if there is going to be a short opening for a particular season, and we can adjust our sampling.

If we know far enough in advance, we can actually adjust the sampling for the APAIS to have some minimum amount of sampling assignments occurring within that window, but, often, these events are not something that we can know well ahead of time, if we've got highly migratory species, for example, or some sort of events that cluster trips in time, and it may be difficult for us to -- Again, if it's not something that we can predict, or know ahead of time, we may not be able to, just by chance, have sufficient samples in the windows when these events occur. Whether we hit them or not, it's similar to the rare-event case. If we hit them, we may have what appears to be a spike, and, if we miss them, then we will likely have no estimate, in that case.

Then the APAIS complex design can sometimes contribute, and so the fact that there are clustering levels within the design, and we might, for example, get a cluster of angler trips on one vessel that is unusual in some way, and, just by chance, by hitting that, intercepting that boat trip with those anglers, we have something that seems unusual in the data, and so maybe we had -- Maybe the anglers were targeting something unusual, and we happened to get them by chance, and this is somewhat related to the rare-event situation, or maybe we have a bunch of anglers that all give us very high release counts for some species, for example, and the another aspect can be the sample weights.

As I mentioned before, we use unequal selection probabilities to improve the sampling efficiency, and so we're going to sites that are more productive more often, and we're accounting for that in the sample weighting, but what that means is, because we go to the high-productivity sites more often, they have a higher probability of being selected, and they have lower sample weights, but the converse is also true, and so, the low-productivity sites, we go less often, and they have a lower

probability of being selected, and so they have a higher weight, a primary stage weight, and so, if when we go to those low-activity sites that have larger weights, if there's something unusual that we see there, or intercept there, or, for whatever reason, there happened to be a lot of trips, just by chance at that site for some reason, then we might end up with unusually high weights, and we try to account for that, generally, through sort of standard weight-trimming techniques, but, occasionally, large weights persist even after that, and it can still -- We are not trying to deviate strongly from standard probability-based design sampling and estimation, and so we don't go in and make just arbitrarily large adjustments to the sample weights, and so those, again, can contribute to seeming outliers, or outliers.

I just wanted to give folks kind of a sense of how many estimates we're sort of producing each wave, and this is just kind of an average count, or average counts of estimates per wave, from I think two or three most recent years, and so, given the states that are currently covered by the programs and the modes and the waves, we have roughly 125 effort estimates, on average, that we produce.

If we have an outlier, from an effort estimate standpoint, that is likely going to -- That has the potential to produce a number of sort of outlier catch estimates, because all of the catch rates for that particular cell would be expanded by whatever the effort estimate is, and so, if the effort estimate is, for some reason, unusually high, or unusually low, that will affect the corresponding catch estimates for those catch rates in the same cell, and then, in terms of catch, we're producing roughly 3,000 catch estimates, just considering the landings and releases, and this isn't including all of the more specific catch dispositions or the mean weights or the length frequencies. I am not counting any of that. This is just sort of the basic landings in numbers and releases in numbers across all of the species and states and modes, waves, et cetera. There is a large number of estimates that are being produced each wave, and, even a rate of 5 or less percent, there is a potential for outliers in each wave.

Then, finally, sort of the calibrations, and, again, I went through sort of the varying degrees of success, the calibration for the APAIS design change, and, in here, typically where we see big changes coming from is when we have species that have catch rates that tend to be correlated, or associated, with design aspects of the survey, and so kind of day. Are the catch rates different on weekend days versus weekdays? If the catch rates were higher on the weekend, then, in the old design, without weighting, those weekend trips were overrepresented, and so, if you had high catch rates on a weekend, you would have had a higher catch estimate.

When we corrected for that oversampling on the weekends, it would have reduced the estimates, or vice versa. If you had higher catch rates on the weekdays, you would have had lower estimates previously, after the calibration. They would have gone up, because we were correctly reflecting the fact that we oversampled weekends relative to weekdays. The same for site pressure, and I won't go through all of these in detail, but if you have catch rates differ between high-activity sites and low-activity sites, you're going to see a change related to the APAIS calibration.

Of course, we have gone through the FES and CHTS comparisons quite a bit, and, again, in this case, you can think of the calibration as sort of this multiplicative scalar, or ratio, and, again, it will vary by mode, between private boat and shore, and it will vary between states, and it will vary between, or among, waves and over years, and so we have presented sort of the average changes, but, down at the most detailed cell level, the calibrations can vary, and so there might be a larger

or smaller adjustment coming from the FES calibration, but it just sort of is case specific on if we had an unusually high or low catch rate in a given cell that had an usually high or low calibration for the FES effort change.

Here are the five species that we'll go through, and I have picked a number of data points within the overall time series from 1981 through 2017 for each of the species, and it's just sort of, again, eyeballing them, the kind of large changes or outliers, and I will be walking through both, in many cases, the landings and the releases, and they're not all -- They don't all correspond to the same year, but there's about twenty to twenty-five point estimates that I will go through here.

Again, I am just going to focus on landings in numbers and releases in numbers, and just sort of some labeling here. This matches sort of what we have on the website, and so, for all of the time series of the catch estimates, there will be a base series, which, again, this is just what we had prior to either of the APAIS or FES calibrations, and this ACAL series, which is an interim, and it just reflects the APAIS calibration, and the FCAL, which will -- It reflects both the APAIS calibration and the FES calibration.

Starting with red porgy, and this one is a little bit unusual. I will be talking about this point here in 2016, just because it has such a -- I guess a little bit of orientation. The landings in numbers, or the harvest in numbers, and these are total landings, A plus B1, and so it includes both the observed and the reported landings, and that's in the top panel, and these are for the South Atlantic, annual estimates for the entire region, 1981 to 2017, and the lower panel is releases, live releases, again, in numbers of fish, again annual and at the region level, and so this is across all modes, private boat, charter boat, and shore, if there are shore contributions.

For red porgy, I am going to focus on 2016, and this one is a little bit different than the others, the other cases we'll look at. Most of the others are usually in years prior to 2013, although there may be a couple of others in the more recent years, and you will see, in this one in particular, although it looks like this large increase is mostly between -- To finish describing the plot, the gray line with the open circles, that is the base series. That is the series of estimates prior to either of the calibrations being applied. The black-solid triangles, that's the ACAL, just the APAIS calibration, and then the black squares, solid squares, is the application of both calibrations.

In this case, we see a large increase going from the APAIS calibration to include both the APAIS and FES calibration, and so, really, that large increase is the effect of the FES calibration. However, you will notice, in both the landings and the releases, there is a large difference between the original estimate and the effort estimate after what I am calling here the APAIS calibration, and, as it turns out, in this case, they are sort of both important.

Here is a table that just summarizes the actual point estimates for the original estimates and how they increase with the APAIS calibration and then the final overall increase with both, and you can see the total change is calculated as the final estimates, the difference between the final estimates and the original base estimates, and then kind of the percentage of the overall change that can be attributed to the two different calibrations, and, really, I shouldn't have labeled this as FES. Well, I'm sorry. That is the percent changes for the FES calibration.

In general, this is the pattern we see. Usually the FES calibration accounts for the large majority of the difference. In many cases though, it's more like 90 percent or greater, but this is one case

where the APAIS calibration is accounting for a somewhat larger percentage, and I mentioned this one is a little strange, and so you may remember 2016 is after we had implemented the new APAIS design, and so, in this case -- Well, let me finish with the APAIS calibration.

The way we can sort of see that the FES is -- How it's contributing is we can look at the ratio of the final calibrated estimate and the estimate that's just got the APAIS calibration, and we can see what the change is, again as FCAL divided by ACAL, and we see that for the landings here and also for the releases, and so it's 2.6 to 3.2. If we look at the ratio, the corresponding ratio, of the effort estimates, again, that reflect the calibration for FES and then that that just has what I'm calling the APAIS calibration, we can see that's close to three, and so that's falling right in between these two, and this information I'm presenting is at a summarized level, and it would actually be -- The application of the calibrations is done at a much lower level, and so, if I had all of the individual cells listed out here, you would see the ratios sort of lining up exactly between the private boat effort for those cells, the change ratios, and the catch estimate change ratios.

That's why there is some discrepancy here, but, again, the overall change ratio in the effort is very close to the change ratios for the two catch estimates, and so that's getting us from the APAIS calibration then to include the FES. To dig into why the APAIS calibration had some effects, in this case, here, what we had to do for the more recent years is -- Because we were making, we continued to make, APAIS design changes, improvements, after we had implemented the sort of basic new design -- When we first implemented in 2013, Wave 2 2013, March/April, we took a pretty serious hit on productivity.

We had addressed all of the sort of deviations from probability sampling that had been present in the old design, but, in becoming more rigorous, from a statistical perspective, we had taken a fairly sizable hit, from a productivity standpoint, in terms of producing intercepts, and so 2013 was a fairly stressful year, and that was the start of this haggard beard for me, and it had just continued to grow ever since, but we had to make a number of changes, again all still consistent with design-based sampling, but ways to improve the productivity of the survey throughout the course of 2013, and, really, into 2014 as well.

Some of those changes included mixed-mode sampling, and so no longer stratifying by mode, exactly, but creating the site groups, which were grouped based on primary mode at the sites, but, when sampling was actually conducted, samplers could collect interviews in any mode, and so that was one major change. Another major change was, initially, we had only the four six-hour non-overlapping time intervals for assignments that cut the twenty-four-hour day, again, into four six-hour slices that didn't overlap.

That was very clean, from a design perspective and for calculating sample weights and estimation and all of that. The problem was that it wasn't productive enough for us, given the available funding we had for sampling, and so we had to create that fifth interval, the peak interval, that overlapped the two daytime intervals. It overlapped exactly three hours for each of those two daytime intervals.

When we created that overlap, we then had to account for that in the sample weighting, because, essentially, the trips that were occurring within that overlap window had a higher probability of being selected. They could have been selected from either of the two original time windows, or they could have been selected in that overlap six-hour time interval, and so we had to adjust the

sample weights to correctly account for that increased probability of being selected for trips that were in that window, and so what I've got here, what I'm showing here, is the difference in the APAIS sample weights from when we applied the correction for this overlap sample window from what the original sample weights were, and so this plot gives you time of day binned to individual hours, and there is just a little jittering here, to sort of separate the data points, and this is a plot of all of the individual differences for the angler trip APAIS sample weights, again calculated as the new weight, the adjusted corrected weight, minus the original APAIS sample weight.

You can see the overlap is in the middle of the day, and so, essentially, it starts at 11:00, and it goes to just before 5:00 p.m. 5:00 p.m. starts the second half of the afternoon interval, and so what happened was weights from 11:00 to 1:00, essentially, were down-weighted, and weights from 2:00 to 4:00 were really sort of a mixed effect, and then weights from 5:00 to just before 8:00 were upweighted, and so trips that came in in this window had their sample weights increased, and trips that were intercepted in this window had their sample weights, on average, decreased, and then, here in the middle, it was just sort of kind of a mixed effect from this adjustment.

Why does this matter? Well, we can plot the same distribution of the differences in these sample weights after this weight correction against the distribution of red porgy intercepts, again, intercepted angler trips with catch of red porgy, and, again, they're lined up by hour, and so you can see that the bulk of the distribution of red porgy trips is sort of later in the day, late afternoon or early afternoon, and so a disproportionate amount of those trips, or a larger proportion of those trips, were upweighted after this correction, compared to a smaller fraction of those trips being down-weighted.

This led to an increase in the catch rate for red porgy, which was then multiplied by the corresponding effort estimate, and so, in this case, changes related to the APAIS actually made a noticeable impact in the change in the overall catch estimates. This is a little bit of an unusual situation. Most of the rest we'll see are more related directly to the calibrations as we've been describing them in the earlier presentations, and it's true -- This is the same explanation for both the landings changes related to APAIS and the release changes related to APAIS in 2016.

Greater amberjack, there is about six points here that we'll go through. For landings, again, in the top panel, it's 1981, 1990, and 2016. In the bottom panel, for releases, it's 1991, 2008, and 2016. Some other features that I will just sort of highlight, and I won't necessarily go into the details on them, but I will just sort of draw your attention to sort of the 1984 to 1988 period of time, and so here's an example where the calibrations were somewhat -- Not completely, but somewhat offsetting, and so the APAIS calibration actually reduced the catch estimates, in this case the landings estimates, so that, when the FES calibration was then applied, which was a systematic increase in effort, the resulting catch estimates were, in some cases, pretty close to the original estimates, and I don't have it right here with me to explain why that's the case, but, generally, again, it's things like catch rates varying by kind of day or high-activity versus low-activity sites, things like that.

Getting into these specific points now, starting with landings in 1981, the big driver for this change is coming from Florida private boat mode, again, east coast Florida private boat, and the top row of the landings table here has the estimates for the entire South Atlantic region as well as the percent of the changes related to the two different calibrations, and so the APAIS calibration resulted in a small reduction, but, again, the FES calibration resulted in a large increase.

When we look specifically at east Florida private boat mode, Wave 2, we can see that that change accounts for a large fraction, well over half, of the total change, and then, if we look at going from the estimate with the APAIS calibration to the estimate with both applied, the change ratio in that cell was an increase of about 4.4-times, and so that's in the landings estimate, and then we see -- If we look at the corresponding effort for the same cell, you can see the two values there, and that change ratio, again, is exactly the same, at 4.4, and so, again, the -- Jay Breidt will discuss more of the details of the FES calibration tomorrow, but, again, it can vary by -- It's not just varying by mode, but it's varying by state and wave and year as well, and so, even though the average change for private boat was more like two to three, individual particular cells could have larger or smaller changes, and, in this case, there is a larger change related to the FES calibration.

Moving forward now to the 1990 point, this one is actually a case where we have a shore -- A contribution to the catch estimates coming from shore, and, again, it's still east coast of Florida. Whenever we have a series where there is typically not much, or very little, maybe generally no, contribution to the catch coming from shore, and suddenly we have a cell that has catch, that's going to make a very noticeable increase, again, because the shore -- The change for shore effort was considerably larger than it was for private boat, and so that can -- Again, it's sort of a combination of a rare event, or at least an uncommon event, having landings in the shore mode, coupled within the calibration, which made larger increases for shore effort, and then that would correspond to larger increases in the catch coming from the shore mode. In this case, it was east Florida shore, state territorial seas, as opposed to being inland, and then in Wave 6.

You can see there that you've got change ratio in the landings estimate of 9.3, and that corresponds to the same change ratio in the corresponding effort estimates, and the APAIS did contribute a little bit to this as well. There was a change in the state resident proportions in that same cell, and so, originally, the base estimate in this cell, 65 percent of the intercepts, weighted intercepts, were from in-state residents.

With the application of the APAIS adjustment, that dropped to 58 percent. Correspondingly, there was an increase in the out-of-state from 35 percent to 42 percent, and so this results in an increase in the effort estimate as well, not just so it's on top of or in addition to the FES calibration change.

DR. DUMAS: With that change ratio of 9.3, or any large change ratio, and so, if this is an example of something that is sort of an outlier, in some sense, and so I think I am following your explanation of why it's an outlier, what might be causing it to appear to be an outlier, but my question is, in the long run, what percent of the time should we expect to see something like this, a change ratio like this, and, if we see something like that, and it's an outlier, is it therefore a number that we should not believe, or, for management purposes, is it something that, because of the way the estimation system works, it's giving us our best estimate, but that best estimate is so improbable that we should not believe it, or is it something that tells us that, yes, we really believe that something changed that much that the landings and the effort really changed by that much? I am getting at how -- Is there some way to estimate how often we would expect to see such an outlier over a time? If we would expect to see an outlier like that, for example, less than 5 percent of the time, over time, then maybe we should discount that outlier, discount that data point.

MR. FOSTER: I will give you sort of my personal thoughts on it, just speaking for myself and not in any sort of official capacity, and certainly not for the program or anything like that. I, personally,

would not want to be in a position where I am constrained to only using the estimates as they are, as a general statement.

I would want the flexibility to, given the information available for the estimate, given any additional information available about historic fishery conditions, weather conditions, whatever auxiliary information might be available to help evaluate that data point, I would want to be in a position where I could consider that information and make a decision based on that, a group consensus decision, given the expertise and experience in the room.

In this case in particular, I don't have any direct experience with the greater amberjack fishery, certainly not -- Well, at all, and definitely not this far back in time, and so, if there weren't auxiliary information available about regulations at the time or unusual fishery conditions that might have contributed to such an estimate, and given the difficult challenge of implementing this large of a calibration, complex of a calibration, given the information available to inform that calibration, I would want to -- I would certainly, personally, urge caution in just using every single number as-is if it seems to be out-of-line with, again, the rest of the estimate series and there is no additional information that would help explain it. I would urge caution in using that directly, as-is.

Now, I will say that I don't necessarily have specific guidance on what to do in lieu of not using a data point like that. We are -- MRIP has a rare-event species project currently ongoing, where we are attempting to develop and evaluate methods that might be used to produce estimates in alternative ways for rare-event situations, but the same procedures might be applied to outlier cases. Again, they range from simple things like moving averages of the data to help smooth apparent outliers like this one or increasing estimation domains, so that it's still a design-based estimate, but it's simply drawing from more data to help calculate that point.

That is sort of where the project is now. In future steps, it might evolve, advance, to more sophisticated small area estimation techniques, but, again, those take a lot of time to develop, and we're just not there yet, and so that's kind of -- Then the last thing was sort of a rate of how frequently we might expect these, and that's going to vary quite a bit by species and how data-rich or data-poor that species might be and how frequently we encounter it in the data, and so, as a simple, first-cut rule-of-thumb, thinking about kind of a confidence interval framework, I don't think it's unreasonable to expect -- I hate to say this for any specific species, but something on the order of maybe 1 to 5 percent we might see.

Again, that's almost just right out of the air, but, because it's so case-specific on how frequently we encounter species and anything that might have been unusual in the effort estimates -- I mean, there's just so many contributing factors that it's hard to say something definitive on it, but it is definitive to say that they will happen. We can't sample at such a high rate that the probability of not encountering an outlier is -- Having an outlier occur is close to zero.

We will likely never be there, and, if we can get to a point where we have sophisticated modeling approaches to smooth it, that everyone can sort of review and agree on, that will be sort of, I think, the best-case scenario for what we could get to long-term, but, in the interim, we feel that it's just best to present the estimates as they are and leave it to folks with the expertise in evaluating those, using additional external information to make the judgment calls on what to do with them. Sorry, but that was a long-winded answer.

DR. DUMAS: Thank you. No, that makes a lot of sense, and another possibility for smoothing the estimates, as we get additional data points, is some type of Bayesian procedure that, for example, could take such an outlier data point and incorporate that information into the estimation, but, getting back to your main point, I think this SSC -- Correct me if I'm wrong, but I think this is sort of where we were and why we were asking for this workshop today, is because I think a lot of us agree that the new estimation procedures that are being used are a great improvement over the Coastal Household Telephone Survey and the way that the intercept survey was done in the past and that all of the improvements you've discussed over the last few days are great improvements, but, when we get to --

I think one of the key issues for us is, when we get to a situation like this, when there's an outlier like this, then is there some way we can get at what is the probability, what are the chances, of such an outlier occurring and is there any way to access that and what are the chance of seeing such an outlier, and then, given those chances, then the council, or someone else, could make a decision, from a management point of view, of what threshold are we going to use to decide when an outlier is just so improbable that we're really going to discount the information of that data point.

For example, if the model assessment procedure could tell us the chances of seeing a data point outlier like this one are only 1 percent, then we may choose to discount that data point, but if the chance of seeing an outlier that large were 30 percent, or 40 percent, then we might give more weight to that in our policy and management process, and so, given your understanding of the estimation procedures as they are now, is there a way to give some type of probability or chances of occurring for outlier data points like this one, outlier data points in terms of landings and effort?

MR. FOSTER: I have some thoughts on that, but I think Richard wanted to jump in.

DR. CODY: I just wanted to make the point that that 1 percent or 5 percent is dependent on what you're looking at. You could have a wave level, mode level, estimate that translates into an annual estimate that's unreasonable say once every four or five years, and so, depending on how you look at that, it can impact your determination.

MS. SAULS: I just want to point out that these are shore landings, and it's very unusual to land amberjack from shore, and so I think, in a data workshop level, this data -- I think shore landings probably wouldn't even be considered as reasonable to include in an assessment. I mean, I think this probably is one of those cases where it's not just an outlier and it's an unreasonable landing estimate.

DR. DUMAS: Right, but we might need some general and consistent method that we apply kind of across all species to decide whether or not an outlier data point should be discounted or not, so that we can be consistent across species and across stock assessments.

DR. AHRENS: If you could comment on kind of how the percent standard error is changing along with those estimates, if you have some indication.

MR. FOSTER: Sure. The application of both calibrations did add uncertainty, of course, to the point estimates, and they were reflected -- The uncertainty, sort of modeling uncertainty, is reflected in the new estimates of variance and the standard errors for the calibrated point estimates,

and so standard error is increased, and, off the top of my head, I cannot remember how much, but it was -- They are noticeably larger compared to what they had been, and, again, a lot of that was -- A lot of the increase was because, for all of the estimates prior to 2004, what was available prior to the application of these calibrations was the original MRFSS estimates based on unweighted estimation methodology, assumed simple random sampling, and systematically did underestimate variance for the point estimates. Because of those simplifying assumptions that were used, ignoring the complex design, it varied, of course, but it systematically underestimated the variances, and underestimated them a lot in some cases, and, typically, it -- That's true as a general statement.

In the cases of these outliers, and I apologize that the figures don't have confidence intervals on them, and it's just, with the three series, it was already fairly messy, and throwing in the confidence intervals makes it a lot messier, but, with the outliers, typically we also see enormous confidence intervals. The standard errors on these things are generally much higher than what you would see typically in the rest of the series, and so I think you could formulate criteria for flagging, potentially flagging, at least potential outliers for future review or consideration.

I don't have recommendations on what those thresholds should be, but, if those were developed, you could evaluate all of the series and see how frequently you've got values showing up, being flagged, to come up with sort of empirical probabilities in some way, to get back to Chris's point. Rob, does that answer your question?

DR. AHRENS: I guess I have one follow-up, and maybe it doesn't need to be talked about too extensively, is that one would assume that, if those variance terms are being carried over into the assessment appropriately, that the assessment itself should be acting as a smoother through that catch data, if you're fitting to the catch in the stock assessment, and I don't know if Kyle wants to quickly comment on that.

DR. WILLIAMS: I will jump in on that one. These models don't estimate catch very well at all. They are catch-at-age models, in the sense that catch is expected to be known fairly well, and, if it's not, there's not much other information in the model to tell you what catch should be, and so I don't think that's an alternative. Now, what we can do and what we do is we put that uncertainty in in terms of a bootstrap procedure, to sort of carry that uncertainty through, but the mean of that uncertainty is still going to be that point estimate that we fed in.

DR. AHRENS: They're conditioned on catch methods basically then.

DR. WILLIAMS: Pretty much, yes.

MR. FOSTER: Okay, and so I will move on. This is jumping forward now to I believe the last landings point estimate for greater amberjack, and, again, this one is -- You will see a pattern. Most of these, again, will be due to the FES calibration. This one is being driven by east Florida private boat mode. Again, you can see sort of, in the top row for landings, the region-level estimates, and then, below it, the east Florida private boat estimates, you can see the total change for that cell accounts for most of the change at the region level, and you can see the last two columns here, the percentages of the change, coming from the two different calibrations.

Again, if we look at the change ratios, they're just about right, at three, and these don't exactly match, because, again, I believe this one had increases across multiple waves, and, to keep the table kind of concise, I just presented the summarized information, and so these are 2016 annual estimates, again, for east Florida private boat mode, and so that's why the change ratios don't exactly match, is because there were increases across multiple waves, but, again, you can see they're very close to three, and so that's an indication that it's the FES calibration that's driving this increase, primarily.

Shifting gears now to releases, again, this one is -- This was a 1991 increase, and this one, again, is a release estimate showing up in shore mode, and, in this case, it's North Carolina. Again, you can see, in the release table, the region-level information and the information specific for North Carolina. In this case, it was Wave 4 in shore mode, and you can see that change accounts for the majority of the change in the region, and the FES changes is accounting for over 80 percent. It's almost 90 percent at the region level, but over 80 percent in this particular cell, and, again, you can see that the change ratios between the release estimates and the effort estimates are exactly lining up. They're at about three-and-a-half-times.

Again, almost 20 percent is coming from the APAIS calibration, and so, just to look a little bit at how that's coming in, this one, as I mentioned, kind of day was a design element variable that was included in the APAIS calibration raking methodology, and, in this case, the catch rates, or release rates, I guess, were correlated with kind of day, and so what's plotted here is the proportion of the APAIS sample weights, the sum of the APAIS sample weights, for weekday in blue and weekend in red, in the original unweighted MRFSS data, and then this would just basically be the split of intercepts between weekend and weekday, and you can see that weekends are sampled disproportionately at a higher rate than weekday.

Then, after the application of the APAIS calibration, you can see how this is adjusted, so that essentially they're brought back into their representative balance between weekend, making up a little bit less -- Still higher than the number of actual weekend days, but less than weekday, after the application of the raking.

Then, when we look at the amberjack intercepts, we see that a high fraction of those --Again, I don't know the actual number, the total number of intercepts, and it's probably a small number, but, of those, the majority of those are occurring on weekdays, compared to weekends in this cell, and so those weekday trips were upweighted compared to the weekend, and so, since there were more weekday trips, there was more net gain of sample weight for these amberjack trips, which increased the catch rate estimate. Again, when we have correlations between the catch rates and aspects of the design, we will see these kinds of shifts when we go from an unweighted situation to an estimation approach, which is appropriately using sample weights and calculating the estimates.

Releases in 2008 for greater amberjack, this one is coming from Florida private boat mode, and it's a similar setup in this case. It looks like the APAIS calibration accounts for about 20 percent and the FES accounts for the remaining 80 percent of the increase, and, again, you can see that the change ratios are very close between the catch estimate change and the corresponding effort estimate change.

When we look at the component, contribution coming from APAIS, in this case, it's not kind of day, but it's related to catch rate differences between high-activity and low-activity sites, and so, again, the higher-pressure, higher-activity sites were sampled at a higher rate compared to the low pressure sites, and so data intercepts coming from high-pressure sites are overrepresented compared to intercepts coming from low-pressure sites when we don't use any sort of sample weighting. We have to use the sample weighting to correct for that over-sampling at the higher-pressure sites.

In this case, the amberjack trips at the lower-pressure sites had higher release rates, on average, than they did at the higher-pressure sites, and, again, this is just specific to this one case, and that's not necessarily a true statement that we would expect to see everywhere. These changes are very data specific, and sort of idiosyncratic, in some ways, and so don't -- I am just trying to explain these individual points, and it's not necessarily true as a general statement across the whole time series, but, at any rate, to look at the plot here, I have sort of split the data into two groups. We have -- VL just stands for very low site pressure, or site activity, and then we have everything greater than that, essentially, and what we're looking at here are the -- The bars represent the sample weight percentages for between the original in blue and after we apply the calibration in red, so you can see -- What this is just showing is that the relative sample weights were increased, relatively speaking, for at the very low sites, and they decreased at sites that had greater pressure.

The numbers here give you the mean releases for trips with greater amberjack releases, and so you can see, at the very low sites, the mean release was about two-and-a-quarter fish per trip, that had releases of greater amberjack, and, at the other sites, it was lower. There was a lower release, and so, in a sense, this increase happened because we increased the weight on the trips that had higher reported releases, and we decreased the weights on the trips that had lower reported releases, and so, because of that, we had an overall net increase in the estimate and the catch rates for this particular case. Again, the take-home is, when we have catch rates associated with a design aspect, in this case site pressures that are used in calculating sample weights, we will see changes when the calibration is applied.

In 2016, this is releases, again, coming from both FES and, to some degree, from the APAIS calibration. It's Florida private boat mode, and the APAIS, again, is accounting for the large majority of the change, about 10 to 15 percent coming from the -- I may have said that wrong. The FES is accounting for the large majority, and APAIS for about 10 to 15 percent.

Again, this is 2016, and so it's that case where, as I described before for red porgy, with the overlapping time interval, we had to make the weight adjustments, sample weight adjustments, or corrections to account for that overlap sampling, and, when we did that, there was -- Later-day trips were upweighted relative to midday trips in the overlap window, and so, in this case, for amberjack, there is a net overall increase in their catch rates, because there is a net overall increase in the sample weights for those trips. I think we just have two or three more species, and then we'll be done.

For king mackerel, there is a 1990 and 2007 that we'll look at for the landings. For the releases, there is this sort of cluster of points here from 2003 to 2007. In general, this is all being driven by the FES calibration, and the APAIS resulted in some minor decreases, the triangle points here, and so the APAIS calibration resulted in some decreases in the release catch estimates compared to the original, the open circles there, but the FES increase is accounting for these changes.

I am going to focus, in this group, on this middle point, 2005, just because it's -- This is a point where the APAIS calibration is having more of an impact. In this case, it's decreasing. The APAIS decreased the estimates a bit, and then we'll also look at 2017. 1990 landings, again, and I'm going to sort of go through these kind of faster, since it's the same basic explanation, but it's coming from, again, a shore mode estimate, and it's sort of generally a rare thing to see for this species, and, again, the shore mode increases because of the FES calibration were much larger, relatively speaking, compared to private boat. You can see the change ratio in the landings estimate of about 9.3, and that corresponds to the effort change as well.

In 2007, again, this is Florida, but now private boat mode, and, if we look at the percent changes, the APAIS actually had a small reduction, but the FES was, again, accounting for all of the final increase in the estimates, and this happened over a couple of different waves in 2007, and that's why the change ratios are not exactly the same, and we're just presenting the summarized information, but, again, you had about a doubling of the landings, and that corresponds to about a doubling of the effort for private boat mode in this case.

Releases in 2005, this is one where the APAIS resulted in a sizable decrease, and I just, again, wanted to walk through kind of a different example of how the APAIS calibration could impact things, and this is for both Florida and North Carolina private boat modes, and I've got the releases, and you can see, again, there were fairly sizable decreases from the original base estimates to those for just the APAIS calibration, and you can see 30 to almost 50 percent decreases in these various cells, and then the FES calibration, of course, is sort of mitigating all of that decrease and then resulting in some increase for the final estimates.

In terms of why the APAIS had such a noticeable reduction for its contribution to the final estimates, this was due to changes in sample weighting across area fished, and so, again, area fished was one of the variables used in the raking process, because it's also used in estimation, and so what did we see?

The plots here, we have Florida private boat 2005 in the top panel and North Carolina on the bottom, and the left cluster of bars gives the distribution of total sample weight by area fished, and so blue is inland and red is state territorial seas and green is EEZ, and so you can see, in the case of Florida, or both, really, but there was a shift from -- There was a down-weighting of the state territorial sea trips and an increase in inland and just a minor decrease in the EEZ trips.

When we look at the king mackerel intercepts, again, the large majority of those are coming from the ocean areas, with just a very tiny fraction coming from inland, and, again, it's probably just an aberration of how the anglers chose to report their area fished, but, again, the overall take-away is there was a net shift away from the areas that had catch of king mackerel, a net shift of weight away from those to an area that had almost no catch of king mackerel, the inland, and so what this essentially did was reduce the release catch rate estimates coming from APAIS, and that resulted in sort of a lower increase than we might have otherwise expected once the FES calibration was applied, and that's basically the exact same thing that happened in the North Carolina case as well.

For 2017, this large increase, again, is due to increases coming from the effort changes with the FES calibration, and there is a couple of cells here. Primarily, it's Florida private boat mode, but there is also a similar change for North Carolina private boat mode. There are minor differences

coming from the APAIS calibration, and so the vast majority, or all of the change, really, is coming from the FES calibration.

Tilefish, and so here's a good example of a fairly rare-event species. It's not the most rare in the series, but it's very uncommon, and, here, I'm just going to look at a few of the landings estimates, and the releases really are -- It's a small sample size issue combined with the FES calibration change, essentially, but we'll look -- For landings, we'll look at 1985 and 2005, which is very interesting. This is a -- Well, sorry. It's interesting to me, and maybe it's just a headache for everybody else, but this is a case where we actually have such a large decrease coming from the APAIS calibration that there was a net overall decrease in the estimates.

This is a case where this was likely originally a -- Well, not likely, and you would -- I think any reasonable person would call this an outlier, and the APAIS calibration actually smoothed this point, and this is not the only time that we ever saw this. There were a number of cases where the APAIS calibration resulted in what appeared to be smoothing of outliers that had originally been in the MRFSS estimate series, and then we'll look at 2009 here as well.

Again, for 1985, here is a case where it's almost evenly distributed, the increase coming from the APAIS calibration change and the FES calibration change. In terms of what the APAIS calibration -- Why it had a large effect, and, again, this is kind of -- It's a result of the APAIS calibration, but it's really being sort of driven by some idiosyncrasies in the data, really one low-activity site being upweighted -- One low-activity site that happened to have a lot of tilefish intercepts, relatively speaking, that got upweighted.

We had a low-activity site, and it was Site 294, and this is east Florida private boat mode, Wave 3, and so there were a number of intercepts for golden tilefish at that site, and I don't remember all the details. I think it was just probably a single-boat trip with a group of anglers, and it was kind of an influential cluster, as I was describing earlier, and so, if we look at the percent of total sample weights coming from this site out of this entire cell, and so, again, only 1.48 percent of the total sample weight for this cell came from trips at this site, 294, under the base, but, when we applied the APAIS calibration, it didn't -- It increased the weights for that site by a factor of 1.5.

Again, it's still not accounting for very much. It's only now 2.76 percent of the overall sample weights for this entire cell for this one site, but that was enough. That led to a direct increase in the catch rate for golden tilefish in this cell, so that there was an increase, when the calibration was applied, that matched that same ratio. Again, it's going from 20,000 to 30,000, which is not 200,000 to 300,000, but it's -- Again, it's just sort of an interesting example of how the APAIS calibration can have an effect, in some cases.

DR. DUMAS: One question there. In this case, that change in that one cell had such a large effect because -- Is it the case that changing that one cell had such a large effect because all the other cells had zeroes, essentially, had zero catch, and so, even though this cell had a small sample weight, all the catch is coming from that cell, and so, when that one cell experienced an increase, even though its sample weight was small, it had a large effect on the final estimate, because it was the only cell providing any catch at all, or probably, and is that kind of your interpretation?

MR. FOSTER: Yes, that's in line. Again, if we had a broad distribution of golden tilefish trips at high-pressure sites and low-pressure sites, a good sort of robust mix, we would have had sort of

offsetting changes in the weights, and we wouldn't have seen a sizable increase, but, again, yes, as you're saying. Because the data were all clustered at this one site, again, it resulted in this unusual change as sort of an artifact of the distribution of the golden tilefish trips, or the sort of lack of distribution. They were all, again, clustered in this one. If the reverse would have also occurred, if all of the tilefish trips had been at a very high-activity site, that would have been down-weighted, and then the estimate would have decreased quite a bit.

Landings in 2005, again, this is the interesting one, and it's coming entirely from the APAIS calibration. In this case, again, it's an overall decrease, and so we went from over 65,000 fish, and these are numbers of fish, down to over 35,000, and so almost cut it half from the original MRFSS estimate, and, again, it's being driven by the APAIS design change. Otherwise, there would have still been an increase, if we only had the FES calibration.

This is North Carolina, and this is interesting. It's North Carolina charter boat mode, and so the FES effect did not apply to charter boat mode, and so, again, there is a little change here, because some of this wasn't just in the charter cell. The big driver was a charter cell, but there were some landings in other cells, and so that's why there is a little bit of a change, in one case, but, here, again, the lion's share of this change was coming from North Carolina charter boat mode, which, again, did not have any FES calibration, and that's why this ACAL values and FCAL values are the same.

Why did this shift? Again, specific to charter boat mode, in the APAIS calibration, there were several variables that we used included in the raking algorithm, because they are used in estimation, and, specifically, it's the coverage adjustment for the for-hire survey, and so, when charter trips are intercepted, samplers record the boat name or the state vessel identification number, registration number, and a determination is made as to whether that vessel is on or off of the for-hire survey sampling frame, effort survey sampling frame. Then that information is used to account for those trips taken by boats that are not on the for-hire survey effort frame.

In this case, the same weights -- What we have here in the plot is, in the blue, we have essentially trips -- The sum of the sample weights for trips that are taken on a vessel that's on the for-hire survey frame, and the red is trips taken on vessels that are not on the frame, and so, again, you can already see that the vessels on the frame were accounting for the large majority of the total trips intercepted, roughly 80 percent, but, when we applied the APAIS calibration, that changed. The weighting changed those distributions even further, to where the on-frame vessels -- Trips taken on on-frame vessels were accounting for over 90 percent, or almost 95 percent, of the trips.

The golden tilefish intercepts were roughly split, and, again, this is not a large number, and I don't remember, off the top of my head, how many total intercepts we're talking about, but they're roughly split evenly between the vessels on the frame and vessels off the frame, but the down-weighting for the vessels that were off the frame was sort of the net driver of this decrease, overall decrease, and so these trips were down-weighted considerably compared to these trips being somewhat upweighted, and I don't have the catch rates here, but it's likely that catch rates were higher for the -- Or they could have been somewhat higher for trips, just by chance, for the off-frame trips compared to the on-frame trips.

Again, you've got trips sort of associated with variables that were used in the raking, and those variables were used because they are either design variables or they are used in estimation, and we

had a change here, a fairly noticeable change, in the sample weight distribution among the levels of this variable, and that resulted in the overall change in the point estimates.

Landings for 2009, this change was driven by Florida private boat mode. Again, over 90 to 91 percent of it is coming from the FES change. Again, the second table here, we see the change ratios are lining up between the landings estimates and the effort estimates, again in this one cell, and so that's why they are exactly matching.

Gag grouper, we'll look at -- In the landings, we'll look at 1981, with this really large increase here, and then I just sort of picked out three points in the release series, 2000, 2003, and 2007. I believe these are all just driven by the FES change, because, as you can see, the solid black triangles and the open circles, which are the base and the -- I said that backwards. The APAIS-calibrated series and the base series, they're all very close, in these cases, and so, in general, that will indicate that the driver for the change is the FES calibration, and so I'll try to step through these kind of quickly.

For 1981 landings, it's Florida private boat mode, and it happened in Wave 2, and you can see that the APAIS calibration made some just very minor decreases in the estimates, but the overall driver of the change, again, is coming from the FES, and you can see the change ratios matching there between the effort and the landings.

For the releases in 2000, again, it's the FES driving all of the sort of final change from the original estimate to the fully-revised estimate, and this was from, again, Florida private mode, Wave 1 in this case. For the release estimates in 2003, it came from a number of cells, and they're all Florida, but it's private boat and a couple of different shore cells. Again, the overall changes are driven primarily from the FES calibration, and, again, you see a change ratio closer to two for the private boat cell, or case, and then almost seven to seven-and-a-half, roughly, change ratios for the two shore cells that contributed to this overall change in 2003. For 2007, in this case, it's just private boat mode, and the FES, again, is the large driver for the changes here, two to two-and-a-half-times increase.

In summary, and I know that was long and tedious, and thank you for your patience and hanging in there with me, and so, again, not surprisingly, the FES calibration is the main driver for most of the large changes. The fact that, yes, on average, the private boat increases are about two to three-times and the shore increases are roughly six-times, but, again, they vary considerably by state and wave and year and mode, and so individual differences, increases, can be noticeably larger, or, in some cases, smaller than that two to three-times average for private boat and roughly six-times average for shore mode.

The APAIS calibrations, again, are secondary, but they can be, as we saw, important in specific cases. One, they are the entire driver for any of the large charter changes, but, even for private boat and shore modes, if the catch rates were highly correlated with kind of day or site pressures, or even area fished, or, in the case of, as we saw with the time intervals in the more recent years with the weight correction, these can have an effect, and they can either be sort of additive to the FES calibration, resulting in an even larger increase, or they can somewhat offset the FES calibration, again, because the APAIS calibration was much more variable, in terms of its effect. We have made it, and so thank you very much. I appreciate that.

DR. SEDBERRY: Thank you, John.

DR. REICHERT: I have several questions, and some of them are clarifications. For gaps in the data in golden tilefish -- That was the only species that there were gaps, and is that because of lack of data?

MR. FOSTER: Yes. When there are gaps like that, the golden tilefish just weren't observed in those cells, and so, for the entire year, for the entire region, all of the modes, there were no observations of -- In that case, I think it was mostly releases, but, yes, there would be no observations in the data.

DR. REICHERT: But how is that different than people don't catch golden tilefish, or is there no difference? Let me pull up that, because there is gaps in both the harvest and the releases, and so, basically, the question is what's the difference between a zero catch and a gap in the data?

MR. FOSTER: Functionally, I would say no difference. What happens is there are cases, if you look in the estimation datasets, where sometimes you will see a zero, and sometimes you won't see any records, and the reason for the zeroes is that that means that there was an observation that had reported catch in one of the catch dispositions, but not in all of them, and so we have some cases, particularly in the early years, but even in more recent years, where maybe we have a landings estimate, or maybe we have a releases estimate, that is greater than zero, but we don't have the other one, and, in those cases, we will just have a zero for the catch type that we didn't have any values for. If we don't have a record with catch in any disposition, that's where we just have a -- It's just missing, and so that's sort of, functionally, how we have those differences in the data series. Again, I wouldn't -- I, personally, wouldn't assume that there were no catch in the cases where we don't have values, but it's just that, because of the somewhat rare nature of the species, we don't always observe them in the data.

DR. REICHERT: I don't think that I still understand the difference, because there are zero catches, correct? If a zero appears -- Like in the golden tilefish, if a zero appears in the harvest, that means that there were no golden tilefish observed in any of these intercepts, or is that an incorrect assumption?

MR. FOSTER: I suspect a lot of those low values may not be -- I would need to go back and double-check. I'm not sure how many of those are actually -- Some of them clearly are zeroes, but I don't know how many of them are zeroes and how many of them are just really low values, but, in general, if we have -- If we only have a landings estimate -- Well, in general, if we have a landings estimate, but we don't have any released data, we usually would have a zero value in the data, and I would have to back and double-check exactly if that's lining up here with what's in the plots. It might also just be an artifact of the plotting and if it is populating points or not based on how far the gaps are, and so I can't speak to whether the datasets actually have zeroes or they just have missing records, based on what we're seeing in the future.

DR. ERRIGO: I was just looking at the spreadsheet, and there are only three years that I have here, and my zeroes are not the same as the MRIP zeroes. These zeroes would be blank in the MRIP dataset, and there are only three of them in the recreational landings that are actual zeroes, but here is fourteen, and this is pounds, and so that's probably like a fish, and John Foster was

doing it in numbers, and so these numbers that are like 600 would be much, much smaller in his dataset, and so there is not as many complete blanks as it looks like.

DR. REICHERT: If you say there is three, there is three blanks in the golden tilefish series, and so those would be the years where there is no -- Well, are those the years that there were no golden tilefish recorded in any of the catch? The reason I'm asking this is because we are -- That's a follow-up question, and I think we discussed it at one of the previous SSC meetings. If we are looking at outliers, we tend to look at the high outliers. Well, again, a zero, some would argue, is also an outlier, because there is no landings of golden tilefish in year X, and so how do you treat that?

Also, you look at the outliers, the high outliers, and correct me if I'm wrong, that tend to be the estimates of the -- The three estimates tend to be further apart if you get a higher outlier, and they are closer together if you are closer to the zero value, and so you tend to not recognize this as an outlier, although it may be an outlier, and maybe I'm wrong with that assumption, and so, anyway, and I've got a totally different question, but can you comment on that a little bit?

MR. FOSTER: Sure. If there is a gap, then that absolutely, yes, does mean that there were no fish observed or reported through the intercept survey for that entire year, in this case for the entire South Atlantic for the given year, and so, yes, that is correct. In terms of the spread, if you will, of the data points as the calibration is applied, yes, effort -- The calibration of effort is sort of a multiplicative scalar, and so you're not just shifting a cloud of points up, but maintaining all of the sort of same absolute differences between them, and you're shifting points up, but you're also spreading the -- You're increasing the differences between them, because, again, we're multiplying by a number like two or six, and we're not simply adding some fixed amount to it.

Again, looking at the full time series, small, absolute points, while they will be multiplied by -- Conceptually, they will be multiplied by some number like two or six, and it's the same as the multiplier for the large values, and, in an absolute sense, they're not moving that much, as you were saying, compared to the large -- The points that start off as large values are going to move, in an absolute sense, a lot more than the smaller points, and so, yes, that's true, and was there another one?

DR. REICHERT: Well, there was another one, but I haven't asked it yet, and I realize that, when they're closer together, when they're lower, then they're less influential in the overall pattern, and so I understand that the higher points, or where they are further apart, that's more influential, in terms of comparing the various estimates, and so thanks.

The other question I have is when I looked at it, and, again, correct me if I'm wrong, it looked like, if you go through the explanation of some of these outliers, it seemed like a lot of them, and I just counted them, but the vast majority were the Florida private boats, and I was wondering what -- Do you know what makes that mode so seemingly influential? I counted like thirteen Florida private, Florida shore was three, North Carolina charter was one, North Carolina private boat was one, North Carolina shore was one, and I may have missed one or two, but the vast majority is Florida private boats, and so I'm wondering, what's so -- What makes that so influential, or do we know, or am I missing a point?

MR. FOSTER: Again, the calibration results can vary across states, but I think probably the bigger, or more important, driver here is just the scale of the Florida private boat fishery. Even before the calibration, it's generally -- For many species, it's generally comprising a large fraction, if not a large majority, of the catch of these species for the region as a starting point, and so it would make sense that -- I think it would just follow from that, that it's often going to be the cell that will account for the large change when we see it in the calibrations.

DR. REICHERT: Thank you.

DR. DUMAS: Could we go back to a slide that has a large -- Like Number 9? For one of the years that has a large outlier, and let's look at 2016, for example, and so you said that the new method with the FCAL numbers have larger standard errors, noticeably larger standard errors, compared to the base. With those larger standard errors, would the confidence intervals, or the error bounds, around the FCAL point estimate -- Would the error bounds around that point estimate include the point estimate from the base case sometimes, or often, with the larger standard errors?

With one of the outlier data points that is really high, if the error bounds around that high data point are really wide, then it might include the lower point, and so they're not really different, and so, even though it appears to be a much different estimate, it might not actually be a different estimate. They might not be statistically significantly different if the standard errors are larger with the new method. That may or may not be the case. It may be the case in some situations, but, if that's the case sometimes, if the standard errors are enough larger that the error bounds around those very large outliers, and also around the very small outliers, if those error bounds include the point estimates from the base method, then the point estimates of the new FCAL and the point estimates of the base might not be statistically significantly different.

In that case, it seems like sort of a take-away from this discussion that we've had for the last hour would be that the old base method greatly underestimated the variances. The new FCAL method better estimates those variances, but the better estimates of the variances are larger variances, and so they have larger error bounds, and so, when we see an unusually large or small value, we are now, with the new method, less sure about it, but we're correct in being less sure about it.

Before, we were overconfident about less-extreme values, and, with the base method, the points don't seem to be as large or small, but we really should -- We were overconfident about those points. With the new method, our points are, perhaps, sometimes much higher or lower, and we see outliers more often, but they have wider confidence limits, but we're more confident about those wider confidence limits, and so the new variances are larger, and so they have wider confidence limits, but we're more confident about it. We're more sure about that than we were before, where we had fewer outliers, but that's because we were using a method that had incorrectly produced smaller variances, and we had, perhaps, tighter confidence intervals, but those were incorrect tighter confidence intervals.

MR. FOSTER: Yes, I believe that I agree with everything as you described it, yes.

DR. DUMAS: Okay. So, if that's the case, going forward, is the only way to improve things to get larger sample sizes?

MR. FOSTER: Well, I wouldn't say that's the only way, but that is certainly a good way, and, if we can increase the sample sizes in the intercept survey, that will only help. Now, in addition to just across-the-board increases in sample size, and I think I may have mentioned this, it's not -- Well, a couple of points.

One, moving forward, we don't have the additional uncertainty about the estimates that is coming in from the calibrations, and so that's an added piece that just affects the historic -- Pieces that affect the historic time series, but, in addition to that, I think we can still make some modest improvements in the precision through improvements to the designs, specifically the APAIS design, specifically through more tailored stratification.

We have limitation on what we can do. We can't -- With the existing sample sizes, we have to have a minimum number of sample units in each of the strata, and so, if we create a whole new dimension of strata, that's a lot of additional sample that's got to come out of the existing strata, and so there is a limit to how many additional strata we can create, but, in a number of places, I think we could create some version of the offshore site group that would allow us to sort of better target sample to sites that have trips that return with catch of these kinds of species, but, again, that's not something that will work everywhere.

Not all states have sites situated and specific geography so that those trips end up getting clustered into specific sites. Some, they are just diffused across a wide number of sites, and maybe those sites are already the high-pressure sites, and it's just that those trips are rare enough, even at the high-pressure sites, that we just don't see them, and so we may not -- We are limited in how much improvement we can make through that sort of a process, but, yes, certainly a simple increase in sample size will help. It's the rising tide raising all boats, essentially.

DR. DUMAS: With the improved methodology, we could better target that increased sample size, so that we get the most bang for our buck and by increasing our sampling in the places where it would best improve the estimates.

MR. FOSTER: Right. With the spatial stratification and the temporal stratification, we can not only target sample to specific sites, or areas of a state, but we can also target sample to specific times of day, and so, if those trips tend to come back in the afternoon, or early evening, we can put sample towards those time blocks as well.

DR. WILLIAMS: I was just going to inject a note of caution about the term "outlier" that gets thrown around a lot and make sure that people are on sort of the same thinking, in terms of what an outlier really is, sort of from a statistical standpoint. Outliers are not necessarily like pornography, in the sense that you will know them when you see them. They can be insidious, and they're always with respect to what your expectation is, and so you have to be very clear about what is your expectation when you're saying something is an outlier.

In the case, when we're looking at a time series of landings, and we see what we call a spike, and we're calling it an outlier, our implicit assumption then is that the expectation is that the landings should be flat, or level, and that's not always the case, and we shouldn't always assume that that should be the case. I am just injecting that bit of caution, that, really, probably the way we should be looking at outliers in this sort of survey is one extreme idea would be to take and estimate a mean for all the cells for the whole time series of the data for MRIP, so we have a mean for every

strata, and then you look at the deviations from each of those strata, and then you want to see if that deviation falls outside of some criteria and then declare it a statistical outlier in that sense, but I just wanted to inject that note of caution about this term "outlier" that gets used sort of, in some cases, inappropriately.

DR. SEDBERRY: Thanks for that, Erik.

DR. LI: I would just like to follow-up on Erik's comments. Outliers are not always influential, and so there is another term of influential points, and so they are different. Outliers are not necessary to be influential.

DR. SEDBERRY: Any other questions or comments on this agenda item?

DR. SCHUELLER: At the risk of opening a can of worms, this is all well and good, and it's great that we can pick some points off the thing and take a look at it and say, yup, there is some interesting things going on here, but then what? From an assessment side of things, and a data workshop side of things, how is this going to be handled, and is it worth it? All of you sitting around the table, is this really worth it? One data point that may or may not even have any influence of the assessment, do we want to go down this hole? That's the question around the table. Then, if we go down the hole, what the heck are we doing, because we can't -- We're going to go down there and say, yup, here's a data point, and it's a real data point, and what do we do with it?

DR. SEDBERRY: Yes, and these data points will come up independently at every assessment for every species and maybe be handled differently every time they come up, unless perhaps we could provide some advice, as the SSC, on what to do, generally, for these kinds of things.

MS. MATTER: As the person who is sitting at the data workshop in the recreational seat, I am very nervous right now, because, as many have said, and Marcel mentioned that nobody is talking about the zeroes, or the lows, where we have nothing, and we're talking about are they really outliers, or are they against perception, where all of these are such great points, about whether, okay, it's there, but is it really going to influence, but, at a practical standpoint, we're talking about assessments in the whole region with different players at different stages, and we're talking about assessments that don't always have a data workshop. Sometimes, where it's a standard, we're just doing it through webinars, and so we don't have a whole bunch of people in the room trying to deal with this.

I credit John for all of the work that he did looking at these, but all I was thinking, while I was going through, was how are we going to do this on an assessment-by-assessment basis? Like, we don't have all of that data coming in to inform when we're looking at these points, and so, without that information, without the consistent methodology in place and consistent people at each assessment, and I'm thinking this is the South Atlantic, but we're also dealing with assessments in the Gulf that I strive for consistency across regions, unless there is a valid reason to, but, if it's a matter of it just depends on who is in the room and what the opinion is as to what the perception is, I just get very nervous.

DR. SHERTZER: I just want to follow-up on what Vivian said. I realize this is an MRIP discussion, but, when we go to assessments, we also have commercial data, and we have indices

of abundance, and we have age comps, and we have length comps, and so, if we're going to go down this rabbit hole for a data point on a time series from recreational landings, we need to do this for all of the time series, all of the different data sources that we have.

DR. REICHERT: That's exactly the point that I was going to make. I remember the conversation with blueline tilefish, where we were talking about outliers, quote, unquote, outliers, and those outliers had nothing to do with the calibration or the new estimates, and they were there no matter what method was used, and so we have always dealt with those issues, and it's a very complicated issue, as you mentioned, but that's a different discussion than the new estimates of recreational catch and effort.

I would like to separate those a little bit. If an outlier, as John indicated, in some instances is -- Again, if a, quote, unquote, outlier is a result of the new method, then, yes, I think, in this conversation, in this discussion, it's important to talk about that, but, in general, the treatment of those outliers is a related but separate discussion, and I'm afraid, if we are going to mix those two, then we may not get anywhere in answering or in addressing our terms of reference, and I would love to hear from people who disagree with that, but, again, I remember the blueline tilefish conversation, where we were re-discussing points that were discussed at the data workshop, and then we as an SSC approved that stock assessment, and then we started discussing those same points again when we were talking about the recalibration, and so maybe I'm repeating myself now.

DR. SEDBERRY: You know, our next agenda item is the mini data workshop, I think, to look at the recreational catch estimates for species currently being assessed, and those species are going to have these kinds of outliers in them too, and so I think it would kind of give us a preview of what's going to go on at the actual data workshops when these kinds of things come up for every species.

Where do we go from here? Do we have any additional discussion on the presentation we just heard? We do have still scheduled for this afternoon the mini data workshop to review some recreational catch estimates for the three species that are currently being assessed, and we -- Golden tilefish, greater amberjack, and red porgy were all in the presentation we just heard, and I'm not sure what else we can kick around with it, but we still have a few minutes left this afternoon, and we can continue to discuss them.

DR. REICHERT: Well, I have a suggestion. Perhaps we should wait -- We have about ten or fifteen minutes left, and we can go back to the terms of reference and see where we need to make recommendations and what the council is asking of us, and that may be, for the remaining time today, a little more productive than starting a conversation, or starting a discussion, and perhaps pick that up tomorrow, and that's just -- I just want to throw that out.

REVIEW OF TERMS OF REFERENCE AND CONSENSUS STATEMENTS

DR. SEDBERRY: I agree with Marcel's suggestion, and Mike is just filling in some details of these terms of reference questions that need to be filled out, and we can possibly get them done this afternoon, and, as Mike just pointed out, the mini data workshop species are the ones that we just covered and talked about the data quite extensively, and so we have the action items, if you

will, from the terms of reference that we do need to address, and so let's look at those and see what we can fill in on them for the rest of this afternoon, if that sounds good to the group. Any objection to that? Okay.

The first item we have up there is describe how the sources of disparity and data issues identified for the five species examined above may affect the estimates for other South Atlantic Fishery Management Council species. Again, this is kind of getting back to Amy's point about we're going to have these issues with every species, and are there some general guiding principles, perhaps, that we can come up with?

DR. REICHERT: Well, I would like to, again, propose to split the two, the one where the new method is the source of the additional variability and then the other one, the inherent variability, irrespective of whether you use the old or the new method, and that may be something for a later discussion.

DR. SCHUELLER: I've been sitting here thinking about your comment, Marcel, about separating the two things, and I agree with it, but I'm not sure, in my head, how I would do that, and I'm sitting here thinking about we're supposed to have this sort of mini data workshop for these species, but, in reality, if this was a data workshop for red porgy, we're going to get the FCAL numbers, and those are the series we're going to be working on. I don't think, in the future, we're going to sit here with all three of these sets of numbers and have these same discussions, and so they are -- I mean, it is a different topic, and I guess this particular term of reference is specific to a data workshop, right, or not? I mean, it's a little -- It's mixed, and so what are we looking for here?

DR. SEDBERRY: I think the data workshop idea was to look at the species that are currently undergoing assessments that are on hold and talking about what we can do to un-hold them, to get them moving forward, and so I think what you just said is exactly right, that we're only going to be looking at FCAL, and we're not going to be looking at these other things.

DR. ERRIGO: I think we just need to -- If there is anything that the SSC is uncomfortable with within these three species data streams, we can go over and see -- Here is why this change happened, and are you comfortable with that, or do you think it needs to be modified or changed, and then you can make a recommendation, perhaps, to the assessment that says, oh, for greater amberjack, we think you should remove the shore catches, because you don't really catch greater amberjack from shore or whatever. For porgy, that looks fine, and it's an artifact of the calibration, but everything checks out, and so we're good with that. We're comfortable with moving forward. I think that's the kind of thing that we need.

DR. BELCHER: I get nervous about us getting into the diagnostics of this, and the only reason I say that is, any time we go to the SEDARs, we're generally looking at what's already been put through by analysts that have already put their comments and concerns into it. Are we going to, every time we see something like this, going to start unraveling and have to have John come in and drill everything back down to a cell level to realize that it's the same things that have kind of plagued -- It's not plaguing the survey, but it's the issues that we know that has already existed for the survey. It's not designed around the species. You get the best estimates you can based on how the overarching survey is designed, which is getting at effort with ancillary data relative to fish species that we're able to marry together.

The only way you really start figuring out where all those things go weird and error bars go wide and all of that is when you get back to figuring out what cells are driving what, and we're, as a group, not going to be doing that. That would mean that, every assessment, we've got to unravel all the way down to the bottom before we're comfortable moving forward.

DR. SEDBERRY: The bottom line is we are looking for a way to go forward with these assessments. The revision assessments and the other ways that the SSC has looked at these new estimates have caused the SSC heartburn and have caused us not to be able to come to consensus on the use of the revision assessments and the data in setting catch limits and fully evaluating these assessments, and so we need to, based on what we have learned so far, what we have learned at this workshop, come to some kind of recommendation for how these assessments can proceed.

We're not going to get any more than what we've got right now, and so what is our recommendation? What is the SSC's recommendation to incorporate the MRIP data into the assessment? Can we move forward with the assessments that are now on hold? Can the data workshops that are scheduled look at this data and make evaluations of these datasets to use in the assessments? We need to come up with a recommendation. I know it's late in the day, and we probably don't want to do anything rash, but we do need to do something.

DR. NESSLAGE: I too am nervous about doing something rash at 5:19 in the afternoon. I thought we were going to put on hold a number of our questions about the calibration until we heard from Jay tomorrow, and so should we really be making a decision about these until we've had the concerns addressed, and that's my comment.

My question was please help me and correct me if I'm wrong, but I thought, ultimately, even though, yes, I remember, personally, getting worked up about some of the blueline outliers, in air quotes, for the record, I thought, in the end, you guys convinced me that it was the calibration for that and black sea bass that were causing the change in stock status and not the outliers, and correct me if I'm wrong. If that's the case, then we're back to, before we can make a determination on those two species, we're back to the calibration issue again, and so I would say put it on hold, even though I know we don't want to, probably.

DR. SEDBERRY: You were right. We'll get some more information tomorrow on the calibration. We'll get that information tomorrow, in the afternoon, giving us very little time to --

DR. ERRIGO: It's tomorrow morning.

DR. SEDBERRY: Bonus. We'll have during lunch to work on it.

DR. REICHERT: What time?

DR. SEDBERRY: Hang on a second. My brain is dead.

DR. ERRIGO: It's not until 1:00 p.m.

DR. SEDBERRY: Right, and so we can -- Marcel, did you have a --

DR. REICHERT: I have a concern. I mean, if this presentation is going to be at 1:00 p.m., and we adjourn at 3:00, that gives us almost no time to thoroughly discuss and come to recommendations and then help George with the report, because we only have three weeks for the report, and so, anyway, that's a timing concern that I had.

DR. LARKIN: I will make this real quick. I know we're all talking about moving forward with the assessments, but we also need to think about moving forward with ACL monitoring, because we're using these peaks as well for closures and accountability measures, and so I just wanted to make that quick point. Thank you.

DR. SEDBERRY: That's an excellent point. I am going to suggest that we recess for the day, and we will strategize after we recess, the executive committee, and come up with a plan to get the additional information tomorrow from Jay and digest it somehow and come up with recommendations and any consensus statements that we need to make and move forward. Again, since this morning at eight o'clock was so much fun, let's do it again tomorrow, because we have a short day tomorrow, and we really still have quite a bit to do, and so let's meet tomorrow morning at eight o'clock. Mike has something else to say before we recess though.

DR. ERRIGO: In order to help us start early with the discussion and conversation, John Foster said he would be willing to answer questions about the FES calibration and that, and it might be beneficial to go over the part of the presentation that he zipped by and see if you have any questions on that, to help us start the conversation, and then Jay Breidt will fill in all the gaps at one o'clock. This way, we can just make some progress.

DR. SEDBERRY: That sounds like a good idea, and thanks, John, for that offer, and so we will do that beginning at eight o'clock in the morning. Thanks, everybody, and I know it's been a long day, and I appreciate all your efforts and patience. Thanks.

(Whereupon, the meeting recessed on August 20, 2019.)

- - -

August 21, 2019

WEDNESDAY MORNING SESSION

- - -

The Scientific and Statistical Committee of the South Atlantic Fishery Management Council reconvened an MRIP Workshop at the Town and Country Inn, Charleston, South Carolina, on Wednesday, August 21, 2019, and was called to order by Dr. George Sedberry.

DR. SEDBERRY: I apologize for our little bit of a late start here. We had to kind of go over where we are and where we want to begin today to make sure that end up, at the end of the day, at three o'clock this afternoon, with some tangible things that we can take back to the council for their meeting in September, and one of the things that we need to do that is we need to go through the TORs and make sure we have addressed all the questions and issues in the terms of reference,

including those letters from the states that are part of the Term of Reference 1, but we also felt like that, in order to be able to fully address the TORs, that we need a little more background information on the FES calibration.

We're going to get some of that this afternoon from Jay Breidt, but we also -- We kind of felt like we needed some kind of fill-in information, so that we could move forward this morning, and so John Foster has agreed to present the stuff that he kind of glossed over the other day, at our request, to go back to it and give us a short overview of that FES calibration, and so that's what we're going to start with this morning.

Then we'll go back and make sure that we have addressed the South Atlantic Council's motion, the terms of reference, that the steering committee and the SSC as a whole has approved, and that the council has approved, and then any other loose ends of things that we might need to address, the few things that have come up over the past couple of days, like hidden fishing effort and things like that, and so we'll start this morning with some more information on the FES calibration.

DR. WILLIAMS: The one thing that I would like to make sure we walk away with is clear guidance on what to do with the assessments that are on hold and the time series, the MRIP time series, for those, and so add that to your agenda.

DR. SEDBERRY: Yes, I think the council would be very much interested in that as well. When you're ready, John, go ahead.

MR. FOSTER: Okay. As George said, I will pick back up with the FES calibration discussion. Again, this first slide is just sort of a review of why the FES estimates are -- Some of the factors that contribute to the FES estimates being different from the CHTS, and I'm going to move past it, because we've covered it, pretty much, at length.

I mentioned yesterday, again, that the overall model framework is a standard one used in small area estimation, the Fay-Herriot model, and, again, it's a linear mixed-effects model, and we'll talk about the different effects in just a minute. This originally was developed, as I think I mentioned yesterday, to produce estimates of sort of mean income, personal income, in small areas at the county level, in particular, developed by a couple of folks at the Census Bureau.

It's a mixed model, and it has both fixed effects and random effects, and so, in terms of kind of the general inputs to the model, or how the model is -- What we're attempting to estimate with it, we essentially have three components that go into the estimated effort that we try to tease out in the model. There is, of course, true effort, which we don't know, and that is modified by non-sampling error that's present in every survey, and these are things like measurement error, coverage error, non-response error, reporting error, recall error, all of those kinds of things that we attempt to account for in how we design the surveys, but we don't have direct control over.

Then there is also sampling error, which is the piece that we do have the most control over, and that's, of course, affected by things like sample size, how strata are created, things like that, and we really don't have a perfect way to disentangle sort of true effort from the non-sampling error that kind of affects how well we are estimating effort.

The components that sort of go into these different pieces, one, we have -- This is sort of a time series -- Conceptually, it's sort of a time series framework, but we have a trend. We have sort of long-term trends, and these are sort of the year-to-year kind of changes, and this is being modeled -- This is input into the model as population estimates from the U.S. Census Bureau, annual population estimates by state.

Then we have sort of seasonal effects, and these are the within-year sort of month-to-month or two-month wave changes in fishing effort, again, that occur with some frequency each year seasonally, and these are in the model as indicators for the five or six two-month waves in each state. Then we have what we're just calling regular terms, and these are terms in the model that we model as random, and this is sort of the -- The other two that I mentioned were considered fixed effects, and these are regular terms, or random, and they're just there to account for sort of changes that don't follow -- That don't come from either the trend or the regular seasonal pattern, and they are modeled with sort of an assumption of a normal distribution, a mean of zero and a variance that is estimated from the model. Then we allow for interactions between all of these effects in the model, terms in the model.

Additionally, the -- I should have started, and I apologize, at the top. We develop separate models by mode, and so the shore model is totally separate from the private boat model. They have the same basic framework, but we didn't include mode, fishing mode, in the model. We just had separate models for mode. We do, however, have terms in the model for state, and, again, these is a fixed effect, and we have the interactions among all of these terms.

While a number of things were looked at, in terms of additional auxiliary information to feed into the model to try to get at estimating the non-sampling error piece, the only one that really held up, through diagnostics in the model, was the wireless effect, and so, again, this just gives a similar plot that you've seen before, and this is just summarized to the national level, but, again, the growing percentage of wireless-only households in the U.S. population, and these estimates were fed into the model by state, and there is actually the National Health Interview Survey that Rob talked about a good bit yesterday that provides us information.

They actually put out estimates of these percentages twice a year, and the model -- We input them into the model at the wave level, and so we just did simple extrapolations from the two estimates that this survey produces every six months, or they don't release them every six months, but that's the time interval that they provide, that they produce the estimates for. We just extrapolated them to the wave level, and it was for each state, and so we didn't use a national or a regional average number. We have state-specific estimates of these wireless-only percentages.

A key point here is that this does not run through the entire time series. The series begins in 2003, and we also attenuated it back to 2000, so that it's essentially out of the model prior to 2000. Then the sampling error is input to the model as well, and that's coming directly from the variances from the two surveys, the Fishing Effort Survey, FES, and the CHTS. Again, this slide is just showing that there were interaction terms in the model for all of these different pieces.

The model itself can actually calibrate in two directions, and this was kind of a key piece that we needed to -- The primary use was to calibrate the historic time series from 1981 through 2017, and that's given sort of by this formulation, where we estimate model terms from the three-year benchmark period, which was 2015 through 2017, where we have side-by-side estimates from the

FES and the CHTS. The model can estimate -- We estimate effects for all of the different terms that I mentioned, including their interactions, to produce coefficients for the mail survey effort, and then that can be used to produce a prediction for the mail effort going back through time.

Then, to go forward, and so for 2018 -- After we stopped the CHTS in 2017, but we can still produce CHTS-like estimates using the calibration model going forward, and we have to do that for -- We do that for a number of -- To support ongoing management for a number of species, which still have ACLs set using the CHTS-based series of estimates.

Those were all of the slides that I have on this. Again, it's a fairly high-level summary of the modeling, and I'm happy to answer additional questions, but Jay Breidt will be able to give you much more detail and a much better description of the flavor of the models. I am not a modeler myself, and I try not to play one on TV either, but I am happy to take questions.

DR. SEDBERRY: Are there any questions?

DR. SHERTZER: When you're taking the model backwards in time, do you control for some of these interactions or effects that may disappear as you go backwards in time, for example the wireless effect?

MR. FOSTER: Yes, and so, as I mentioned, that effect is only in the model until 2000, and then, essentially, it's just zeroed out. The indicators are zeroed out through the rest of the time series. We don't have specific terms in the model for all of the other potential sources of difference between mail and phone. They are just, essentially, kind of as an aggregate effect of the difference, and, again, that's kind of a central limitation of this approach, is we only have those three years of side-by-side benchmark to try to estimate that difference, but, yes, the wireless term is out of the model, essentially, prior to 2000.

DR. LI: I do not know what exactly the fixed effects factors and the random effect variables factors are.

MR. FOSTER: Again, the fixed effects are essentially a series of indicators for state and wave, interactions of those, as well as the state populations, the human population estimates for the coastal states, and those are all fixed effects. The random, or irregular terms, they are a series of indicators in the model, and they are there essentially to account for changes in the effort that don't, essentially, fit well through either the trend or the seasonal terms, and, again, the seasonal terms were those indicators for the two-month waves.

DR. LI: The reason I asked this question is the selection of predictor variables is very important, and so, if you have some variables that are not the critical variables driving the dynamics then, it's meaningless to do prediction. Also, I have read the materials, and it seems like, because you have to backward project, and then the selected variables have to be the common variables, they will not be available back in time, and also available in your benchmark time period, right? That will limit the selection of variables, and there might be some important variables out there, but we cannot use it, because it's not available both back in time and the benchmark time periods. I am not sure if this is the case, but this might be something to look into, and, here, you used the population size, and I understand that part, and you used the seasonal indicator and the weight, and I understand it's a category variable, but you just -- And out-of-state.

I am not sure how important those factors driving the fishing effort -- Just based on our knowledge, understanding, and not based on the data, and I know your data is only three years of data to come up with this relationship, but I just want to say, in this case, I would rather select variables based on my understanding, my knowledge, about the fishery, rather than based on the three years of data, because there might be some factors that are very important, but not being used here, because of the limitation of the time period.

MR. FOSTER: Sure, and so I think Jay Breidt will be able to speak to sort of the model formulation process and how they went through sort of -- It was more than simply kind of a step-wise regression approach, but the sort of diagnostics that they used to identify the terms that would remain in the model, and I believe that -- Again, he will be able to describe it better than I can, but there was sort of, initially, kind of a Bayesian approach to building the models and evaluating terms to include in the model, but the final model used to produce the estimates, or for the calibration, was simpler to implement this using this Fay-Herriot approach, which, again, is a sort of standard way of doing it, compared to what they had done with the sort of Bayesian approach to identify terms to include in the model.

I guess a key point here is this approach was trying to come up with a way to estimate the differences between the two surveys. It's recognizing that neither survey -- Even though the FES has a series of improvements over the CHTS, we're not looking, in this approach, to try to estimate the true effort. We just don't have the information available to really do that, and it's just an attempt to model the difference between the two surveys and apply that backwards through time to generate FES estimates or forwards through time to generate CHTS-like estimates after the CHTS was ended.

There's a lot of -- You could think about a lot of additional auxiliary information that we might want to include in the model if we were trying to estimate true effort, but, in this case, it's really just trying to get at the difference between the two, and, again, there was a number of variables that were looked at, but the wireless effect was the only one that, through their process, was significant and retained in the model, and I know that's not necessarily a good response, or a full response to your question or comment, but, again, I think he'll be able to talk us through, in more detail, the sort of model-building process that they used.

DR. LI: Just to follow up, let's go back to Marcel's question yesterday afternoon after you looked at those trends, time series trends, with outliers, and you explained that most are from Florida private boat, and then we wanted to know why Florida private boat, and so I'm thinking the variable you selected here to predict the trends, and then maybe something more than what we already used, more variables, important variables, are there that haven't been used yet. For example, like Florida private boat. Maybe, when you look into the details, maybe some management changes or something there, and maybe they opened up new fishing grounds in some years, and we don't know, and maybe the management relaxed a little bit, and so there is more things there, I feel like more important things there, than like maybe the waves or something. That's just a guess, and I don't know yet. I don't know the answer yet.

Then another comment is you mentioned Bayesian, and Bayesian is an estimation process, and it's not a model -- I would not say it's a model-building process. Like, when you said what variables you select, that's model development, the building process, and Bayesian just helps you with how

to estimate parameters, and it's nothing to do with what variables you decide to put in the model, and, here, in terms of the estimating process, and I'm just thinking back to Chris Dumas's comments yesterday.

If we believe non-sampling error is big in this case, and then, for example, you see the coverage, the survey coverage, explaining most of the differences between the two surveys, which means that now the sampling error might be a big component, in this case maybe you want to try out different estimation methods to incorporate more uncertainty there or you can increase the sampling size. Then, if, after we incorporate more uncertainty in the model, and then maybe, after you estimate the confidence interval between the two surveys, then they overlap for those outliers, and that means they are not significantly different.

MR. FOSTER: Okay. You covered a lot of ground there. The first thing I guess I would say is the -- The CHTS effort series that go back through time that were design-based estimates, all of the effort estimates are inputs to this model, and so, going back through time, the CHTS series, to the extent that it picked up those sorts of changes in effort that could have been associated with regulations or weather events, whatever may have made a particular individual point estimate deviate in time from an expected trend or seasonal pattern, it's input into the model, and it would be picked up, potentially, in the terms as those irregular terms, but, again, it is an input into the model coming from the CHTS series, and so, even though the benchmark is only for those three years, the CHTS effort series for the entire time series is still informing the model as well as the key piece, I think, that's the real limitation here, among just not knowing true effort, is the fact that there is just such a time series of mail estimates that was available for the benchmark compared to almost forty years of the phone estimates.

Again, once we have sort of a much more robust time series of the mail estimates, I think, at that point, it would absolutely be worth revisiting the calibration, because there will be a lot more mail time series available to inform the model from this side of the benchmarking period, and then, again, as the -- I'm not exactly sure how best to address your comments on the variances or the precision.

The variances from both the CHTS effort estimates as well as the FES effort estimates are terms in the model. I mean, they are also inputs into the model, and so the model -- The variances that come out that are for the model predictions reflect -- They incorporate the uncertainty that was originally estimated for the point estimates, as they are design-based variances, and so those are inputs into the model as well, and I apologize, but, again, I will have to leave it to Jay to sort of better describe how they are incorporated and how they are reflected in the final estimates of precision that are associated with the individual model predictions, but I'm happy to attempt to answer.

DR. LI: Thank you, John, and I'm sorry to throw out a lot of things, and I want to say like, if the estimation method for this model is Bayesian, which is good, then we can adjust priors for those sampling errors to incorporate more uncertainty if you believe this error is big. Now, if you are not using Bayesian yet, then maybe you can consider some method, like Bayesian sampling method, to incorporate more uncertainties.

DR. VAN VOORHEES: I just wanted to add that, when we set out to do this, with getting help from our expert consultants, we were looking around and thinking that somebody else might have

done this, going back and calibrating back through time to preserve a time series of information when they have made substantial changes in their methodology, the survey methodology, but, in fact, we really couldn't find other examples, and we have kind of realized that we're unique, in the sense that your customers require a comparable time series of catch statistics, and so we have decided that we needed to do the best we could to come up with an appropriate way to do that.

In most cases, people were just saying, well, we changed our survey methodologies, and so the estimates we're producing now are not comparable to the estimates we have produced in the past, and so you know that, and you need to figure yourself if you need to calibrate a time series, but we're not doing it for you, and so it really is challenging, and we don't deny that, and, as John pointed out, only having three years of a side-by-side comparison certainly limited the ability to go back through time to accurately hindcast everything.

I also wanted to add that the transition team that we formed actually looked at this as we were approaching it, and the model hadn't been developed yet, and we were trying to figure out, okay, how long do we do the surveys side-by-side, in order to get the information needed to support a calibration, and there was a bit of fuss over that, because we thought there is such a big change here in the estimates that we need to account for it as soon as possible, and so there was discussion about possibly doing it after two years of side-by-side, just to do it sooner, but we finally decided that, no, we would settle on three years of side-by-side and do the best we could. We couldn't really wait longer to account for this difference, given the many assessments that were in line to be done, and you wanted to get the most accurate information into those assessments as soon as possible.

DR. REICHERT: Thanks, John, for that overview, and I may have misunderstood this, but you said the wireless effect was the only factor that was retained in the model, or did I misunderstand what you said?

MR. FOSTER: I will clarify that, and so, in addition to the other terms that I mentioned, state population, the wave indicators, the interactions of those, wireless was the only additional term that was included, evaluated, for model use that was related to the non-sampling errors, and so those other terms were all still in there, and this is in addition to the CHTS effort estimates, their variances, the mail survey estimates, the FES estimates, their variances, and all of those are in the model as well, but the wireless was the only one that was looked at, again, specifically for the non-sampling error.

It was the only one that held up, and I apologize, because I'm trying to think if we had other terms in there that we looked at, and we included things like non-response rates for the CHTS and the mail survey, and we had, I think, a proportion of -- From the CHTS, I think we had proportion of respondents that were female, or proportion of respondents that were proxy reporters, and so one person answering for the whole household. The CHTS attempted to speak directly to all of the members of the household that took fishing trips, but, in almost all cases, it was just one person answering for the entire household, and so we tried to look at those things.

We evaluated some other external variables, I think, and I believe we looked at license counts that the U.S. Fish and Wildlife has for states, but there were a lot of issues with a number of the external time series, because, again, over such a long period of time, there is lots of inconsistencies in how some of the external series were -- The information available for the external series, and it would

have just led to sort of additional modeling efforts to try to calibrate those series to then use in this calibration, and so there's just a number of things that limited the utility of the external series.

DR. REICHERT: Thank you, and, earlier, you mentioned that, prior to 2000, the wireless effect was zeroed out, and so there were residual effects, and so, earlier, I believe you mentioned that the wireless effect and the gatekeeper effect were the primary effects, and so, prior to zeroing out the wireless effect, do you think it was mostly the gatekeeper effect that would account for some of those differences, or were there other factors that may have played a role, in terms of relevance?

MR. FOSTER: The wireless effect is the only sort of -- Of all of the kind of net differences between the two surveys, that's the only one that is sort of explicitly in the model. The others are sort of just kind of in there in aggregate, as kind of a -- You could think of it as sort of an overall survey design effect, but it's just the aggregation of all of the other differences between the two surveys.

Certainly gatekeeper is -- Of all of the remaining differences, I would say, yes, gatekeeper is one of the largest ones, but there is just a number of differences, and Rob sort of spoke to a number of them yesterday. The fact that the CHTS -- Not only is it a phone survey compared to a mail survey, and there are often differences just from the mode of contact, and they have been demonstrated in more general survey research literature, but the fact that the CHTS was a cold call with no time for a respondent to sort of think about the questions and consult, potentially, with other members of the household, or get some sort of advance notification that the survey request was coming, compared to the mail survey, again, where it's in their house for as long as they choose to keep it before mailing it back, assuming they do respond.

The CHTS also attempted to get trip-level information, profiling all of the trips from all of the anglers in the household, again a much sort of heavier, cognitive burden than the mail survey, which just asks for summary counts of trips, and it doesn't attempt to profile and get detailed information on each trip. That structural design difference is pretty important, and it led to, again, a number of issues with the CHTS, in terms of incomplete data, which led to lots of imputation, and so, essentially, I think, as Rob mentioned, in the large majority of cases, anglers would -- Respondents would usually only report on one or two trips and then recognize that -- Prior to getting into the trip profiling, the CHTS respondents would be asked to provide the total number of trips taken by the household for the two-month wave in each of the modes, and then, based on that number, they would go into the detailed trip profiling.

If an angler, a respondent, had said that we took twenty trips, and then they realize, after the first trip has been profiled, that I'm going to have to do this nineteen more times, again, they would just say either all the rest of my trips are the same or just thank you for -- I don't have time for this and hang up, and that's not at all present in the CHTS. We have very little missing data, but, again, it's because we're not attempting to get that level of detail, and so I'm sort of getting a little into the weeds, but these are just sort of -- There's a lot of differences between the two designs, and, again, the model is -- Other than the wireless, it's just in there sort of as an aggregate difference between the two.

DR. REICHERT: Thank you.

MR. FOSTER: Sorry. FES.

DR. NESSLAGE: I have a question for John and a question for Kyle, and so I will give them a chance to be forewarned.   The question for John is, if I'm understanding this correctly, there is really two main effects. One is the difference in effort estimate in those three years between the mail and the phone survey, and that's a huge difference, and then the other main covariate that popped up as being important was the wireless trend over time, and those are really -- That is what is dominating the calibration, and is that an oversimplification, or is that kind of a super drilled-down essence of the take-home message?

MR. FOSTER: Yes, and that is a simplification, but I think it's essentially correct. It's just that the difference between those two is sort of being estimated by the various other terms in the model, but, yes, the three years of benchmark -- I mean, if you want to think of it very simply, a very simple version of this model, with a lot of the effects kind of -- I shouldn't say zeroed out, but assuming that we just had coefficients of one on all of these, and that would just be the ratio of the estimates in that benchmark period then applied to the rest of the series with the modification of the wireless effect going back through 2000, and so, yes, that is -- It's a little more nuanced than that, and I apologize that I haven't done a real good job describing it, but, conceptually, yes, you can think of it that way.

DR. NESSLAGE: If that's the case then, have we seen -- You would expect that, for many or most of the species, we would get the pattern we've seen, where the calibrated harvest estimates are higher across the time series, but then you would expect the deviation to increase from 2000 forward, right, and is that what you have seen across most species?

MR. FOSTER: The short answer is yes, and it's easiest to see that in the effort estimates themselves, because, yes, the difference between the two series of effort estimates is fairly consistent from 1981 through 1999, and then, once you get to 2000, when the wireless effect comes into the model, the differences between the effort estimates start to increase, and they increase -- They continue to increase until the end of the calibration time series in 2017, and that's where they are generally the largest.

For example, with private boat mode, on average, there is a difference in two to three-times between the FES effort estimates and the CHTS, but that difference is often smaller than two in the earlier part of the time series, and then it increases from 2000 to 2017 to whatever the maximum difference is, two to three-times, and so, in terms of catch estimates then, yes, the differences between the estimates do tend to be relatively stable, or similar, until about 2000, and then they do increase as well, but, again, that's heavily modified by the APAIS calibration and the specific catch data in the APAIS series.

DR. NESSLAGE: Thank you. If that's the case, then I would turn to Kyle and Erik. Going back to the revision assessments, remind me. When we were presented with those, was that -- Did you give us the MRIP revision average scale version originally or the -- I thought that was your interim approach, as opposed to the actual MRIP revision data, and is that correct?

DR. SHERTZER: They were annual adjustments, based on the ratio of the new estimates to the old estimates, that were applied to the landings that were used in the benchmark assessment, which were the Southeast Fisheries Science Center version of the MRIP estimates, and so the landings that Vivian provided for that particular assessment that were scaled by the MRIP ratios annually.

DR. NESSLAGE: Then I guess can you explain why one would want to use this -- Why we were even considering the average scale one, given how all of the trends -- How they're changing over -- The covariates are changing over time, and there's a year effects that's being incorporated, and why would we want to do an average? Why was that even considered, and I'm just --

DR. SHERTZER: It was purely exploratory. It was to tease apart this effect of scale, overall scale, versus trend at the end of the time series.

DR. SEDBERRY: Any other questions? All right. Before we move on, I need to do a little organizing here, and so I thought we could take a short break and give people who need to check-out the opportunity to check-out, and so let's meet back here at five after nine.

(Whereupon, a recess was taken.)

DR. SEDBERRY: Welcome back, everybody. What we're going to do now is bring up -- Mike is going to bring up on the screen the terms of reference, and we want to make sure -- We still have a presentation and some things to go over this afternoon, but we want to make sure that, before we leave here, that we have captured some statements and recommendations and advice and consensus statements from the SSC that address the terms of reference.

We also want to make sure that, in those statements, that the SSC can state, if it's true, that the SSC is satisfied and agrees that MRIP is the best scientific information available or that the SSC can clearly state specific reasons why they don't support MRIP as the best scientific information available, with some solutions to fix what's wrong, and so we need a statement from the SSC that says, yes, we have received all of this information, and we have digested it at this workshop, and that we agree with the agency that it's the best scientific information available.

I think we also need to address what Erik mentioned earlier this morning, some kind of a statement that the assessments that are on hold -- It would be useful to have a statement from the SSC that those assessments can proceed with the calibrated MRIP data and that the SSC endorses, approves, or whatever wording we choose, that we approve that those assessments should go forward, and so, as we go through the terms of reference, we want to make sure that we address those things and we have strong statements that the SSC agrees on.

The Term of Reference 1A, I think that also addresses the letters that we received from the states, and those are something else that needs to be addressed, and so we want to make sure that we address those concerns that are in those letters too, and so I guess the way to start this is just to bring up Term of Reference 1A, which Mike has done, and for us to insert some wording about our conclusions regarding that term of reference.

The general term of reference is review and describe the sources of disparity between CHTS and FES estimates of recreational effort for SAFMC-managed stocks, considering the impacts of the effort survey change and the Southeast Science Center post-processing. We had several presentations on this, and then the Sub-TOR 1A is describe, for a set of managed species currently in the SEDAR process, how the sources of disparity between the telephone and FES surveys affect FES catch estimate time series, with attention on trends, uncertainty, potential outliers, et cetera,

for those five species, specifically. We have, up on the screen, on the left, and I guess the screen on the right is the original term of reference is --

DR. ERRIGO: They're all the terms of reference.

DR. SEDBERRY: All the terms of reference. Then the screen on the left has the wording that we have so far from this meeting about how the SSC feels we have addressed that term of reference.

DR. REICHERT: This is a minor point, but I think it would be good for us to, if we make general statements about the FES and CHTS, to move them up, so that there is no confusion that those statements refer to the red porgy, greater amberjack, king mackerel, golden tilefish, and gag. That may help us in the future, and so just as a little aside, and then we can -- If there are any statements or issues that are specific to any of those species, we can put them below, and that's a minor thing, but that may help us down the line.

DR. SEDBERRY: Thanks. I agree that that will help us down the line. The farther we get away from the workshop, the harder it is to remember exactly what we did, and so some very specific wording is really important in these statements.

DR. SCHARF: In terms of the other reasons, clearly John has provided us with lots of information on other aspects of the design that contributed to the differences between the FES and the Coastal Household Telephone Survey. We could list those in terms of the gatekeeper effect, the cold-calling effect, the higher response rate, et cetera, or we could just include those in aggregate, and I'm not sure if all of those contributed, particularly to the pre-2000 differences, and so whether we -- Mike has the transition from the households, from landlines, to mobile only as the primary reason for those differences, and we could list those other differences as well, or we could just make a broad statement about the overall design of the survey and how that sampling framework has contributed to higher response rates, and probably more accurate response rates.

The other thing is in direct response to the letters from the states, and so the differences in the sampling domain between the surveys and their presentations about the sources of hidden effort that are captured by the FES, that are likely not captured in those state-specific surveys, that can contribute to big disparities in effort that the states were concerned about, particularly Florida and Georgia. I think those things should be pretty explicit.

DR. SEDBERRY: Yes, and I think the first thing you mentioned I think will help provide some context and some specificity to what we're talking about, and that's one thing that sometimes really gets lost in our SSC reports, is what exactly were we talking about. It's clear to us right now, but it may not be this time tomorrow, and so I think that's a good point. Then the part about the letters from the states, did you -- Mike is working on some wording there. Make sure that that captures what you were saying.

DR. DREW: I know I'm not technically part of the SSC, but I also just wanted to make the comment that I think this comment would also apply to the U.S. Fish and Wildlife Survey estimates, which were also brought up. They do use -- It appears to me, based on their description, that it's a combination phone and personal interview survey, as opposed to a mail effort survey, and so things like the gatekeeper effect and the recall bias would also be applicable to the Fish and

Wildlife Service estimates, which are much more in line with the Coastal Household Telephone Survey for saltwater angling, which was definitely something that the states brought up.

DR. SEDBERRY: Very good point, and, if we can make those kind of specific references back to those letters, the better off we are.

DR. LI: Just to follow-up to Fred's comments, the first one, the transition of households from landlines to mobile only, this is driving the major differences after the mid-2000s. Before the mid-2000s, there is no transition, and so I don't know if we want to add something at the end of the sentence to just make sure that this is the major driver for the differences after the mid-2000s. Also, if we like, we can list the other factors, the gatekeeper and other factors, that has contributed to it overall, the whole time series.

Another comment is the recommendation, the last point of the recommendation of the simulation study is testing -- We recommend a simulation study testing the process of estimation, like testing the process of estimation and the sensitivity and what is the sensitivity. To me, it's not clear.

DR. ERRIGO: It was clear to me yesterday when I wrote it, and a lot has happened since then. It was sensitivity to some of the changes or perturbations that was being discussed yesterday, but now I have lost it, and I was hoping that somebody took notes on that.

DR. LI: I have a feeling that here we are talking about how they estimate or calculate the weight stuff, and that's all the things from FES itself, and now we have concern on the calibration, and it's not FES calibration, but it's just calibration, but it's all FES information, and so I want to make sure this simulation study, what we want to test, is we test the process of sampling or process of estimation of weights and the process of calibration.

DR. ERRIGO: This had nothing to do with calibration. When you were discussing this, it had to do with coming up with FES estimates like now and forward.

DR. LI: Now forward. Okay. Now I just need some clarification about the sensitivity of what to what.

DR. ERRIGO: If I put that in error, I can just take it out.

DR. SEDBERRY: Do you think it's an error, or do you think that it needs to be more specific?

DR. NESSLAGE: When I made that suggestion, I was worried about two things. One was just error checking the code was the first question, was have you actually tested the whole estimation procedure using simulated data, so you know you're getting what you expect out of the numbers, and that's kind of the base level, and then the second thing that came up is really there are a lot of different weights and additional data sources going into the final FES estimation besides just the mail survey. There is the American Community Survey and how all of that interacts.

In theory, it should -- All the weights should play out the way you would hope they would. I guess I don't have a good handle on whether you're actually getting the expected performance that you would hope, given some of the uncertainties in those ancillary data sources, and perhaps that's something we can talk more about, to try and get at more specifically what could be done, but then,

offline, we also discussed the ability of one of these -- A simulation study like this to explore the effects of sample size and possibly guide the MRIP program moving forward in determining how to continue improving in areas where we're seeing inadequate sampling causing concerns in the ultimate estimates. I don't know, and there's a lot of things going on there, and I don't know other people feel, and it was just me that talked about it, and so I would love to hear some responses to that.

DR. REICHERT: To that point, the two points you brought up, correct me if I'm wrong, but the first point was asked and answered, and that was done, and weren't we comfortable with that? No? Maybe I am missing a recollection, but I thought that John addressed that.

MR. FOSTER: One thing I didn't mention yesterday that I will is, in terms of sort of a review of the code, while that's not been done in the context of a simulation study, the code and the methodology has been reviewed. The weighted estimation approach actually went through the MRIP certification process, and so it had independent reviews of it. The method was reviewed independently, and then we've also had sort of an independent program, or review, of the actual code itself and how sort of independently confirming the estimation using sort of kind of a -- I think a spreadsheet framework to sort of do the calculations manually, or manually with some assistance, to confirm that the 10,000 lines of code that we have to do the estimation actually return the right values, and so that piece has been done, but just not in a simulation setting.

DR. REICHERT: Thank you, and that was kind of -- Does that address your number one concern?

DR. NESSLAGE: That does address number one, yes.

DR. REICHERT: This is ultimately going to be a research recommendation, and so I think, as an SSC, we need to be very specific what we're asking. Does that make sense?

DR. NESSLAGE: That does, and so the phrase "process of estimation" is probably too vague then, is where we're going with this, and so testing the code appears to have been done, and that's awesome. That makes me feel a lot better. I guess can you can speak to whether uncertainty -- I mean, the population size estimates have uncertainty bounds around them as well, and how is that incorporated? That survey is not done on an annual basis, but you're providing annual estimates, and so I guess -- And all counties, and so I guess I'm wondering how some of those uncertainties play into the FES calculations and if that's been explored.

MR. FOSTER: Uncertainty about the American Community Survey estimates? Okay. The short answer is that the uncertainty about those estimates is not incorporated. The American Community Survey estimates are used essentially as known control totals, and while, yes, that is -- That is a shortcoming, but it's also standard practice, and I won't say that it's ideal, but it's a common practice to use the census estimates as control totals.

If it helps, the sample size for the American Community Survey, the annual sample size is, I believe, over three-million households, and they get a response rate -- Their final responding sample size is somewhere between two and two-and-a-half million households annually, and so it is a piece of uncertainty that's not incorporated, but their estimates are exceedingly precise, and so I think the CVs on their estimates are about 5 percent or smaller for most of those large state-level estimates, which is -- We're not trying to do anything with county-level estimates, and we're

working at the state level, and so those estimates are very precise. If we incorporated that uncertainty, which would be difficult to do, but not impossible, it would make a very minor difference in the variances, and it would make no difference to the point estimates.

DR. REICHERT: I just wanted to make sure that we then are -- That our recommendation is crystal clear, what we are asking, and, if that's the case, then I'm okay. That's the last sentence on the screen that's half cut-off there. If that's crystal clear, I have another clarification, because I'm afraid that, if we read this down the line, we may not exactly remember what we were talking about. The hidden fishing effort, it may be good to clarify that a little bit.

MS. SAULS: I also was wondering why you were identifying state surveys as not measuring hidden fishing effort. I don't understand what this is talking about. The same would go for the Fish and Wildlife survey. I believe all of those capture, in some way, total fishing effort and not just public access sites.

DR. SEDBERRY: I'm not sure how the wording ended up that way, but we do need to clarify that make sure it's accurate.

DR. REICHERT: I remember one of the things that we talked about was the licensed anglers. The non-licensed anglers were included, at least in part, in the FES, and some of the examples, and correct me if I'm wrong, but some of the examples that were used in the letters were based on the list of licensed anglers, and so you potentially are missing part of that effort, but correct me if I'm wrong.

MS. SAULS: I don't know which surveys you're referring to, but some of the things in the letter weren't necessarily surveys. They were just observations based on numbers of fishing licenses and taking effort and dividing it by that number of licenses, and so you're not really talking about a survey at that point, but there are other surveys that are in development and being tested in other states, on the Gulf coast of Florida as well, that do account for those things, and we didn't hear those details at this meeting. I just want to clarify what you mean by a survey.

DR. REICHERT: Well, I may have misused the word "survey". All I'm saying is that we need to be a little clearer in what we mean with that hidden fishing effort, again, because, down the line, if we are -- We want to make sure that we are very clear in our language.

MS. FRANCO: Maybe you could say, instead of state surveys, just state observations, because observation pretty much covers what some of the stuff in the letters was, but Florida has a survey that they were --

MS. SAULS: It's an MRIP-certified survey, and it's not a state survey.

MS. FRANCO: Yes.

DR. DREW: I think maybe the issue is sort of two things are being conflated here, which is, number one, that concept of hidden fishing effort, and is it -- Also, the question of is it hidden from the survey, or is it hidden from other anglers or from personal anecdotal observation versus something like comparing legitimate surveys like Florida's MRIP-certified survey and the U.S. Fish and Wildlife survey, and you can't just directly compare those estimates, because of

differences in methodology and sampling frame that we've talked about here may have -- They are not perfectly comparable.

If we're doing like a comparison of the estimates that come out of that, we should also be considering some of the effects that we talked about here that are maybe weaknesses of a phone-based survey versus a mail-based survey, differences in how the estimates are expanded, that we haven't really had a chance to talk about in I think some of these alternative surveys, and so maybe the recommendation should be that -- Not a recommendation, but that, when you compare these surveys, they are not direct same methodology, and there's differences that may be contributing to differences in the estimates, and so, just because one says one thing and the other says another, it doesn't mean that one is wrong and one is right or that there is differences that are contributing to why you're seeing differences in effort that we kind of already touched on here with the FES.

DR. REICHERT: Thank you for stating that a lot more eloquently than I did.

DR. SEDBERRY: Now we need to make sure that we capture it in writing. It really needs to make it into the report.

DR. LANEY: Along the lines of Beverly's comment about surveys, I think, Katie, just for clarification purposes, the U.S. Fish and Wildlife Service survey you were talking about is the every five-year survey of hunting, fishing, and outdoor recreational activities, or I think we should put in there whatever that official title is, but I think that's the one that you're referring to, right?

DR. DREW: Yes, that was the one that I was referring to, and I believe that's the one that the Florida letter referred to.

DR. SEDBERRY: Thanks for that clarification. It's really important to get the wording right in these.

DR. LI: For the statement of the recommendation, do we want to remove the process of estimation, or we can say like testing the sensitivity of certain factors to the process of estimation, for example the uncertainty in population size estimates or something.

DR. SEDBERRY: Mike is removing it, or he's doing something with it.

DR. NESSLAGE: I was waiting my turn, because we switched subjects. Can we take one subject and deal with it first and then switch back? Can you not delete it until I get a chance to say something?

DR. SEDBERRY: Okay. How does that look to everybody?

DR. BUCKEL: Just so when we come back we remember -- The sentences you were just working on look good to me, but if you go up to the one that was dealing directly with the letters, maybe preface that sentence, that paragraph, with -- It starts with the differences in sampling frame, and preface that with respect to the letters, the state letters, sent to the council with the differences, so then that's clear that that's dealing with the letters. Okay. Then, but not incorporate into the state observations, and then maybe a parenthetic statement, and those were largely based on the

saltwater license numbers, and so maybe -- Some weren't, but maybe a parenthetic statement after that of which were largely based on saltwater license numbers.

MS. FRANCO: Well, it's not -- All the state letters were different, and I don't know if it's fair to generalize that statement.

DR. BUCKEL: I am just trying to -- If there's another way we can deal with it, but that's one of the big differences, to me, between the numbers, is because of the effort from unlicensed anglers.

MS. FRANCO: Also from the private sites, because that's also kind of hidden as well from everything.

DR. REICHERT: I was just thinking the same thing, that it's not just the unlicensed anglers. There are other hidden efforts that we discussed yesterday, and I think you just mentioned one of them, and it would be good to think about those sources, again to be specific, in terms of what those hidden fishing efforts were, or where they may originate from.

DR. DUMAS: Another example of that was the non-coastal areas of the state were included in the FES, but not in the CHTS, as I recall.

DR. ERRIGO: Just to clarify, private fishing sites are hidden fishing effort, but there is no survey. The FES does not get at that. It's the FES does. It's the APAIS that doesn't. Sorry,

DR. SEDBERRY: Okay, and so we want to make sure that our wording is clear and there's enough context provided so that we'll remember what we were talking about and that the state representatives are happy with the way that the content of the letters has been presented here and addressed, and so are we all good with what we have so far? Okay.

DR. ERRIGO: Genny, did you have something else to help hone this recommendation?

DR. SEDBERRY: Okay. Luiz, if you are there, we're trying to connect with you.

DR. BARBIERI: Thank you so much, George and Mike. I was just going to ask -- Right now, for those of us who are on the webinar, we cannot see what's being written, and I don't know if there's a way to project on the webinar the actual consensus statement that's being put together now.

DR. SEDBERRY: Luiz, we're not using the computer that's connected to the webinar for the statements, but we can switch that, we think.

DR. BARBIERI: That's fine if that's going to be too difficult, and I can review that later, and that's fine, but I just wonder -- To tell the truth, I just looked at the agenda, and I just looked at the terms of reference, and I am not seeing where here, explicitly, it was ever part of our charge to respond to the letters that were written by the states, and I am bringing this up simply because I feel that those letters were written based on a number of observations and impressions, and I think that they express concerns, but I don't know whether we will be able to properly address those concerns or understand the perspectives that folks who wrote those letters actually had when they

wrote the letters, and I am trying to see how the SSC can effectively achieve this in a way that is clear and objective at this time, at least at this time.

I mean, looking at all the terms of reference, everything that we have on the table that we have to achieve to complete our terms of reference and meet them, and I think it was Dawn, or maybe it was Yan, but somebody mentioned that the different letters have different content, and they reference different things.

The State of Florida, for example, does not have a survey in the South Atlantic, on the southeast coast, at all, and so we discussed some of the issues that were brought up by the perspective that will be gained by looking at the results of the Gulf Reef Fish Survey, because we have been working for six-plus years with the MRIP office and their consultants in developing procedures that we felt were really sound, scientifically sound and statistically sound, that were reviewed by the MRIP consultants and others and has been certified as an MRIP survey, and so those questions are not to say state survey or -- Which, by the way, it's not really a state survey. It's a supplemental MRIP survey that was developed with the MRIP program for that purpose, specifically to serve as a supplement and target a stratum of reef fishes that was not as appropriately sampled by the general survey.

I am bringing this up because there was this discussion about state surveys or other surveys, and, I mean, the discussion is just trying to find out -- I think this workshop is trying to find out whether the data that we have coming in right now can be used to develop stock assessments, stock status determination, and catch advice, and we developed specific terms of reference for that purpose, and so I'm not trying to sound grumpy, and I'm sorry if I'm sounding that way, but I'm just seeing that we're kind of going a little bit into a rabbit hole of trying to address general letters that were written to either the MRIP program or the council and that we were never explicitly asked to respond to.

DR. CROSSON: Luiz, do you think we should remove that section in our write-up about the respect to the state letters, just remove it entirely? What is it that you're proposing?

DR. BARBIERI: Yes I do think that we should remove it, or, if we want to do it later, if we have time -- I am not opposed to trying to do it, but I'm just looking at our terms of reference, everything that we have on the table to address, and the discussions that we are having, and different people are saying, okay, what state surveys, and how do we respond to this, the different letters, and it becomes -- Scott, to me, it's like how difficult it is to define what's the message in the letters and what are we trying to really address and are we prepared to respond to them explicitly, as we are stating right now, or do we prioritize following our terms of reference and actually addressing those first?

DR. SEDBERRY: Luiz, I think you've made some excellent points there.

DR. ERRIGO: Luiz, thanks for that. First of all, I got the notes up on the screen for you there. Also, what we can do is most of the information written in that sentence has to do with differences between the FES and actually the CHTS, as well as any other types of surveys that are similar, and so that information can stay in, and we can re-work that sentence or put it up in a bullet up above as sources of disparity, because we are tasked with coming up with sources of disparity between the two surveys.

We don't necessarily have to address the letters specifically, but, if we put information in there about all the sources of disparity, then others can take that and come to whatever conclusions they would like. If we're asked specifically to address the letters, we can then come back and specifically address them, if that's how we want to proceed.

DR. SEDBERRY: The terms of reference do not specifically ask us to address the letters, other than, as you said, in light of the disparity between CHTS and FES. There is information in there that contributes to the -- That describes some of the disparity, but you're right that there's nothing in our terms of reference that says that we need to address those letters, and I may have sent us down that rabbit hole, in terms of -- But, really, I'm thinking of it in terms of just addressing the disparity and is there anything in those letters that describes the disparity.

DR. SCHARF: I understand Luiz's concerns, and we can reword this a little bit, but I do think that what we've heard here today contributes to our understanding of why the FES-calibrated estimates are so much higher than what are being observed based on what was contained in those letters, which were part of our meeting the last time that we met, and so I think that we can address them at least generically. I mean, we don't have to -- We are not addressing those letters explicitly, point-by-point, and we're just saying that some of what we've learned at this workshop has informed why there might be some large disparities, and I think that we need to have that language in our report.

DR. SEDBERRY: I would be comfortable with that language as well, but we will hear from other SSC members.

MS. LANGE: I think all we need to do is remove the first phrase of "with respect to the letters sent to the council" and just start off with the "differences in sampling frame and hidden effort are a part of the issue", and I also would suggest that moving the consensus and recommendation to the end of the discussion would be appropriate.

DR. SEDBERRY: Thanks, Anne.

DR. BELCHER: Some of it, to me, as Fred was saying too, as someone who was part of the state people who helped write the letter, there was a lot of clarification. Again, we're looking at it from thinking that the saltwater license registry was what the base was of how this was being designed, and finding out about the households and the fact that you had license matched and unlicensed matched has a very different implication for us. I mean, I understand it better, from that standpoint, and, again, this 40 to 80 percent unidentified effort, as Dawn recommended, finding out exactly what we can bring back to the states to know of things that aren't technically licensed. I mean, to me, that's amazing to see that half of that is unknown licensed people, whether it's private docks or whatever, and still that's a pretty hefty statistic to take home.

DR. SCHUELLER: All due respect to the terms of reference, but they're just a minimum amount of work or statements that this group needs to make. There is no statements that say we can't go above and beyond what the terms of reference have asked for, and so us addressing these letters is completely within the scope of this workshop, given the discussions we've had historically, and so I'm against taking this out. I think it's fine to include it. Just because somebody doesn't ask for it, it doesn't mean that we can't give it to them.

DR. SEDBERRY: Good point.

DR. SCHARF: To that point, we wrote the terms of reference.

DR. SEDBERRY: So the steering committee giveth and the steering committee taketh away. I don't know what wording you may have changed during all of that discussion, Mike, but --

DR. ERRIGO: I took out the beginning of the sentence that said with respect to the --

DR. SEDBERRY: That was Anne's suggestion, I think.

DR. ERRIGO: Yes.

DR. SEDBERRY: Okay.

MS. LANGE: To Amy's point, I don't think that we necessarily have to include "according to the state letters". Everything that's in here is relative to what the states were concerned with. I mean, everything we're working on is all related to their letters in one way or another, and so I don't think we need to explicitly say -- I think leaving it out is a good idea, at least that phrase.

DR. SCHUELLER: That's fine with me, but the initial recommendation on the table was to delete the entire section, and I'm not okay with that.

DR. SEDBERRY: I think the general consensus of the committee is that we're not going to delete that, but, the way it's written now, probably with some tweaking here and there, is the consensus of the committee. Does anybody object to anything about the way it's written right now?

MS. SAULS: Well, I mean, I don't know. Maybe I will let Luiz speak for Florida, but my presence was requested at this meeting, and I was invited to participate, and I was prepared to provide a presentation, but I was not asked to do so, and so I don't really feel like we were able to give you enough information for you to say that everything in those letters is -- I am just not sure that you adequately considered some of the concerns in those letters.

DR. SEDBERRY: You may be right about that. We didn't have a formal presentation on it, but they were included in the briefing materials.

DR. BARBIERI: If I may jump in for a second.

DR. SEDBERRY: Go ahead, Luiz.

DR. BARBIERI: Well, I guess I was misunderstood, my statement was misunderstood. Our first term of reference says to review and describe the sources of disparity between CHTS and FES estimates of recreational effort for South Atlantic-managed stock and consider the impacts of the effort survey change and the post-processing from the Science Center on -- Obviously, the topic of the disparity is relevant.

I just felt that, if we were going to the point of starting explicitly responding to the letter, I didn't feel like -- That's just my own personal opinion, but I didn't see that as something that we could do or that was part of our charge. If others disagree, and I heard Amy's comment, and, if the committee decides that responding directly to the letters is a good idea, by all means, but I was just trying to stick with the charge here, in terms of what the SSC role is, as a scientific reviewer, and the charge that was put in front of us, but I will respect the decision of the committee to proceed otherwise.

DR. SCHUELLER: It's regretful that the states weren't able to present something based on their letters, and so that's -- We can't fix that at this point, and so I guess my question to the state folks is, given that this is sort of directed at the FES surveys versus some of the other options out there, are there specific words in here or statements in here that you disagree with based on your knowledge of what's going on in your state, and do you have a recommended change to what's up here, because I know this is an SSC meeting, but I don't necessarily complete view it that way. I think that the state participants should speak up, and I think that they should put their opinion into this, because they're invited participants to the meeting, and so, if you see something up here that misrepresents what your state is doing, please tell us.

DR. SEDBERRY: Yes, absolutely.

MR. ESTES: I'm Jim Estes, and I'm with Florida Fish and Wildlife. I think what you are trying to get at, and I don't necessarily agree with it, but let me see if I can help. I think what you were trying to get at is that you believe that the new FES represents reality, is an accurate representation, reasonably accurate representation, of reality. I think that gets to what you want to go do, and, therefore, it seems to me like that would follow that we need to move on with the assessments, that we need to accept this and we need to get on, and I think that's what you're saying, but you're getting all jumbled up in, I think, the words and trying to be really careful, and that's just one opinion.

MS. SAULS: I don't think there's a state in this room that would tell you that you guys should not do any more stock assessments and not use any data that you have available to continue managing our fisheries. I mean, we understand that these are the data you have to work with, but we did want -- There were some concerns that were brought up that maybe there is still some more work to be done, and I think that people would appreciate having that acknowledged.

DR. SEDBERRY: Well, eventually, we're going to have, as part of this series of statements we're making, we're going to make some research recommendations as well, and perhaps what you're talking about would fit in there, but, if you have a -- If any of the states, or anyone at the table, has a statement they would like to make to be inserted into the record, we can do that.

DR. ERRIGO: Also, when the charge was given for this workshop, it was -- The explicit intent was to come up with a way to move forward now, knowing that there is a lot that is being worked on that may change or develop as time goes on. There are long-term things that are being looked at, and so that is explicit in this workshop, that we are looking at the short-term.

DR. SEDBERRY: That's correct.

MS. LANGE: I agree, and I just -- Again, I think the consensus and recommendations, we have pretty well -- Not the recommendations, because we still need to flesh out what's here, but I think they just need to be moved to after our discussion. Everything else there is sort of answering the question, and then the consensus should be at the end, and the recommendations should follow.

DR. ERRIGO: That's just an artifact of how the conversation took place. That consensus and recommendation was yesterday, and what's under it happened today. That's why it's like that, and I will just move it.

DR. SEDBERRY: Yes, we can move that around, and we will, and I see that, at some point, we'll have kind of a summary of consensus statements and recommendations at the end of the report, and so, right now, these are -- We're trying to make sure that we get everything down accurately and that whatever statements we make have context, so that we'll know what we're talking about two weeks from now and that it all makes sense and that there's consensus and that we agree with that's being stated.

DR. DUMAS: One point related to the differences between the state and Fish and Wildlife estimates and the new MRIP estimates that was brought up yesterday was that the definition of a fishing trip could be different between the different surveys, and that could account for some of the differences between the estimates, and so an example was given of the definition of a fishing trip in the Georgia survey is different from the one used in FES, and so the definition of fishing trip could explain some of the differences.

MS. FRANCO: Can we just interject really quick? Georgia isn't a survey. It's just observations on fishing licenses, and I just wanted to clarify that for the record.

DR. DUMAS: Thanks.

DR. REICHERT: To Jim's point earlier, I agree, although I wouldn't call -- I wouldn't conclude that we think that that represents reality. I mean, it's the best estimate we currently have, and that doesn't mean that there aren't means to improve the methods or estimates or whatnot, and so that's the first point.

The other thing is I think it would behoove the committee, in terms of our consensus, and that goes to the point that Erik was making earlier and the additional point of this is not only used -- The new MRIP estimates are not only used for stock assessments, but also to track ACLs, and so I think, unless you disagree, I think it would behoove us to make a statement that we feel that the new estimates should be used for stock assessments and tracking ACLs, until other information becomes available or better estimates are available or something like that. That's a little more -- That goes a little further than we just consider it the best scientific information available, because I think, ultimately, that is probably something -- That is a statement that the council is probably looking for from us, but that's if there is a consensus in the committee that that is something that we are comfortable with, and I just wanted to bring that up.

DR. SEDBERRY: I agree, Marcel, that those kinds of specific statements are very useful for the council, and that's what they're looking for, and so, if we can put that up there and make sure that we all agree with the wording.

MS. MATTER: I just wanted to add that the FES estimates wouldn't be used to track ACLs until an ACL was set using the FES estimates in an assessment, and so it would be quite some time until that would happen, but, yes, it would follow then that the FES would be used to track an ACL that was set using FES estimates. Right now, we're using CHTS-like estimates to track ACLS that were set using the older data.

DR. REICHERT: Thank you for that clarification.

DR. SHERTZER: Just to second that, because this list of five species is an incomplete list of the stock assessments that are ongoing. We also have scamp and cobia that are ongoing and ones that are coming up soon, such as Spanish mackerel and some others, snowy grouper.

DR. SEDBERRY: Marcel, are you looking at the --

DR. REICHERT: I think that's good, to make sure that we add that language to the report, and so, again, that it's clear if, how, and when this is going to be used, and so to avoid confusion.

DR. SEDBERRY: Okay, and so Mike is finished typing the consensus statements from Marcel and Vivian and Kyle's comments, and are we agreed? Does the SSC agree that that is our consensus? Any objection?

MS. LANGE: I am not sure that that second consensus statement reflects what we're talking about. We endorse using the FES estimates to track ACLs set using the -- I am not sure what the "until improved datasets become available" -- I mean, if we're going to use these once the new assessment comes out, based on the FES, until then, from what Vivian said, we'll be using data from the old system. I mean, I guess it's just that "until improved datasets become available", and I'm not sure what that means.

DR. REICHERT: Can I make a suggestion perhaps that will clarify that? Rather than "until improved datasets", say "once the FES data are --" Vivian, can you help me out with the language? Maybe "are used in an assessment", or the language that you were using earlier of FES will not be used to track landings until FES estimates are used in a stock assessment.

MS. MATTER: The first sentence sounds fine, to track ACLs set -- Track ACLs, and you can say that were set using the FES estimates, and you can say existing ACLs that were set using CHTS would continue to be monitored using CHTS-like estimates, until they are re-assessed.

DR. REICHERT: Erik made a good point. Just call them ACLs.

DR. SHERTZER: I thought you said something about using them in assessments as well, which didn't get captured in this statement.

DR. ERRIGO: That is another consensus statement, but I just wanted to get this one out of the way, but this is just about tracking using the FES, and then there's the assessment one.

DR. DUMAS: I just wanted to remind us that, yesterday, we learned that the standard errors around the estimates that come from the new FES/APAIS methodology were noticeably larger than the CHTS estimates, and so, with those larger standard errors, if the confidence limits around

the new estimates include the old estimates, then the new estimates are not significantly different, and so that should be remembered. I think it might be important to see what those confidence intervals are around the new estimates and see if they are so large that they include the old estimates. That might be important going forward. Thank you.

DR. LI: I just want to make sure that here, when we say track ACLs, it means in the future, right? Okay.

MS. MATTER: I was going to mention that, too. It could say "that will be set using FES-based data".

DR. SEDBERRY: Any additional comments or clarifications or edits? Is there anybody on the webinar that has any additional comments or questions or edits?

DR. WILLIAMS: I try to avoid these things, but we still don't have the stock assessment consensus statement, and the last recommendation that was just put up is pointless. I don't understand the point of it. I mean, we're beyond CHTS, and what is that comparison going to tell us? Nothing.

DR. SEDBERRY: That was Chris's comment, and I thought it was just a comment, and I'm not sure that it was a recommendation, was it?

DR. DUMAS: We're concerned about whether the estimates using the new MRIP method are -- The reason why we're here is because they were very different from the estimates using the old method, and, if the estimates from the new method have wide confidence intervals that include the estimates from the old method, then the new estimates aren't really different from the estimates from the old method, right?

DR. SEDBERRY: Is there a recommendation?

DR. DUMAS: The recommendation is to calculate the confidence intervals from the estimates of the new method and compare them with the estimates of the old method and see if the estimates from the old method fall within the confidence intervals of the new method, and, if they do, then there is really no difference in the estimates between the new and the old method, and so, for the purpose of setting fishery management policy --

DR. SEDBERRY: But we're still going to use the new method.

DR. DUMAS: But, if the new estimates are not significantly different from the old estimates, then there is no justification for changing policy.

DR. GRIMES: Haven't we already recommended using the new method? Why justify using the old method if we agree that the new method is --

DR. SEDBERRY: We have already done that, at our previous two SSC meetings.

DR. ERRIGO: I think what Chris is saying is, if there is no significant difference between the two, you can use the FES estimates to track ACLs set using CHTS data. However, there has been

shown to be a systematic shift in the estimates up. It may not be significant statistically, but it's there, and, when we track using point estimates, without the uncertainty bounds, we will hit our ACLs quicker, and that's the problem.

DR. DREW: I think this is just -- Obviously, I would defer to John Foster and the MRIP crew on this, but ASMFC has dug into these estimates a lot for the species that we're interested in, and, while the confidence intervals are -- The PSEs are larger for the calibrated CHTS estimates than they are for the original base estimates that we've been using this entire time, especially in the early part of the time series, the CVs, or the PSEs, on the most recent sets of data are really comparable to what has been done with the improved APAIS estimates from that 2004 forward period, and so I don't think there is -- The differences that we're seeing in total catch and removals are outside of those bounds, for our more data-rich species, and so, obviously, there may be exceptions, especially in the South Atlantic, where you have more data-poor species and more rare-event species, but I would say that I don't really -- It's certainly something that your assessment teams can look at, if they have any curiosity about that, but I don't think that it would show what you're expecting to show, which is that the FES estimates, especially in the most recent timeframe, are significantly similar to the CHTS estimates. I think the PSEs are comparable to what they were before, in the most recent set of time series, and the estimates are large enough that they fall outside of those bounds, and so there is a significant difference.

DR. DUMAS: Thank you. That clarifies it, because I didn't know whether there was a significant difference, and so good, and so that analysis has been done. Thank you.

DR. DREW: To clarify, I'm talking about my experience with ASMFC's managed species, which, again, there may be specific species within the South Atlantic that are different, but, from everything that we've seen so far, the difference persists even with the new calibrated PSEs.

DR. SEDBERRY: So do we still need that recommendation, Chris?

DR. DUMAS: I would recommend that we do what you just described that you did for your species.

MS. LANGE: I was going to suggest removing that recommendation for this part, but I'm not sure, and I also was curious about, and I think Genny was as well, about the initial recommendation, the wording of that, for the prioritize simulation study, testing the sensitivity of the process, and has that been clarified?

DR. SEDBERRY: It has not.

DR. NESSLAGE: I'm sorry to tell you how to run the meeting, but have we decided to strike Chris's recommendation? We should probably end that discussion before we start another big one.

DR. SEDBERRY: I am not sure that it adds a lot. I mean, I'm not sure that it's something that we need to address.

DR. DUMAS: If it's not a consensus recommendation, I think then strike it.

DR. SEDBERRY: Okay.

DR. LANEY: I was going to weigh-in on that first. Given that Katie pointed out that the analysis they did was for ASMFC-managed species, and further qualified that there could be SAFMC-managed species, especially rare-event ones or ones that are data-poor, it still seems, to me, from a purely academic perspective anyway, that the analysis that Chris has suggested would be beneficial for the SSC, at least in seeing where those point estimates lie relative to the new estimates. Again, I don't know how much work that would entail.

DR. SEDBERRY: So you think it would be beneficial to the SSC?

DR. LANEY: I would personally find it beneficial, but I will, obviously, as the new kid on the block, defer to the rest of the SSC for their input on that issue.

DR. ERRIGO: I have done a lot of work with both of the estimates for a while now, and I have looked at it a lot of different ways, and, from my experience, what I think we'll see is that the species that we assess, for the most part, that have a larger recreational component, not like golden tilefish, but the other species, you will see that there is a significant difference between the two and that it gets larger in the later part of the time series.

For species like golden tilefish or knobbed porgy or something like that, you will find there is probably not a significant difference in most years, because, in most years, we have almost no data. There is nothing there, and so there may be one year where there's like a huge spike, where we encountered like three trips that had knobbed porgy, and there may be a significant difference in that year, because there is data, but that's probably what we'll see. The species that are very rare, that we don't have any data for, there won't be much of a difference. The species that we assess, for the most part, that have a recreational component, there will be.

DR. SEDBERRY: Genny, did you --

DR. NESSLAGE: Are we back to the simulation thing again?

DR. SEDBERRY: I think so.

DR. NESSLAGE: It's time. Okay. John, I have a question. Help me understand, because I'm not a sampling guru. Your estimates should be unbiased if the strata are appropriately describing the way the system is working, correct?

MR. FOSTER: The estimates are unbiased with respect to the estimation approach. The sample weights are calculated appropriately, given what the frames cover, and they are unbiased with respect to how sampling error is estimated. Now, that's not to say that there aren't things that we really can't account for that could be creating differences between the estimates and whatever the unknown true values are, like the fact that APAIS can't cover private access sites, for example, or any of the sort of potential non-sampling errors that can be present in any survey.

From the methodological standpoint, yes, they are unbiased, but there could be things that we can't account for that could be creating biases, and, again, it's just -- That's present in any kind of a

survey that is dealing with people providing responses and not just sort of a theoretical survey, essentially, or a simulated survey, and so did I address it? I'm not sure if I answered the question.

DR. NESSLAGE: I think so, and then the follow-up question is I think Rob had mentioned that the sample sizes were selected based on the variance observed in the previous survey estimates. Is that correct?

MR. FOSTER: Yes. Specifically for the FES, that's right. It's based on what is called a Neyman allocation of standard approach that uses historic variance information to help optimize -- For a fixed sample size, it helps optimize the allocation among the strata, using the variance information, yes.

DR. NESSLAGE: That variance information is coming though from the Coastal Household.

MR. FOSTER: No, the variance information is coming from the three previous years, and so, for the 2018 FES allocations, those were based on the results from the FES from 2015, 2016, and 2017, and so we've been conducting the FES since 2015, but, for those first three years, it was conducted side-by-side with the CHTS, as part of the benchmark period, and so, again, we have that time series of FES estimates and their variances that we use in this process, and, moving forward, we'll continue to use a moving window of the most recent years to continue to inform the allocations for upcoming years.

DR. NESSLAGE: Those variances are calculated at the various strata or clusters, correct, in order to determine the appropriate sample size for each sub-section of the survey, correct?

MR. FOSTER: That's right, and so, for this allocation purpose, we are calculating variances for estimates within the strata, yes, and that's a more detailed level than how we calculate it for standard estimation, but, again, it's just for the purposes of sample allocation.

DR. NESSLAGE: Then I guess my question is, if the strata are not necessarily capturing the, and I have got to make sure that I say this right, but the structure of the process that you're trying to estimate, could that be a source of potentially biased estimates from the FES? I know the states have mentioned some concerns and some places where the strata may not be -- It may not be adequately characterizing their population that they want to be included in the sampling frame, and so I guess what I'm wondering is, if we're talking about a simulation study, is it possible, and/or worthwhile, to even explore additional or alternative strata?

MR. FOSTER: Well, if the estimation approach, and, in this case, essentially, it's, if the sample weights are calculated using the sample selection process in the design, you will get unbiased estimates for the population that's being covered by the survey frame, and that is how FES estimation works.

If we were, for example, making simplifying assumptions, like, if instead of accounting for the design stratification, we just assumed that, well, we have just a simple, random sample from the entire state, and we ignored all of the strata, and we calculated sample weights as if we had just taken a simple random sample from the whole state, that would be inappropriate, and that could lead to, potentially to, biases, if there was some correlation or association with fishing effort and reporting fishing effort and the strata that we were ignoring. Yes, that could create a bias, but we

don't do it that way. The sample weights are calculated correctly, by reflecting all of the different levels of stratification that are present in how the sample is selected.

Again, that is just -- The process by which the samples are selected, the process by which the sample weights are calculated, is all just -- It follows standard sampling theory for the FES's stratified design, and so, again, it's simple random sampling within the strata, and that sample selection process is completely accounted for in how the sample weights are calculated, and then we use those sample weights to calculate the effort estimates, and so, again, it's just -- It's the standard approach to sampling, design-based sampling, and so I don't -- Again, if we were ignoring aspects of the design in how we did estimation, then there could be a problem, and that's, essentially, what the problem was with the old MRFSS approach, was it ignored complex aspects of the design in its estimation approach, but we're not doing that now. We fully account for all of the design aspects, including stratification, in the estimation approach for both the FES and the APAIS.

DR. NESSLAGE: I believe you, but I'm not sure that's the question that I am asking. I guess someone help me out here if they -- I see Katie thinking, and I'm just not asking the question appropriately. I guess my question is what if the strata aren't the best we could be using?

MR. FOSTER: Okay, and so I think I have the question now. If your strata -- If we are covering -- Changing the strata could lead to systematic differences in the estimates if changing the strata changes the coverage, and so, if there is suddenly people we are or are not covering when we change the strata, then, yes, that could lead to systematic differences in the estimates, but, if you change the strata, but you haven't changed the overall coverage, and so you've got the same sized pie, but you have just changed how the slices are cut, that's not going to lead to a bias, or a systematic difference.

It will make your estimates, potentially, more or less precise, and you can improve the precision on the estimates by changing the strata, and so, ideally, with stratification, you are partitioning your frame into groups that are homogeneous within and heterogeneous among, or between, and that's the goal, one goal, for stratification, and, to the extent we can do that, it will improve the precision on the estimates, but, if you don't use any stratification, and you take a simple random sample, and you calculate appropriate sample weights, that's going to be an unbiased estimate.

If you then stratify that sample and you, again, calculate appropriate sample weights, given that you now have stratification, you get an unbiased estimate as well, and, if you happened to, using those two approaches, get the exact same sample units in two different draws, you would get the exact same point estimates, and so the differences would come in if, again, you have changed the actual coverage of the population in some way, but, as long as that coverage is maintained, then changing your strata won't change or introduce systematic differences in the estimates, and so, again, changing the strata, if the coverage remains the same, can only result in changes to the precision on the estimates. It can make them more or less precise for the same amount of sample.

DR. LI: What John has said is completely right on the condition that the population is evenly distributed or a homogeneous distribution. If they are clustered, then, depending on how you design the strata, it could be different.

MR. FOSTER: Well, I'm not sure if I know what exactly you mean, but, if your frame covers all of the same units -- If you have different strata, you can change the strata, but you shouldn't be introducing a systematic difference. Now, I'm not sure exactly what you mean by how the units are clustered. I mean, with the FES, we're sampling households, and so the stratification -- The strata have to be true strata. They need to be exclusive, and so a household shouldn't be in multiple strata, right? I mean, it's only going to be in one stratum. If it ends up in multiple strata, then we could have an issue with sort of duplicity. A household could be overrepresented, and we have to figure out a way to account for that, but the way the design works is we don't have households that can show up in multiple strata.

DR. LI: I think my comment was like, for example, if you are sampling a pizza, if you are sampling the number of mushrooms on that pizza, if those mushrooms are evenly distributed around that pizza, and then it doesn't matter the strata and how you cut, and you will estimate right. However, if the mushrooms on the pizza are clustered in this corner more and in that corner a little bit, then, depending on if your strata when you design -- If you can capture the structure of that corner or whatever, because we can never sample the whole pizza, and that's the thing. You cannot have coverage of the whole pizza, and so we cannot sample the whole true population, never.

MR. FOSTER: Well, right. We're not going -- We can't census the population, but it's not -- In that example, let's say all the mushrooms are clustered on one slice of the pizza, in one stratum. It's hard to keep the analogy going, but, in the context of repeated sampling, you will still get an unbiased estimate, so long as you have sample in all of the strata.

Now, the tight clustering kind of gets to kind of a rare-event sort of scenario, where there may be -- In the context of repeated sampling, there might be a lot of samples that don't get any observations with mushrooms, but you will get some mushrooms, such that the expectation will still be unbiased in the context of repeated sampling, but, here, while fishing is still a fairly rare event, and a small fraction of the overall population is fishing, reports fishing, it's not so rare that we never get any fishing reports.

I mean, we are always getting -- We get reports of fishing. In fact, again, as I mentioned briefly, that's one of the sort of calculations we do when we look at sample allocation, in addition to just the Neyman allocation, is what's the probability of us detecting at least one, or having at least one, household report fishing in our strata, and so we, again, allocate samples so that we have reasonably large probabilities of detecting at least one household -- Having at least one household report fishing, given that fishing has occurred.

By reasonable, I mean like 85 or 90 percent chance of getting a household reporting fishing, given that there is fishing going on by households in the strata, and so the bottom line is that I agree with you that, depending on how the strata are set up, you may be more or less likely to detect some event, particularly if it's a rare event, or highly clustered, as you're saying, but, in the context of repeated sampling, it's still not a bias. It's just -- It is still a precision issue, even though, in any one realized sample, the samples -- It may perform better or worse, and I'm sort of rambling on, and I apologize, but it's just -- We have to keep in mind that the sampling, the unbiased-ness refers to this context of repeated sampling, and so, if we were able to take many, many repeated samples from the same population, the mean of that distribution of efforts that we could produce from all of those sample will be giving us the right answer. It will like up with the true population value.

DR. SEDBERRY: We are still on Term of Reference 1A, and I want to make sure that everybody is happy with the wording that we have for recommendations and consensus statements, but I also want to make sure that we get in a consensus statement about proceeding with the stock assessments, and I'm not sure that we have actually --

DR. ERRIGO: We have not gotten to it.

DR. SEDBERRY: We have not gotten to that one yet, and so let's be thinking along those lines.

DR. NESSLAGE: To not belabor the point, I think, personally, and, again, no one really else has commented on this, and so I would appreciate that this is more than just me talking about this, and, if I'm the only one, I will just withdraw the recommendation, but I think there is value in demonstrating to the world that you're right, that these should produce unbiased estimates, and one way to do that is a simulation study with a known set of answers that you can reproduce.

Then, as we also discussed, there is value in exploring the precision of those estimates and how that would be affected by your sample sizes, and that can be done through simulation as well, and so I still think there is value in doing that simulation study. I don't know where that falls in NOAA's priorities, but I stand by that statement, but I would love to hear from the rest of the SSC.

DR. WILLIAMS: I am not on the SSC, but you will hear from me. I appreciate John's wonderful tour through remedial stratified sampling design, but that's what it is, and there is no reason to test it. It's in the literature, and I suggest a book by Thompson, and it's very good, on sampling design, and there is no point in simulating it. It is what it is. It's exactly as John just described it, that the strata can vary, but you're still going to get an unbiased estimate, and it's all a matter of efficiency and minimizing variance.

DR. SEDBERRY: Okay. So the recommendation about simulation should be left in there, and, when I go to the council next month and say that one of our recommendations is that OST prioritize a simulation study, if I use those exact words right there, someone is going to say, well, what do you want them to do, what does this mean, and I'm not sure that I know.

DR. LANEY: Well, you had asked if the rest of the SSC was comfortable with leaving that statement in, and I am, for what it's worth.

DR. SEDBERRY: Is there anyone on the SSC that disagrees with that, that would like to see it removed?

DR. BELCHER: I would just -- I guess some of this is, when we ask for these things, okay, and so we ask for it, and does it get done, or does it not get done, and, if it does get done, how does that help us move forward? Is this something that we need? What if it comes back and the thing blows up and it doesn't look good? Does that mean that we do something different?

I kind of hesitate, only because, again, we can ask for it, but, if it's not going to be addressed, or John has indicated that some of that has been done, and so is that something that, if it's not been written up, or if it has been written up, maybe it's already readily available, and I just don't know what we do with it, and that's the only thing, is that I don't disagree with the idea for the

understanding of it, but, from our standpoint right now, I don't know how that is going to impact us in moving forward.

DR. SEDBERRY: Right, and I can see that, when I present this to the council next month, they're going to say, well, what are you going to do with this, and I don't have an answer for that.

DR. LANEY: A question for Genny and for John. I mean, John, you have indicated to us that this model was constructed using standard techniques, and I would ask Genny -- My perception, again, not being a modeler at all, is that this sort of simulation is also standard operating procedure, isn't it, most of the time, when you build a model?

MS. LANGE: Well, I have the same question I think that you had, George. What exactly is expected? When you go to the council and say that this is our recommendation, what is it exactly that we're recommending, and what will be the use of the outcome from it? I am not sure, and my expectation and understanding is that S&T has gone through all of the rigorous things that need to be done to validate their design, and, if that's the case, I'm not sure that this is going to add, and it may take away from other things that can be done, and I'm not sure.

DR. NESSLAGE: If everyone was convinced that these numbers that are being generated were unbiased, we wouldn't be here. If the states were happy with the precision of the estimates, which they never have been happy with, and I assume there will still continue to be concerns with the precision of the estimates, we wouldn't keep coming back and questioning MRIP or MRFSS. I am just suggesting this as a way to help the council, to help stakeholders, to help the states move forward in feeling better about these estimates.

I don't need them. I don't know at this point that the SSC needs them, but I am suggesting it as a way to move everybody else forward to get onboard, and so, if the SSC doesn't want to suggest that, and you think it's wasting the MRIP staff's time, that's fine, and I happy to drop it, but I'm just saying this might help the situation, which goes way above and beyond our technical comment on the technical aspects of these estimates.

DR. SEDBERRY: Thank you for that. That helps me. Now, when I posed the question about leaving this recommendation up there, I didn't hear anybody say, no, take it down.

DR. SCHUELLER: I agree with that Genny just said, and I have no disagreements at all with anything you have said, and so I think it's fine to leave it up here. I think that we could maybe make a few more statements about what we want out of this, and so this explore the effect of sample size on precision estimates is a much more complicated statement than it appears to be here, and so there are some species that are perfectly well sampled with the current setup, because they're more common species, and so we can -- I mean, with something like this, you could probably make statements about sample sizes with respects to different types of species, including things like rare-event species, and so this could help guide other things, such as where could they put sampling effort into in order to improve estimates for certain species of interest, and so it's not just about the first bullet. The first bullet is sort of a step, in my mind, to the second bullet, which is what we really care about in a lot of circumstances.

MR. FOSTER: Just a few comments, I guess, from -- Well, from my perspective. One is this simulation idea is a reasonable thing to do. It is not an unreasonable request, I think, but I would

like to -- Because, as we've discussed, it's something that we already had in the works to do, but it's just not been something that we can get completed, because of competing priorities and limited resources. Again, from our perspective, it is primarily a tool that we would use to optimize sample allocations, but, if results from it can be used for other purposes, that's great.

In terms of a validation tool though, the other thing that I would point to though is the fairly lengthy number of peer reviews that have already been done on different aspects of the program, and so the estimation -- The weighted estimation approach, and not just for the individual components, but for everything together, that has already been through an MRIP certification process, which includes an independent peer review.

The individual survey components, the FES, the APAIS, their designs and estimation procedures for the components coming from those, those have been through the MRIP certification process, which included peer reviews, and they have further been through the National Academy's peer review, conducted in 2016 with a report released in 2017, and so, from that perspective, there are a number of reports, documents, that already exist that could be used, I think, as references for have any of these things actually been looked at, and the answer is, yes, they have been reviewed fairly thoroughly, and so I would just hope that folks can keep those things in mind as an awareness when we're talking about communicating to others what has been done to sort of evaluate the program, but creating the simulation engine and using that is not an unreasonable thing to do.

DR. LANEY: To John's point, if the SSC would find it useful, it would be nice to just go ahead, for the council's benefit, to list all the other studies that have been done, for their reference, that tested the various components of the whole program.

DR. ERRIGO: I can get those offline, and then I can actually get the actual reports and everything, and I can send them to the SSC and have them available for the council.

DR. SEDBERRY: All right. I would like to suggest -- Go ahead, Erik.

DR. WILLIAMS: Just to that point, I understand what John is saying about creating a simulation engine, and maybe perhaps a shortcut would be to take one of our most data-rich species and then sort of sub-sample it down to where you get spiky, crazy estimates, but that might be a simpler way to sort of address this, what I agree is a good thing, is a variance sample size tradeoff sort of plot of some sort, and maybe something like black sea bass, which is data rich, might be a way to work from that, rather than ginning up a whole simulation code.

DR. SEDBERRY: I am going to suggest we take a break here, and maybe Mike and I can work on the wording of this recommendation, to make sure we have it right. Then, when we come back, let's be prepared to talk about -- To finalize a statement on the assessments that are on hold and the assessments in general and can we move forward, based on what we have accomplished at this workshop, and so let's come back at 11:15. Thanks.

(Whereupon, a recess was taken.)

DR. SEDBERRY: Just as a reminder, we have the calibration presentation scheduled for 1:00, which means we need to go to lunch early, which means we're going to break at 11:45, and so that will give you an hour-and-fifteen-minutes for lunch, and we usually have an hour-and-a-half, but

I would like everybody to get to lunch early, so you can get it out of the way and get back here by 1:00 for that presentation, and so we'll break at 11:45.

DR. REICHERT: It may be useful to also mention that this is not the final report and that we'll wordsmith some of the statements to make it clearer, but also, because this is the text that's on the webinar, I want to make sure that everyone realizes that this text may change a little bit relative to the final report.

DR. SEDBERRY: Right, and so, before I could present this to the council next month, I would have to -- I would need some context to go around that recommendation there, so it will have some meaning to me and I would be able to explain to the council better what we meant by this, and so we're going to wordsmith some of these things offline, so that we can move on to a recommendation regarding stock assessments, the stock assessments that are on hold and just in general how the SSC feels about moving forward with stock assessments, given the information that we now have on MRIP. Mike has the cursor in place to start the words for this. Does anybody care to start?

DR. REICHERT: I think we have a statement that's above that, and so I'm going to come back to a statement that I made earlier. I think it's important for us to make sure that we make a general statement, and then there may be a statement or a recommendation that we want to make relative to these five species, and so I think, if you go up, we made a general statement about using the MRIP for ACL, and I think we can come up with a similar statement relative to stock assessments, and I recommend to make that a general statement, unless that statement is very specific to the five species that are listed under there, because we already have said that we consider that -- We have reaffirmed that we feel that this is best scientific information available.

DR. SEDBERRY: So let's make this a consensus statement. Mike is drafting a consensus statement there, and, of course, we're going to hear more about calibrated estimates this afternoon, and so, if anything changes, we can come back and edit this further, but I wanted to get something down on this, because time is going to be short after the presentation this afternoon, and so want to get a start on this, and, if anything changes, we can come back and edit it.

DR. LI: I just want to make a notion that I am personally comfortable with using FES calibrated estimates post-2000. For pre-2000, I am still waiting to see the calibration presentation this afternoon.

DR. SEDBERRY: Okay, but, depending on what happens this afternoon, yes, I understand.

DR. WILLIAMS: Does the phrase "going forward" exclude MRIP revision analyses that were presented to you in October?

DR. SEDBERRY: Marcel, did you have something?

DR. REICHERT: I am trying to remember. Erik, you made a statement relative to those revision assessments, in terms of the use for management, or is my memory completely lacking me?

DR. SEDBERRY: You had a statement that you read into the record at the end of the discussions at the last SSC meeting.

DR. REICHERT: I mean, that's relevant to this, right? Just to clarify, what you're asking is are we or should we not include those in this statement, and, in other words, should we use that or should we ask for new assessments, or new estimates.

DR. WILLIAMS: I don't want to confuse your statement. I think your statement should be very concise and just say that we endorse using FES in assessments. I guess what I'm suggesting is, given that statement, then what does that say about the MRIP revisions that were done? Does it change anything about what was decided by the SSC with those MRIP revision analyses? The answer may be no.

DR. SEDBERRY: I have that statement. Maybe we can paraphrase.

DR. ERRIGO: There you go, and I will highlight that, to make sure that we come back to it after the discussion with Jay Breidt, in case we want to make any modifications.

DR. DUMAS: Looking at the example species that we talked about yesterday of red porgy and greater amberjack and king mackerel and golden tilefish, all those, looking at red porgy, out of that 2016 year, and looking at the FES estimate versus the other estimates, the FES estimate is very different, and I still don't think we know why. If you look at the FES estimates compared to the other unadjusted estimates -- We were comparing three types of estimates, the base, ACAL, and FCAL, and FCAL were the FES and APAIS-adjusted estimates.

Looking over the time series, the three types of estimates are very similar for a lot of years, but, in 2016, the FCAL estimate, the FES adjustment, was very different, and there are similar years for the other species that we looked at, and I think the public is going to want to know why. If we think that this FES APAIS-adjusted and FES-adjusted method is a better method of producing estimates, and I believe we do think that, because it's been through extensive review, and the review says it's better methods, and, based on the presentation that we heard yesterday and today, I'm convinced they are better methods, and so the method is better, the method is producing similar estimates to the old method, for many years, but, in some years, the new method is producing very different estimates. Why?

Is it only because of the landline, the difference in landline households, and, if that's the case, maybe we should be seeing very different estimates in the years 2013, 2014, 2015, and 2016 between the FES estimates and the old method, but we don't. We only see the difference for a few years. I am thinking that the new method is better, but that these large differences that occur between the methods in some years are due to differences in the input data, that we've got some type of outlier input data that are driving the differences in the estimates.

If that's the case, we need the method to be able to identify the outlier input data, and I think we talked about that briefly, but not a lot, and so, if you look at the example species, the time series graphs, the data points, do we know -- In 2016, that high estimate for the FES, that is very different from the other estimates, from the other methods, but, in the twenty years before that, the new method produces similar estimates, and so why is the estimate so different?

If it's just sampling error, then the confidence intervals around that have to be really large, I would think. I am being shown the confidence intervals around that, here by my colleague on the right,

and the confidence intervals are very large around that. The confidence intervals appear to be so large that they include the estimates from the other two methods, and so that's saying that there really is no statistically significant difference between those three data points. If we were considering making management changes based on the differences between those three data points, then maybe we should not make management changes, or think twice about making management changes, because those three data points are not statistically significantly different.

DR. REICHERT: But we have concluded that this is a better method, and it's the only method we currently have in moving forward, and so that creates an issue if we are not changing management, because we concluded that we have a better method, and it's the only estimate we have, and so I'm not sure where --

DR. DUMAS: But if the better method produces wider confidence intervals, then that means the better method may be telling us that we're less certain, and so the better method may be telling us that we cannot discern a difference, and so we can't say whether something has changed or not, in terms of catch or in terms of landings or harvest.

DR. REICHERT: But changed in terms of the time series or changed in terms of the difference between the current method and the old method, because we discussed the differences between the old method and the new method and why those differences may occur, and that goes back to that we concluded that this is a better method, and am I making sense?

DR. DUMAS: Right, and I would agree this is a better method, but the better method may produce an estimate that's not statistically significantly different from the estimate produced by the old method.

DR. REICHERT: But is that the case in all cases, or is that the case in the points where the -- The high points, because I think, yesterday, John mentioned that, the higher the point, the higher the confidence intervals.

DR. DUMAS: Right, and it's not the case in all cases. It's only in select cases, and so, for some species, the new method may produce different estimates, statistically significant different estimates, from the old method, and, in those cases, I would go with the estimates produced by the new method, because I believe the new method is better, but, in some cases, the estimates may not be different. They may be so -- The confidence intervals around the new estimates may be so large that they include the old estimates, and am I understanding that the reason why we're here in the first place is because the new estimates appear to be very different from the old estimates?

DR. ERRIGO: I actually understand what you're saying, and I agree. Like if you're trying to -- If you see a spike like that in a given year, and you're deciding whether you're going to close your fishery or not, but the confidence intervals are huge, you may want to think twice. However, in this situation, in the historical data, we're mostly thinking about putting this into an assessment. Now, if they're not statistically different, I think we still would have to use the FES estimate with the uncertainty estimates around it, because, if we use the base estimate, we would have to use the uncertainty estimates from the base, which we know are incorrect.

DR. DUMAS: I agree with that completely. I just think the uncertainty estimates around the new point estimates are very important to management, going forward, especially if the uncertainty

around the estimates is larger than the uncertainty around the old estimates, and the uncertainty around the new estimates is a better estimate of the uncertainty, and so we should go with that. The uncertainty around the old estimates was perhaps too small, and so the new method produces better estimates of uncertainty, but we should be sure to include those and consider those in management decisions and management recommendations.

DR. LANEY: With regard to this particular time series, and that particular 2016 point, it brings to mind the discussion we had yesterday about outliers and Erik's comments about outliers, and doesn't that come into play here as well, at least with respect to the assessment teams that are working with these individual species? I mean, in this case, looking at the whole time series, there is just that one point, at least in the lower graph, where there is an extreme departure from the previous methods for estimation, and so --

DR. ERRIGO: This is the discards here, and this is the landings up here.

DR. LANEY: Okay, but does that outlier discussion come into play here as well?

MS. LANGE: For the assessment purposes, I believe that's why we have the data workshops, so that, if there is an outlier, or what appears to be an outlier, the members of the data workshop can evaluate whether or not that needs to be adjusted or used as-is, and that would be based on digging into how that number came about, and so I don't think there's an issue, really, with it.

DR. VAN VOORHEES: I just wanted to clarify, because there were concerns about precision of the FES-based estimates. In the years 2015 to 2017, when we conducted the FES side-by-side with the CHTS, the sampling levels we used for the FES were lower than what we are using from 2018 forward. We have actually increased the sampling in 2018, and we'll be maintaining a sampling level than what we used during the benchmarking period, and that was partly because of limited resources, in terms of being able to afford doing both surveys for three years. I just want to make sure that people understand that what you're seeing in 2016 is somewhat influenced by that lower sampling level that we had during the benchmarking period, but the estimates should be more precise, moving forward.

DR. SCHARF: Just for clarification, isn't this data stream treated like any of the other data streams in the assessment, in terms of the fishery-independent surveys, where the CV, the coefficient of variation, is incorporated into the assessment, in terms of -- Maybe not necessarily explicitly in a weighting scheme, but in the -- In other words, the error around the estimates is incorporated into the assessment.

DR. WILLIAMS: Just to answer that, yes, but not probably in the way that people think. Just to clarify this, and I've said this many times, but catch-at-age models are conditional on catch. You assume that the catch is pretty well known in sort of the base run, but then what we do is, in an MCB uncertainty analysis, we then re-sample the catches, given the PSEs, and then re-run the model thousands of times, and so, in that sense, the uncertainty is put in, but the clarification that I wanted to make on this is that a high estimate of landings is not going to get drawn down in any way in the assessment. If it's fed in as a high number, it's going to go in as a high number, but it will have uncertainty around it during the bootstrap procedure, and so that's just one distinction, to make sure that everybody understands that.

DR. BUCKEL: I think this is where the simulation may help out, and so we're looking at this 2016, because this gets to Erik's point yesterday of identifying outliers is tough, and there may be some others in here that have issues, and maybe the simulation -- I think, for this case, there was some issue in 2016 with the time of the day that the samples were taken, and so, maybe through simulation on some generic reef fish -- You guys have created the generic reef fish life history, and that could be simulated, and you find that time of day is really critical, and there may be three years in this time series where time of day wasn't sampled adequately, and then those three years could be dealt with in the data workshop, and it would point us in that direction, and so I think that's -- Just to get back the simulation, that could be used to help identify what about the sample size or time of day, et cetera, that could point to these problematic years. They may not look like an outlier on the graph, but they -- Thanks.

DR. BELCHER: Just to remember too that that's a product of FES and APAIS, and so the simulation is actually being talked about relative to the effort and not the catch side of that.

DR. SEDBERRY: Can we go back now to the consensus statement that we were working on regarding going forward with assessments? There is our consensus statement that has the highlighted part that we might need to come back to after hearing the presentation this afternoon, but let's make sure that's the consensus of the SSC for now. Is there any objection to the way that's stated right now? All right, and do we have a similar statement about ACLs, or are we finished that one? Okay. It's just above it. I see it now.

DR. REICHERT: So these are our general consensus statements, or at least the first draft, and they may change a little bit, and so, in terms of the five species that we have looked at, that we have seen some data for, are there any specifics that we need to address in terms of the issues that we discussed in the last couple of days that we need to add to that statement, but then below, when we are talking about the individual species? We talked a little bit about some of the, quote, unquote, outliers, and the analysis was very good, in terms of, okay, what potentially can explain those outliers, and we talked a little bit about that, but I am not sure if there are any specific statements that we need to make under these species.

DR. WILLIAMS: Just, given your consensus statement, when my team moves forward with assessments, we're going to take the FES numbers and plug them right into the assessment as they are, and so, if anybody has any issues with that, this would be the time to bring it up.

DR. NESSLAGE: In a typical data workshop, all of the estimates, data sources, are scrutinized. I don't see why we have to have the FES-based MRIP estimates shoved down the throats of the analysts just because the SSC says this statement. One of the reasons why we balked at the revision assessments was because they were just slapped in in a turn-of-the-crank run, and, in my experience, turn-of-the-cranks blow up in your face more often than that. I'm not a big fan of them, and I know the SEDAR process has some allowance for that, but then you end up getting bit in the butt, because people don't like the answers.

I have a real problem with just saying stock assessments, and what does that mean? Is it all stock assessments, or do we want all new stock assessments that are going to use the MRIP estimates to go through a data workshop first? I thought that's where we were going at the end of our last meeting. if I'm wrong, please correct me.

DR. SEDBERRY: No, you are right that we did have a stipulation about them going through the data workshop of the SEDAR process.

DR. SHERTZER: I think we're also moving towards sort of a different paradigm with doing assessments with having research track assessments and operational track assessments, and so, in that spirit of trying to produce more assessments for the council, not every stock is getting a data workshop for every assessment, and so we're not going to have that luxury of going in with a big group of people to look at every single data point and every single data source, and maybe there is some tradeoff there, but that's what you're proposing, if we're going to do data workshops for every species, we're going to have to slow down doing assessments.

MR. CARMICHAEL: The reality is we haven't done that for every assessment in the past. It's true that we're going to a research track, but that's really an extension of what had been the benchmarks, and we have done many standards and updates over the years without doing the full-on data workshop. There is a group of people that are assigned to those, in most cases, certainly in all of the standards, and not in the updates, but they have a chance to look at the data. They do several webinars, and they go through the data, and, in fact, it was one of those very webinars that was going on for greater amberjack that happened after this SSC had looked at those revision assessments that led to questions about that data.

The group just had trouble resolving what to do with that, and they were -- I guess, in their minds, was the thought of what's going to happen on a bigger picture with all of these data, and there was a hesitancy to move forward, and so we do have a process in place to look at the data, where it doesn't have to be used blindly, and there certainly can be some discussion of how the analysts could use sensitivities or other things to deal with any concerns that folks have with the data, which is I would say exactly what we do with every other data series that we work with, and so I think, within SEDAR, we have a process that makes sure that can happen, and they are not obligated to just plug in the numbers blindly, particularly if some look suspect or you believe they deserve some further scrutiny.

DR. SHAROV: On the issue of this discussion at this moment, I think it's directly related to the previous consensus of the SSC finding this survey as the best available scientific information, in this case, the scientific methodology and the information, when you are using the numbers, because I think the meaning of this consensus here, essentially, is to say that the SSC did not find any -- Did not identify any specific issues with the FES or APAIS and in the total estimation process, and, therefore, you are endorsing the use of these estimates from 2017 back, and so that's my understanding.

I wanted also to mention that, in multiple places here, you are focused just on FES. Here, you have it highlighted in yellow that we endorse the use of the FES-calibrated estimates, and it's not just FES, but it's APAIS as well, and FES really is the effort survey, is it not? John had shown, yesterday, that, in some cases, for certain species, the correction or adjustment back in time, as a result of the APAIS recalibration, has had a significant impact as well, and so I understand that you are endorsing not just FES, but the APAIS recalibrated estimates, and so, essentially, the overall MRIP estimates, whatever the acronym for the survey is that you want to use.

DR. SEDBERRY: Thank you, Alexei. We just had that highlighted to remind us to come back to it, but I appreciate your comments. Right now, it's 11:47, and we really do need to break for lunch,

so that we can be back here in time for the scheduled webinar speaker, and so I hate to cut people off in the middle of questions, but if you can hang on to them, and can we wait until after lunch to deal with those, or maybe even after the presentation to deal with them? Okay. All right. We are recessed for lunch. Be back here at 1:00, because it will start promptly. Thanks.

(Whereupon, a recess was taken.)

DR. SEDBERRY: Okay. Our presenter, Jay Breidt, is online and ready to go, and so we will resume here. The agenda item is revising the FES calibration model in greater detail, which addresses TOR 1A, and the presentation is an in-depth presentation on the FES calibration model by Dr. Jay Breidt of Colorado State University, and so I think he's ready to go, and we will just jump right into it.

REVISITING THE FES CALIBRATION MODEL IN GREATER DETAIL

DR. BREIDT: I think you all know the problem that motivated this work, and I'm sorry to be jumping in in the middle, or not even in the middle, but late in the stage, and so I don't know what all the conversations have been, but the basic problem here has been two surveys with partial overlap leading to very different estimates, and we were asked to address this problem with a calibration methodology that would allow sort of a currency conversion between the two different survey methods, and I don't think I probably need to show the magnitude of the differences between the mail survey estimates and the telephone survey estimates here, but it's here on a logarithmic scale for Florida private boats, and this is Alabama shore fishing. Again, the pink triangles are the mail survey effort, and the blue dots are the telephone survey, and so there are big differences and small overlap with which to understand those differences. On the log scale, that's what it looks like, and so there's the problem, which you all know.

The problem we were tasked with was is there a way to really convert from these sort of telephone units to the mail units and vice versa without making a judgment that one method is correct, or the gold standard, or, even better, and they're just different, because they are different methodologies, and can we come up with a defensible statistical procedure that would allow us to sort of make the conversion between these two different currencies, and this was, of course, something you don't want to have to do, and anything to do with the situation is going to have to rely on some model development and some modeling assumptions, and so the intent here is to make those assumptions pretty explicit, to show what goes into the machinery of currency conversion.

The way we do these things in statistics is to start by thinking about sources of variation and what is causing the different sources of variations in the two different data streams. The mail and the telephone estimates are different from each other, but they also have other variation. They vary in time, and they vary in space, and they vary state-to-state, and they vary across years, and they vary by wave.

Some of those are somewhat predictable. The variation by wave, there's a certain structure, and there's some trending in the year, and so we tried to look at those different sources, as sort of depicted by these box plots here, and think about how we can start to decompose those different sources and describe them.

When we talk about spatio-temporal variation, what we want to do here is try to explain as much of the spatio-temporal variation that is shared across these data sources, explain as much of that as possible, and then look at the differences that are left over we remove that sort of major source of variation and try to -- The thinking there is that both of these data sources should be seeing the spatio-temporal variation. They should trend. They should have seasonal patterns, and anything that's left after that we would call in statistics the irregular part, which is what is known as the classical time series decomposition to the trends plus the seasonal plus the irregular. If we can remove those trends and seasonal pieces, or explain part of them, then we can focus in on the irregular and try to decompose that.

The trend, we know there is going to be effort that varies over years, and trend doesn't necessarily mean that it's consistently going up or down, but it's just sort of a long-term pattern that we can see, maybe partly due to population changes. Seasonal, there is going to be effort that's varying wave to wave, and that variation is going to change state-to-state. It's going to be more pronounced in some states than others, and then there is the irregular part, and so, even if you could explain the trends perfectly and the seasonal perfectly, that true effort would have some additional sort of idiosyncratic variation that wouldn't be explained by just the plain old trends plus season pattern. There is other stuff, sort of idiosyncratic stuff, that maybe we could explain if we had further information, but, otherwise, it's just sort of additional variation above and beyond the trend and seasonal. Our basic models for effort, if we could get at the true underlying effort, would be that trend piece, the seasonal piece, and an irregular piece for each state.

As I was saying, this model is called a classical decomposition model. The effort is trend plus seasonal plus irregular, and the problem is we don't get to see the true effort directly . What we see is the effort that has added to it sampling error, because we have to draw a sample and base our estimates of effort on a sample, but, also, there is additional effects that are due to the method of sampling, and so those include a lot of different effects, but we'll kind of look into those in more detail down the road.

The model here says let's work on the log scale, and so just take natural logarithms of the estimates of effort and the -- Whenever I don't say otherwise, I'm working on the log scale. It says that the estimates from the telephone survey, which are written as "telephone" in black, and so the telephone estimate is made up of these different pieces. Telephone estimate, in part, includes the true effort, and so it's an estimate of true effort, but it's got two other pieces in there. There is the effect due to the telephone method of sampling, and then there is the sampling error that comes from the telephone sampling methodology, and similarly for the mail. This is the mail sampling error.

The effort is the part that we have broken into the classical decomposition, trends, seasonal, and irregular. Then there is the two sampling errors, and then there's these method effects, and so I want to talk about those, in turn, and the effort -- As I said before, we want to think of that as trends plus seasonal plus irregular, and the trend part, we sort of expect, in some way, depending on population, and log population is a pretty good predictor of overall trend in effort, and it has decent predictive power, and we have those numbers from the U.S. Census Bureau, and so we've got state-specific population, and so that's a way to get at the kind of overall trend, to explain some of that structure in the overall trend.

Then we've got the trend plus the seasonal part of the model, and we want that to vary by state, and so, basically, that is what this piece of the model is doing. We say that we have the six two-month waves, and each state will get its own set of six two-month waves, which we parameterize by constructing some dummy variables for our regression, and we put that in, and that gives us a way to describe the wave-to-wave variation. Together, we've got state effects that kind of shift the overall trend up or down, and we've got each state has its own coefficient of log population, to get that kind of overall trend, and then there are these state-by-wave effects, so that every state has its own particular seasonal pattern.

Just that very simple model actually accounts for a lot of the variation in the telephone survey, and so that's not too surprising, that trend and seasonality explains a lot of this variation, and even just different values, sort of reflecting that a lot of that variation has been captured in this simple model.

What's left over? The pieces that are left over are the irregular part of effort, the two method effects, either the phone method effect or the mail method effect, and then the sampling error, and so the irregular part here -- What this means is that, if you can actually measure true effort, and so if there were no sampling -- If you didn't have to do any sampling and you could actually go out and measure the actual effort, and you have this trend plus seasonal pattern, there would still be additional variation in it that wouldn't be explained by that simple trend plus seasonal, and so there's kind of idiosyncratic variation, and we can't really say much more about it, except that it's idiosyncratic, and it's variable, but it's real, and it's not a sampling effect, and it's not a method effect, and so we model that as a random quantity, and this is a standard approach to this sort of problem.

We model it as a random quantity, with mean zero and some variance that is denoted with this Greek letter here, and, here, it's just a normal distribution, and so we just say I don't know what this is, and, on average, it's zero, but it could go up or down, and it's just sort of this idiosyncratic piece that I can't explain, but it's real. It's not an effect of the measurement methodology.

That is the basic model for the effort, and then we will turn to the sampling error. The sampling errors are actually the nice errors, because they are pretty well understood. They can be either design observational studies, and we know a lot about the sampling error, and it comes from the fact that we have stratified unequal probability samples, and we know a lot about the structure of these, the telephone surveys, and we know a lot about the structure of the mail surveys, and so we know that the sampling error is, on average, zero mean for either the telephone or the mail, and so, if we look at the survey-based estimate based on the telephone, it has in it this telephone method effect, which we haven't yet talked about, and it's got the true effort, which we're trying to measure, and then it's got this telephone sampling error, which at least we know is zero mean on average, and this is just the sampling error, so that the telephone data, the telephone estimate, is actually unbiased for the telephone target.

The telephone target is the true effort plus a piece that we would rather not have in there, but it's always in there, in every real survey, and there's this effect due to the method, and so that's the telephone method effect. It's the part -- The telephone estimator targets this telephone target, which includes the true effort, but it also includes this unwanted piece of the telephone method, and the sampling error here, the telephone sampling error, has zero mean, and it has variance, which is unknown, but we can estimate it, and we can estimate it from the sample itself, and we

do that with the design variance estimate, with a little bit of a tweak that I will talk about in a few minutes, and so that's the telephone piece.

We also know that the sampling error comes from a within-state stratified sample of fairly reasonable size, and so it's not too crazy at all to assume that the telephone sampling errors are independent normal random variables of mean zero and with variances that we actually estimate quite well from the survey itself, and so the telephone sampling errors, and the T here is for telephone, the S is for state, and the t is for the time point, which is the wave, and these are independent normalized variables, and so mean zero, and with a variance for the telephone that depends on where you are, which state, and which timepoint we're talking about, but we have a way to estimate from the data itself.

Similarly, the mail sampling error, it's the same exact argument. It says that the mail estimates are unbiased for the mail target. The mail target includes the true effort, which we don't know, and then whatever effect the mail sampling method has in moving that target away from the true, and so the same kind of story here. We know something about these variances, because we get to estimate that from the data, and so those are the sampling errors, and those are two independent surveys, and so the sampling errors are independent of one another.

The sampling error variances, we have estimates of them directly from the data, but what we really want is estimates of these quantities on a log scale, and so the details of this are not that important for this discussion, but there is a methodological approach we have that's kind of novel that allows us to make nice estimates on a log scale and make everything match up, and so that's kind of a secondary point.

In this particular slide, it actually gives great detail of that novel bit of work, which is, I think, not very important here, but it's just that we can do it, and so it's just the point that we sample data on the original effort scale, but we modeled it on a log scale, and so we have to do a little bit of fixing, and so, skipping ahead then to the method effect, we have a mail method effect, and we have a telephone method effect, and these are the non-sampling errors to come from each of those survey modes.

We have sampling error, because the sampling frame -- It's the same thing as the sample. We didn't sample everybody, because we couldn't afford it. It wouldn't be possible, but we have these method effects due to non-sampling, and the non-sampling errors are many, but the biggies are coverage error, and that's where the population is not the same for the sampled frame, and there is non-response error, because not everybody you sample actually responds, and then there is measurement error, because true responses are not the same thing as measured responses, and all of those can play a role in the method effects, and these kinds of errors can interact with each other.

A lot of surveys, a big part of what you do is you design your sampling protocol, but you also have to do a lot of stuff to the measurement protocol, and those two, designed together, are designed to minimize these method effects, these non-sampling errors, and I'm just calling these non-sampling errors and combining them and calling them, together, the method effects, the mail and the telephone method effects.

The problem with these method effects is that we cannot figure out what they are from the sample itself. Like, previously, we were able to quantify the sampling error from the sample itself, but

you can't do that with these non-sampling errors. The further problems are these method effects change over time, and so you just cannot entirely disentangle them from the effort, which is also changing over time, and so this is a problem, and the confounding -- These method effects are confounded with the effort, and so what does that mean?

The telephone is a biased estimator of the telephone method effect plus the true effort, but those non-sampling errors that show up in the telephone method can have all of the features that effort has, and so the telephone method can -- It can change over time, and why can the telephone method effect change over time? There is error, and why would it change over time? Well, maybe the quality of your frame changes over time, and so your coverage error is increased. Maybe you have a change in your response rate over time, and so your non-response error increases. Maybe, at some point in time, you have changed your measurement protocol, and so your measurement error characteristics at one timepoint are different from your measurement error characteristics at another timepoint.

All of these things can change over time, and so the telephone method has its own trend, its own seasonal, its own irregular component, its own idiosyncratic, non-sampling errors from state to state and wave to wave, and so this is a fundamental problem, and it's the same thing with mail. Mail is an unbiased estimator of the mail method effect plus the effort, but the mail method effect is going to have its own trend and its own seasonal and its own irregular component.

When you're designing surveys, you're trying to make these method effects as small as you can possibly make them. You are trying to use a really high-quality frame, and you're trying to really encourage response and follow-up on your non-respondents, and you're trying to use really good measurement protocols, but there is always affects in there. You are just trying to make them as small as possible, and now we've got these two surveys, each with their own method effect.

We can't disentangle these method effects from the true effort, and this is always true. In every survey, there is going to be some method effect that you have tried to mitigate as much as possible, and you can't just estimate these method effects from the sampling itself. If you could, you would always get rid of them, because they're of no interest, and you really don't want those in there. You want to measure true effort, and so you can't get rid of them directly, but what we can do is estimate the difference in the method effects, and so, if you look at the combined model so far, it says that the mail survey estimates have this non-sampling error, this bias due to the mail method, and there is the true effort, and there is the sampling error, and, similarly, telephone has these non-sampling errors and the telephone method effect, the true effort, and the telephone sampling effort.

We can't disentangle the trends, seasonal, and irregular of the mail method from the trend, seasonal, and irregular of the true effort, and we can't disentangle telephone trend, seasonal, and irregular from true effort, trend, seasonal, and irregular, but, because we've got these overlapping estimates, we can actually the take the difference.

We can look at the difference from mail to telephone, and that gives us mail method minus telephone method, the true effort subtracted out, because they both have true effort in there, and that's now gone, and what remains is the difference in the sampling errors. We've got the difference in non-sampling errors and the difference in the sampling errors, and the nice thing about this quantity is that it's actually an unbiased estimator of the difference in the method effects,

because sampling error is unbiased for zero in each case, and the difference is still unbiased for zero, and this is an unbiased estimator of the difference in the non-sampling error effects.

That is a key point here. Because we've got these overlapping data, these overlapping time series, between mail and telephone, we can actually estimate that difference in the method effects, and we don't have a lot of overlapping data with which to explore the difference, but we have some, and so this is what we have to use, and, if we can model that difference, then we can use that to extrapolate either forward or backward in time to understand this sort of currency conversion and what would have happened to telephone back in the past, if you had used mail, or what would have happened in the future to mail if you were using telephone and so on.

The only way this is going to work is if you can actually look at the unbiased estimate of the method effects and explain it somehow with some observable covariate, and, if we can do that, then we can extrapolate forward in time, and so that's the calibration idea. It's estimate the difference between the two method effects, the non-sampling errors, to the two different modes of data collection, and then explain that difference, in terms of observable covariate, and then extrapolate that forward or backward in time to get our calibration.

This is extrapolation, and you have got the unusual dangers, and we don't know if this model continues to hold over time points on which we didn't estimate it, and, if it doesn't hold over the full range of time, the calibrated values can be wrong, and we want to try to -- We have limited data with which to do that, and we have to think about these things. When we look at the difference between the mail method and the telephone method, what makes those different, and how might we explain it, and so, for example, one of the reasons the mail method and the telephone method might be different is because they have different measurement protocols, and measurement error might be changing over time, and maybe we could find some covariates that would explain that change, or maybe the non-response patterns are changing over time. Maybe we could explain some of that change with the covariate, and, similarly, maybe the coverage error is changing over time.

That is one that we already know. For sure, the coverage error has been changing over time, because the Coastal Household Telephone Survey relied on landlines only, and, of course, there is lots and lots of wireless-only households, and that has grown over time, and so we know that the landline-only frame increasingly missed wireless-only households, that increase in coverage error, and so that's one thing to look at.

There are some data available on wireless-only households that basically comes from the National Health Interview Survey, which is conducted by the National Center for Health Statistics, and they have half-year estimates for each state over a window of time of interest, the 2007 to 2014, and what they get from that survey's estimates is the proportion of wireless-only households, and what we did is just extrapolated this forward and backward in time, and I used that as a covariate, and so it does approximately what you would expect, that, if you extrapolate it back in time, you get approximately zero wireless-only households from the year 2000, although it has grown rapidly since that time.

When we took wireless into the model and we interacted with the state, the wave, the log population, that helps to explain some of the variation, and it turns out that wireless is, in a sense, highly significant statistically. There is a lot of evidence that its effect is not zero, but it's effect

is not that large, and so the practical effect of this adjustment is not all that great, but it does explain a little bit of the residual variation.

The model for this mail method minus the telephone method includes the wireless, its interactions, the trends in seasonal factors, and it does some sort of sensible things in time, and there's just not a lot of other covariates that we have available to explain changes in these methods, and so what remains is basically a level shift, and so this is a huge assumption that any difference in measurement error between two methods is constant over time beyond that wireless effect that we can explain. Another extreme you could try would be to just say, well, there is -- Don't calibrate at all, but, if you look at this line, what this says is that there's a level shift, and then this is the wireless by covariate interactions for the various trend in seasonal factors that we've got.

Putting this together then, again, the telephone estimate is the telephone method effect, that non-sampling error, and it's the true effort, and then the telephone sampling error that we know applies to that, and so that's the telephone target plus the telephone sampling error. Similarly, the mail, it's non-sampling error, the mail method, the effort, and the mail sampling error, and then that's the mail target plus the mail sampling error.

We know a lot about both of these sampling errors, and we can estimate the telephone target and estimate the mail target, and the big part of that is that we have this mail method minus the telephone method part of the model, which is -- This is an inherently hard problem, and that's where the biggest assumptions lie.

In terms of notation, replacing words, the telephone estimates on a natural log scale is this T-hat, and the telephone target is this T without the hat. Telephone sampling error is this e, and M-hat is the estimate of mail effort, and M is its target. e^M is the mail sampling error, and this is the Greek letter Nu, and st is the irregular term. These are the modeling assumptions, that the sampling errors are normal, and they have variances, which we don't know, but we can estimate well in the data itself, and then there is additional fixed term, which we don't know --

We put all of this together, and we get this horrible-looking expression, but the important thing to know about this horrible-looking expression is that this is a well-known standard model. It is what is called the Fay-Herriot model in survey statistics. It comes from a paper by Bob Fay and Roger Herriot, which has been around for a long time, and it's a very standard and well-studied methodology that is used in what is called small-area estimation. There is lots and lots of citations, and it's tied into our whole statistical techniques of estimation and prediction, and there is software. There is packages in R and things like that, and so it's just a well-known, well-studied methodology, and the way we formulated this problem, the way we wind up doing our calibration, is exactly an application for this Fay-Herriot methodology, which is well developed.

The way we do the Fay-Herriot approach is we have to estimate this parameter ψ, which is the variance of that irregular trend, and we have to estimate the parameters of our models that connect the two data sources, these parameters beta, and then, once we've got that, we can predict anything we don't know using a standard methodology called empirical best linear unbiased prediction, which abbreviated as EBLUP, and we can ask questions like what would be the mail target equivalent for state s, and past-year wave t if there is not mail estimate available, and we can just compute the empirical best linear unbiased prediction, the EBLUP, of M_{st}, and what that will do is combine all the available information.

If there is no mail data at all, then it's just going to take everything it knows, including the telephone survey estimate for that state and that year wave, and use that as its estimate, as its predictor, and so on. Similarly, if you wanted to come up with what would be the telephone target equivalent, if there's no telephone estimate, you can compute the EBLUP for that telephone target, and that would combine all the information you've got. Then the remaining step would be to estimate the mean squared error of those resulting EBLUPs, in order to quantify the uncertainty in those estimates.

The alternatives to this EBLUP technique is one would be that you just do a really simple calculation to just say, if I wanted to go from telephone to mail, I would take the telephone, and I would just add on the difference in these method effects, the mail method minus the telephone method, and, if you were doing that with just a level shift, if you just said that the difference between these two is constant over time, that would just amount to taking your telephone estimate and adding on a constant everywhere, and that's on a log scale, and so, on our original scale, that would just be a multiplicative adjustment, the same multiplier everywhere.

Because we have both the level shift and the wireless, and the wireless can change from state to state and wave to wave and year to year, we have this slightly more complicated predictor if we didn't do the EBLUP, and this is kind -- This isn't the EBLUP, and it's not optimal, but it's simple, and it's pretty useful for comparison, to see what the EBLUP is actually doing, and what the EBLUP is actually doing is not only this bit, but, additionally, it's estimating that irregular term.

These are some examples of telephone estimates that are calibrated to the mail target, and so this is back on -- This is on the log scale, and so the blue dots are the telephone survey estimates, and the pink triangles are the mail survey estimates, and we go through this fitting technique, fitting trend and irregular seasonal patterns, and then we get these different predictive values here is the EBLUP, and the black, on which the gold is lined, is not the EBLUP, but it's sort of a step simpler, and it's just take the cellphone estimates and adjust it for a shift in the wireless effect, and then the gray is if you did a shift and ignored the wireless effects and just shift everything. You can see that the estimates are all pretty similar, but there are some small differences, due to the different effects of different pieces of the model, and similarly for Alabama shore fishing. That's the original estimates and the calibrated estimates.

The summary here is these models are accounting for various sources of variation, and they include the trend, which is not only the irregular, the sampling error, and then the non-sampling errors get dumped into the method effects, and we model the difference of the method effects, because we have unbiased estimates of the difference in method effects.

We have to rely heavily on modeling assumptions here, and the modeling assumption is that these differences in the non-sampling errors are stable over time, the way that the measurement and non-response differences are stable over time, because we don't really have an alternative way of predicting. The model also assumes that the coverage area has changed over time, due to the growth in the wireless-only households, and that part we actually can understand completely, and we have information about the proportion of wireless-only households.

When you formulate the model in that way, the calibration methodology turns out to be a very standard statistical procedure, this Fay-Herriot small area estimation technique, and so, under the

assumptions of this model, you know statistically what to do. Once you have got that part down, under those assumptions, the calibrated values are exactly the empirical best linear unbiased predictors, which you can get from standard software, and that's a nice, reproducible feature of the methodology, with the caveat, of course, that this is heavily modeled, and, in order to do this problem, you need a model, and that model relies on modeling assumptions, and so that's the approach that we have used, kind of in a nutshell, and there is details in there that I have glossed over, but that's sort of the guts of it.

DR. SEDBERRY: Thank you, Jay. Do we have any questions? Does anyone have any questions or discussion points on this presentation?

DR. ERRIGO: Everyone is absolutely perfectly clear on how this works? That's awesome.

DR. LI: Just some clarifying questions. Thank you so much, Jay, for such an excellent presentation, and so I personally have a much better understanding about the whole process. Then the question is, eventually, you are modeling -- The response variable is the difference between two observed efforts, right?

DR. BREIDT: Right.

DR. LI: Because, earlier in the presentation, I understood that it was a difference, the response variable is a difference, but, later on, when you showed the whole formulation down there, it's separate. Your formulations in your modeling, the response variable, they are two separate response variables, and so I am confused.

DR. BREIDT: No, you're exactly correct. What we actually do -- Going back to that slide, the complete model specification here gives us the T-hat and the M-hat, but then what we actually model is the difference between two, and, once, you take the difference between the two, that's when you get into the Fay-Herriot structure, and that's exactly what we do, is model the difference between the two, and this is just all the pieces, and the last line, that isn't actually here, is we take T-hat minus M-hat and model that difference, and so it's a good catch on your part. That's what is not showing, but that's what we actually do, is model the differences.

DR. LI: Okay. Thank you, Jay. That's great. I mean, it's a smart way to model the difference, instead of absolute variables. Another question is, given that we are modeling the difference, and so after you subtract each other, and then the rest you have left is the non-sampling errors, or do you call that difference methods, and the difference in sampling errors, these terms, and then my question is, for this term, you use predictor variables to model each term, or do you just assume that it's an intercept, random intercept?

DR. BREIDT: The difference between the two method effects, the mail and the telephone method effects, we model with the wireless, with the state, the wave, the state-by-wave interaction, and that level shift, and so it's all of those terms, and so those terms are allowed to -- That difference is allowed to have a trend and seasonal pattern as well as an irregular, and so that's what we modeled with the covariate.

DR. LI: Okay. Just to recap, to make sure my understanding is correct, the whole formulation, the whole model, is the response variable is a difference in the survey estimates, or the survey

observations, and that equals -- Here on the left side of the model, we have the difference in methods, and that's one term, and then plus the difference in sampling error, and that's a second term, and each term, each of these two terms, is modeled using variables, states, population size, and the wave, and the values, and their interactions.

DR. BREIDT: Correct, yes. I guess I shouldn't have quite said it that way. I mean, that's the gist of what's going on, but the way we actually get, in the Fay-Herriot -- I am worried that this is going to be confusing. That's the underlying idea, as you just said, but the way we actually implement the Fay-Herriot is slightly different. We don't actually use the differences directly, and that's more of a conceptual way to present the model, why it works at all, but the way we actually enter the data into a small area estimation package is in terms of telephone only, if we have telephone only, mail only, if we have mail only, and we actually use the average of telephone and mail if we have both, if we have an overlap.

This actually works, and, conceptually, the -- I didn't explain it in the clearest way, but, conceptually, the reason it does work is because we have an overlap, but, because of that overlap, we can estimate these two pieces, but, yes, those method effects are -- We are describing those in terms of covariates. Sorry that was a little confusing. It is not the case that we just take the differences between mail and telephone and use only those data in the small area model. That is not the case, and that's more at the conceptual level, like why we're able to make these models work at all. We actually use all the data, all the telephone data, all the mail data, and with the overlap.

DR. LI: Thank you, Jay. In terms of those variables that you used to explain those differences, the population size -- State, I can understand, and we can clearly see that the state variation explains a lot. Wave, I can understand that, and the population size. For these variables, the variance for sure we understand. It's a big changing methodology, and so, other than these four variables, have you considered other variables, or is there any other variables available that you think that could be critical?

DR. BREIDT: We definitely thought about that and tried to think of other covariates that might be helpful to explain, and the conditions for those covariates would be that they would need to be available now, in the past, and in the future, measured in a -- Maybe it could be explanatory for the effects we're seeing, maybe explanatory for the measurement error, maybe explanatory for the coverage area, maybe explanatory for the non-response, and we talked about different things, and we fiddled around with a few things that we did have available to us, but we never really could come up with much that was useful that would be predictive and available to us.

You can see the things that might be useful, and so a lot of the -- We couldn't actually get state-by-state, wave-by-wave, and so we're always open to suggestions, but with a fair amount of -- We didn't get much beyond this basic step. We tried some demographics within states, not just population, and that didn't make much difference, and we tried different age groups, different cohorts and things like that, but not a lot out there that we were able to come up with that was useful.

DR. DUMAS: Thanks for that great presentation. That helped a lot, and so, after doing this, did you do -- I would assume you did like stationarity checks, or did you have any autocorrelation remaining in your errors and your residuals? That's question one.

DR. BREIDT: A lot, as you would expect, a lot of this autocorrelation structure disappears once you remove the trend in the seasonal, and so what remains there is we have modeled the irregular plus sampling error, and the basic models are not correlated over time, and that's a pretty decent model. We have done some checks on that, and most of the structure is gone. You might be able to refine this a bit and add in a little bit about a correlation, but you wouldn't expect to gain a whole lot out of it, but we didn't find much there. A lot of that has been averaged out and then removed by these regressions.

DR. DUMAS: Great. Thanks. The other question is did you have any correlation, spatial, across states?

DR. BREIDT: We didn't really look at spatial structure beyond just basic exploratory stuff, but we didn't actually do much at all, I would say. I may be forgetting, and I may be misspeaking on behalf of my graduate students, but I don't think we did much, in terms of spatial structure there, and so there may be something.

DR. DUMAS: Thanks.

DR. LI: One more clarifying question on your Slide Number 32. Why the single calibration from telephone target to mail target is telephone plus the difference in two methods? Is it telephone target plus --

DR. BREIDT: Could you say that one more time?

DR. LI: The equation here, the first equation, why the mail target equals two telephone plus the mail method minus the telephone method? I just cannot get that equation. Is this supposed to be mail target equals telephone target plus?

DR. BREIDT: This is what if I want to convert to mail units and I have the telephone estimate, and so I have the telephone estimate, which is just denoted as telephone, and now I have to add to it a shift to put it into mail units, and the shift that I add is my estimate of mail method minus telephone method, and so this isn't meant to be an equation. It's just like, if I want to convert that telephone estimate to a mail estimate, what do I do. The simplest thing I can do is take the telephone estimate and add onto it my estimated difference between methods, mail method minus telephone method.

DR. LI: Yes, I understand that procedure. Here, this equation, is it telephone or telephone target in that equation?

DR. BREIDT: It's telephone. If you look at the next bullet with the level shift only, if I wanted to make that adjustment using only a level shift, if I have this estimate in telephone units and I want to convert it to mail units, I would just add on the shift. It would be telephone estimate plus this mail shift gives me mail units, and this is similar. If I wanted to kind of take this telephone estimate and convert it to mail units that account for wireless, I would shift it by the mail shift, and then I would adjust it for wireless. It's just kind of a notation to say how would I take this telephone estimate, which is your T-hat, and convert it into kind of a mail equivalent, and the way I would

do that is I would just add on that appropriate piece. Did that make sense? I see your question of why isn't it telephone target.

DR. LI: I understand why. I am just doing the calculation myself and trying to match, based on the understanding. If T-hat plus M-hat, that gives us M-hat, right, and not M target, but M-hat?

DR. BREIDT: Yes.

DR. LI: Thank you.

DR. BREIDT: It's meant to be an estimate of the mail target, and so T-hat is an unbiased estimator of telephone target, and, to adjust telephone target to mail target, you have to add on the difference, which would be the method, and that's what this is doing, sort of in pieces. Under the simplest sort of model, under the model of improved wireless, what would you do?

DR. LI: Thank you, Jay. Now my question is for the SSC, and I am thinking -- Think about this. We use the coefficient estimated based on three years, and, also, it's very current, like a recent three years of data, and those coefficients for the differences -- We use that one to predict, to explain, the differences back in time. Is it acceptable, or is it good -- I don't know. I just feel --

DR. SEDBERRY: That's a question you're posing to the SSC or to --

DR. LI: One direction goes to Jay and one direction -- I mean, that is what it is, what they have, what we can have, and then, given that, the question is to the SSC is what shall we do?

DR. BREIDT: Right, and I was just going to say that this is not something you would choose to do. This is gun to your head, and you have to come up with a calibration technology, a way to do this currency conversion, and how are you going to do it? Well, our approach is let's write down an explicit model with all of the assumptions listed and tell you what are the assumptions. They are strong assumptions, and we can say, if those assumptions hold, then this is a reasonable thing to do.

If the thing you can do is something that can be reproduced, it is a way to make this conversion, and there is probably a window over which it's quite reasonable. Do those assumptions hold back over time? Probably not. Those are strong assumptions, but it is a reproducible way, and a defensible way, to do it over the window in which that you're pretty confident that your assumptions do hold. If you extrapolate way back in time, you know all bets are off, but it's at least something you can do and reproduce and explain what's going and what assumptions you are making.

DR. LI: I understand that, Jay. Is this method the only one available, or is there any other options that potentially we can test?

DR. BREIDT: I'm sure there is other things that you might test. Our approach, we thought about a number of different approaches that included various time series models and different kinds of measurement error type models, and then we felt, on this one, this method of formulation, thinking let's describe the major sources of variation and then put it into a technology that we know well,

and so the small area stuff is using Fay-Herriot, the real well-established technology, just the ideas of using mixed models and empirical best linear advice predictors.

Once we had that formulation, it seemed like a pretty good way to go, because it's a well-established technology, but I certainly wouldn't claim this is the only way you could do it, and it's definitely not the only thing that we considered. We considered various kinds of time-series-type models, state-based formulation and things like that, but this one is one that allowed us to do a number of things, which was handle the trends in seasonal, bring in relevant covariates, handle the fact that there are sampling errors that we understand and non-sampling errors that we understand a lot less well, but all those sources of variation at least show up in this model.

We are not sweeping anything under the rug. They're all in there, and we've made some attempt to address all of those sources of variation, and, with all of those sources of variation, we have a technology to construct the calibrated estimates that we want, and so that seemed to be the right spot, in terms of a statistical methodology, but absolutely there are other ways you could approach the problem, and so we did give it some thought, but we didn't try to fit -- A lot of those we ran into problems with pretty quickly, some of the measurement error approaches and things like that.

DR. LI: Thank you, Jay. Here are my thoughts. Given we are using a coefficient estimate for current -- Based on the current three years, and we back-predict for further, like back into history, like ten or fifteen years earlier, the difference, I am thinking, along this long time period, Magnuson changes. For example, management options, like code or however, but they may change, and then the fishing behavior may change, how people conceptually understand and do the fishing activity may change.

Those things, I am not sure how much of those changes can be captured in the four variables that are included here, and so I am thinking those estimates from this calibration that are closer to the benchmark time period, I feel more comfortable with them, but those are further back, and I don't know.

DR. AHRENS: Can someone just remind me if the uncertainty associated with this method gets carried forward into the estimates of effort, and, if it does, is that sufficient then to deal with the fact that there have sort of been all these other things going on?

DR. BREIDT: I was just going to say that one component of the model that does capture the additional things that we cannot explain are these -- Those are just additional random terms with an associated variance, and so that stuff that we cannot otherwise explain that may have happened in the past, changes in fishery behavior that we were not capturing with any covariate and changes in management protocols that we're not capturing with other covariates, that just gets thrown into this irregular term that has the variance, but we can't really predict those terms, and so, to the extent that they -- If they have any predictable pattern to them, if they trend with population, if they have state-level shifts, or if they have seasonal variation, then it shows up in the model, but, otherwise, they get dumped into this irregular term, and that's the only place that is in there, and so that uncertainty is reflected in that irregular term.

DR. AHRENS: Does that get carried forward when you have the standard error that comes out of this? Does that get carried forward into the uncertainty associated with the effort estimate?

DR. BREIDT: Yes, it does.

DR. AHRENS: Thank you.

DR. DREW: I think there's also an element of those changes that are driving effort are also going to be sort of picked up in the Coastal Household Telephone Survey, and so what matters is not necessarily whether it's -- For this calibration purposes, it's not necessarily whether it's getting at the correct what's driving effort, but maybe what's driving the difference between the CHTS and the FES, and so the CHTS will sort of be, in theory, tracking that differences in effort caused by differences in management, socioeconomics, et cetera, and then what we want to add on top of that is the difference for the FES, and so is that going to affect how you respond to a mail survey versus a phone survey, or is that just strictly being picked up by the effort through the telephone survey.

DR. BREIDT: That's very much along the lines of what I was saying before, that a lot of the non-sampling errors can't be disentangled from the effort itself, and so, like back in the past, some of those differences -- The way you respond to the survey might have their own trends in seasonal effects and things like that that are -- They are completely confounded with the true effort itself back in the past.

DR. SEDBERRY: Any other questions or discussion? Thank you, Jay. We appreciate you being able to join us. I know we had a hard time scheduling it, but I'm glad you could fit this in.

DR. BREIDT: Thank you so much. Thanks for making the time.

CONSENSUS STATEMENTS AND RECOMMENDATIONS REVIEW (CONTINUED)

DR. SEDBERRY: Thanks again. We had, back in our review of the terms of reference this morning, put in a little note to ourselves that we -- That the SSC accepts the FES calibrations, but it was kind of pending the presentation that we just heard, and, to me, it didn't change anything, but it might in other people's minds in the room, and do we need to go back and revisit that term of reference before moving on to other things? It's the highlighted area there that Mike is showing. No changes? Okay. Marcel, did you have something?

DR. REICHERT: From all I've heard, and from the discussion, I'm comfortable with that methodology, given the constraints and the caveats that were laid out in the presentation and the other documentation, and so I'm comfortable with it.

DR. SEDBERRY: Okay. Thanks, and I think our wording up there captures that, as part of our report and our report back to the council, and so I don't think that we need to change any of the wording in there. Now, the next thing I have on the agenda is that Erik wanted to say a few words. We also have some leftovers from prior to lunch, and let's do that first. Before we broke for lunch, three people had their hands raised to talk about whatever we had up on the screen back then, and I will try to remember what that was, and so we had Erik, Vivian, and Wilson. Do you remember what it was that you wanted to talk about back then? Go ahead, Erik.

DR. WILLIAMS: I wanted to just clarify some things. I had mentioned, very clearly, that, given the consensus statement, the way we're going to proceed with our stock assessments is we're going to take the estimates that come out of MRIP and put them straight into our assessment models.

John Carmichael came up and suggested that there would be an evaluation of that data through the current SEDAR process that might pick up outliers and do something with them, and I don't think that's the case, and I don't think we're going to do that, given current operational assessments. The only time those data might get evaluated is through a research track evaluation, but I put forth that this is a larger issue that needs further discussion and thought on how to deal with it, and I put it to you that the assessment scientists cannot deal with outliers, because, a, we're not capable, really, of identifying them, and that gets back to my statement yesterday about what is an outlier.

In one sense, you can look at a spike in a landings time series and suspect that it's an outlier, but there is possible Type I and Type II kind of errors, in terms of are you identifying all outliers with that method, and so I would say that's not a -- It's a very incomplete method for trying to identify outliers, and so then you're left with is this an ad hoc method for adjusting outliers, and it's not very rigorous, and so what I would put forth is that this topic of dealing with outliers is going to need a larger discussion and more thought, in terms of developing more standard and rigorous techniques for actually identifying outliers and then dealing with outliers.

I guess what I'm saying is I just want to make sure the SSC is clear that, when we proceed with assessments now, we're going to just take the straight time series that comes out of FES, or out of MRIP, and put them into our assessments, so that, when we come back to you with an assessment for your review, and it has a, quote, spike in the time series, I don't want to have to revisit that discussion, and so I want to figure out how we can get past that point.

DR. REICHERT: To further clarify, I think, earlier, I was talking about there may be some efforts going on to talk about outliers, but that's the working group that looks at rare events, and that's not outliers that were discussed, and so that's an entirely different topic, and so I just wanted to clarify myself there. Is this something, Erik, that, in some instances, you explore via sensitivity analyses, or that is something that is mostly done during the research track, or is that still something that you guys will explore during operational assessments?

DR. WILLIAMS: We are not going to evaluate it during operational assessments, which is what these upcoming assessments largely are. We might evaluate them during research track, but the argument I'm putting forth is the way we have been evaluating them in what was benchmark data workshops is a very ad hoc method of dealing with outliers to begin with. It's not very satisfactory, and so I'm actually suggesting that we need a better method for dealing with outliers.

Now, the other thing to add to this is that, as I said earlier, we are incorporating the uncertainty from these landings into our uncertainty analysis for stock assessments, and so, if you're satisfied with the fact that, okay, there's spikes that are going to occur, but they also come with large PSEs, or standard errors, we're dealing with that, in terms of uncertainty characterization, and, if that's satisfactory, then we're all okay, but I get the sense that that's not always the case.

DR. REICHERT: To that point, I am not saying that that discussion, or that further discussion or analysis, about outliers shouldn't occur. I completely agree with you that that should occur, but I am just wondering, in terms of sensitivity runs, which are a normal part of the benchmark

assessment, and sometimes of the standard assessments, remind me, and is that still occurring for operational assessments, because there may be things in the new data that are presented to the stock assessment team that may warrant some exploration.

DR. WILLIAMS: Yes, we definitely do sensitivity runs and would gladly look at things, but, again, I will refresh your memory on what sensitivity runs are and what they should be used for. Sensitivity runs are exactly that, in the sense that you are taking the model and tweaking something about it and looking at the output reaction to that change in the input, but they are not alternate base runs to be used for management, and that's something entirely different.

MS. MATTER: My comment was very similar to Erik's concern. The only thing I would add is that -- A couple of things. We just finished a State of Florida assessment for yellowtail snapper in June, and we did use the FES data, and we did investigate these, quote, unquote, high peaks, and it's something that goes back to SEDAR best practices.

We actually talked about that, and we said this is what our plan is. We're going to potentially, again, ad hoc, maybe look at these and say, hey, does this look weird, and I will continue to do that and say, okay, well, this estimate is coming from this strata, or these groups of strata, and say it was this many intercepts, these are the anglers, and we'll still continue to provide that data for the report, and so it will carry through and inform, and, in addition, we will continue to provide the variances along with those estimates, but, like Erik said, if the assessment group is saying, you know, we don't have the expertise to tweak these, certainly a step further with the data provision step, we will not have the time, with all of the assessments that are going through the whole entire region that we're responsible for. We just do not have the manpower to do the type of analyses.

Think about all of the things that everybody has presented, that Jay has presented and John has presented, in terms of all these calibrations, and I guarantee you that I cannot do anything better, and tweaking those kinds of things -- You are just introducing bias that they worked extensively to eliminate.

DR. SEDBERRY: Thank you for that input. That's helpful.

DR. LANEY: I discussed my comment with Marcel, but the only observation I was going to make was that, with respect to Term of Reference 1A, where we were asked to describe, for a set of SAFMC-managed species currently in the SEDAR process, how the sources of disparity between CHTS and FES affect the FES catch estimate time series, with attention on trends, uncertainty, and potential outliers. John Foster did exactly that for us, and he carried us through that whole analysis in infinite detail, and so I think perhaps the only thing we need to do there, since Marcel pointed out that all of that information is in our briefing books, and the council has access to all of that information, is just to point out that John did that for us and that information is available to the council, and so we might just want to stick that in our report somewhere.

DR. SEDBERRY: So we need a statement to that effect somewhere around TOR 1A as a -- Not a recommendation, but a conclusion, perhaps, or --

DR. LANEY: I would just as an observation that John Foster did that, John Foster and the staff, I presume, and not John alone perhaps, but they did that for us, and he walked us through each one

of those species in infinite detail, and so that information is there, and that TOR was addressed, as far as I'm concerned.

DR. AHRENS: I think it would be worthwhile to point out for this that there was no clear single explanation for the patterns that we're seeing.

DR. REICHERT: Well, with the exception of the point that I brought up yesterday that -- I need to go back to my numbers, and John addressed that point, that about 70 percent of those could be traced back to the Florida private boat data point, and John said that was part -- I forgot the exact explanation, but that's because of the -- What was it, John?

MR. FOSTER: Well, just sort of the relative magnitude of the Florida private boat fishery compared to the other states, and also the fact that these species are primarily, if not almost exclusively, from the boat modes.

DR. REICHERT: Thank you for that.

DR. SCHARF: So I'm not sure where we are in moving forward. Given what Erik was saying about the difference between an operational and the research track, and the operational not incorporating a formal data workshop as part of the SEDAR process, and so we have these assessments that are in front of us on the table, and then we have future assessments that are lined up behind them, and it reminds me of the runway in Charlotte Airport, where there's just planes sort of lined up ready to take off.

It seems like we're sort of between a rock and a hard place, because we either have to -- Regardless of our comfort level with some of these disparities between the calibrated FES numbers and the original numbers, we either have to move forward with these operational assessments, using the data stream as-is, or not, and I don't know what that "or not" means. I mean, it doesn't seem like that "or not" is even an option, in terms of -- You know, we can't go backwards and convert all of these to research tracks all of a sudden, and so is there -- I guess what I'm asking, maybe not for the ones that we're sort of backlogged on, but the ones that are next on the table, is there a way to sort of hybridize this operational track and incorporate some sort of a shorter data workshop for these MRIP numbers before those assessments proceed, because, regardless of how we decide to deal with outliers, it seems like the best place to deal with them is during that setting, as opposed to the assessment group dealing with the outliers.

I mean, there can be advice or guidance from that group, or a protocol for how you deal with them, but it seems like it needs to happen there, with the folks that understand that data best, but I guess I'm just trying to figure out what our next step is.

DR. SEDBERRY: I don't know the answer to that question, and maybe that's something that would have to be taken up by the SEDAR Committee at a council meeting, and I'm not sure, but let's see what other people have to say.

DR. BELCHER: I think I'm going to go back to what we were saying earlier though. I think we need to be careful with, again, the use of the outlier, because there's a lot of times I feel that, when we go into the data workgroup and people are presenting their white papers, those are the people who are most intimate with that data, and so they're not just putting the stream out and not putting

their own due diligence to what, in their mind, is a gut check with what seems like an odd data point. I mean, they don't just do the run and hand it off and not discuss it, and so I think, in the sense of the true outlier, those are the people who are going to identify the problem points and have ways, solutions, to best address how that point is going to be dealt with.

For us, I kind of feel like the question is more -- Like the operational, to me, are business as usual. The question is do we have too much heartburn, again, which I don't feel that we do, with the FES approach, and so those checks are still there. John is still doing that. The data streams that are coming through are still getting those outlier checks from the analysts, and so are we kind of adding too much of a level of check on top of this?

DR. AHRENS: I think, for me, it's, if we've done the best possible job of characterizing the uncertainty in the data that is going in, and if it's captured within the outputs of the assessment, then perhaps there is something within kind of the P* type process that may be not capturing some of the concern that people have, that it's not really an issue with the data as long as the uncertainty is characterized, but it's more kind of what we do once we get those results and make a decision.

DR. ERRIGO: I just wanted to ask another question to Kyle or Erik, if I could real quick. The P* analysis, if there is more higher levels of uncertainty in the parameters that go into the projections and the P* analysis, then that contributes to the buffer, yes? They are saying yes. So it does affect -- The levels of uncertainty that are predicted for the data do affect the P* analysis and the estimates of ABC. I just wanted to make sure, because, honestly, I wasn't 100 percent.

DR. WILLIAMS: Yes, that's true. It would affect the P*, which would then affect the ABC and the buffers, but what it doesn't affect is stock status, because we don't have probability-based stock status criteria. If we move to that, then it would affect that as well.

DR. REICHERT: What I was thinking is, in the new ABC control rule, I think we took out the stock status, and so that may not affect -- There is a mitigating issue, in terms of it affecting the ABC, the ultimate ABC.

DR. WILLIAMS: Yes, and I was just referring to the declaration of overfished or overfishing. That's just a yes or no, and there's no probability statement associated with it.

DR. REICHERT: I agree, and I was just thinking about the next step and where that may percolate further in our ABC recommendations. Thanks.

DR. SEDBERRY: Where does this leave us?

DR. ERRIGO: We're doing great. For TOR 1C, I already started a list of all the critical factors that have been discussed here at the meeting, and I think that it's going to be fleshed out offline, using the briefing materials. They're all in there, if people are okay with that. I think we just need to address what's happening with TOR 2.

DR. SEDBERRY: Yes, because that question just came up, and Marcel has a question, and I need to remind myself and everybody that we need to take public comment at some point before 3:00 as well, and so, Marcel, ask your question, and you and I, Mike, can talk about TOR 2.

DR. REICHERT: I think it's important that we include what Erik just said about -- Because there's a specific question about uncertainty, and I think it's important for us to include in our report that the uncertainty around the estimate is included in the assessment, in the operational assessment, and, in that respect, that is accounted for, and so it's not that we completely -- Those high points or low points, that the uncertainty that that takes with it is ignored in that process, and so I think it's important to put that in there somewhere, and I think it's actually above the text here. Again, we can add that to a draft for the group, or the committee, to review later.

DR. SCHUELLER: I am not sure that we've come to a resolution about our group's recommendation related to the outlier topic, and my sense is, from what's been said around the table, is that, basically, take the data as they are, but then, as they're processed for preparation in assessments, in whatever format that is, they will get some attention, as all the data do when they're being prepared, and so maybe a red flag would raise at that point, if it's necessary, and maybe I'm wrong in what I've been hearing around the table, but I just thought I would put that out there and see if that's what we're saying or not.

DR. SEDBERRY: I'm glad you brought it up, because, these things that still seem to be unclear, we need to clarify them, because they always come back to bite us.

DR. SCHARF: I mean, that's what I would -- That's my perception too, and I just don't know how it would occur. Just operationally, how would that happen? Like the data that John provided to us in that presentation for our five species, where he selected certain data points within those streams and talked to us about what was driving those was great, and I don't know if that kind of information can be readily made available to the folks in those states that have an understanding of that fishery, and so this big FES calibration difference is due to the Florida shore mode during Wave 2 in this region, and then they could evaluate whether or not that was likely to be real, because, most of the time, it's due to the weighting.

You have a low-activity site that has a lot of effort, and it's being upweighted, because it's a low-activity site, and so there may not be anything wrong with it, but it's just because of the weighting scheme and the fact that it typically doesn't have a lot of effort, and suddenly it does, but, again, I don't know how that happens during an operational assessment.

DR. SEDBERRY: I'm not sure about that either, and I'm not sure what we need to say here to make sure that, at some point, that it is evaluated.

DR. SHERTZER: What John showed was more in-depth than what we would typically see at a data workshop, and, even when you get that information, you're still left wondering what you do with the data point. Is it real, or is it not, and you can track down where the estimate is coming from and what's being inflated, but you don't get the answer of whether it's a good estimate of landings, and, just because it's an outlier, it doesn't mean that it's not real, and so that's the next level of dilemma that you run into with what do you do with this datapoint, even when you understand its origin.

DR. SEDBERRY: Is there some place in our notes here, Mike, that we have addressed outliers? Does the word "outlier" appear?

DR. ERRIGO: In the TORs, no.

DR. SEDBERRY: Okay.

DR. ERRIGO: Does this capture what the SSC is thinking at this point? If you want, you can also make a recommendation that it be prioritized at something, a workshop or a working group or something, or addressed -- A SEDAR address the issue of how to deal with outliers or something, in a systematic way.

DR. REICHERT: I am still reading that sentence.

DR. SEDBERRY: It kind of gets at it, but it's not very --

DR. REICHERT: Because we can't evaluate outliers as in the past for the -- It depends on what you mean by "in the past", because we don't have -- If you're talking about what was during benchmark and standard assessments, that is no longer happening, and operational assessments, and so I would like to get much clearer language there, in terms of how we are using that.

DR. WILLIAMS: I would even add that that phrase "procedures for evaluating outliers" is misleading, because we don't even really have a procedure. I would say the way it's been handled in the past is very ad hoc and situational.

MS. SAULS: We look at the data at the data workshop, but we're not provided with any information on why a calibrated estimate was so much higher than the original, and so states will comment on things like the shore landings for amberjack in Florida in 1990, and we'll say that our experience is that that's an improbable event, and maybe we should not be including shore landings for that species in this assessment, but, as far as -- We never even saw the calibrated estimates until they were released to the public, and so we never had any --

DR. SEDBERRY: Again, that's done on an ad hoc basis, and only when there's a data workshop, and so we're talking about a situation here where there these things show up -- I don't know. We just incorporate the data without examining them for outliers.

DR. REICHERT: I would like to take over Erik's recommendation, in terms of it may be good to come up with a -- I don't know what procedure, a workshop or whatnot, to discuss outliers and how to identify them and how to deal with them and associated issues, because, time and time again, we come back to that, and, time and time again, we say that it's ad hoc, and so I think that's what Erik mentioned earlier. I think it would help if we came up with a recommendation in whatever form that is going to be addressed. As I said, I'm not sure whether that's in the form of a workshop or whether that is something that should occur outside the SSC, but I think this is something that should definitely be occurring.

DR. SEDBERRY: Can the SSC make a recommendation to the SEDAR Committee, or the council, that some process be established to deal with these outliers? Again, what is an outlier and how important they are is, to me, questions that I can't even contemplate.

DR. SCHUELLER: I know, a few years ago, there was a workshop about best practices, and a document came out with a timeline, and I thought that, in that, there was going to be some analytical or quantitative-type group that was going to take up topics as they arose, in order to

address them for the entire SEDAR process, and does that group actually exist, and is this something we could recommend that they take up as a topic?

DR. SEDBERRY: I don't know if that group exists. If they do though, this would be a good topic for them. There might be someone in the room who is on the SEDAR Committee, and I don't know.

MR. CARMICHAEL: There is talk of such groups, and, if I'm thinking of it correctly, we did the data best practices, and there was talk of similar things for other issues. The reality is that there just hasn't been the resources, time, money, people, et cetera, to do it, but I think there is still some interest, and the outliers topic, in particular, is one that we felt is really a region-wide issue that needs to be addressed, and that's one reason it wasn't the primary topic of discussion here at this workshop, because this is one SSC, and this outlier thing is something that needs to be handled consistently across the region.

That's one that definitely, I think, yes, should be handled at a higher level, and that was our recommendation at SEDAR when this first came up, that there should be some regional effort to deal with this, and, as far as getting like standing groups and stuff to do that, that's been a bit tougher, but we do have the ability to bring groups of people together and have them talk about these issues that cross over all of the stock assessments, and I think, if you listen to what Clay has said over the years, that's one of the things that he has thought to be an issue within a research track process, and it's just been really, really slow getting all of this stuff off the ground and getting it moving.

DR. SEDBERRY: Thanks, John. I think what Mike has typed up sounds good to me, and, in light of what John said, we could expand it to make it a national thing, but I think the wording that you have there, Mike, kind of covers our local issues and could be taken to the National SSC as well. Does anybody on the SSC disagree with what Mike has typed there, our recommendation being that SEDAR uses the best practices working group to address a systematic way of identifying and dealing with outliers in the catch data? I know it just kind of kicks the can down the road, but we can't possibly address it at this workshop.

DR. BELCHER: I agree with that statement in the context of what we're talking about with the FES, but I would still argue that goes to every survey-generated data stream that comes through SEDAR. There is going to be -- Marcel has had funny spikes, and I'm going to call them spikes, because the outlier thing is still kind of getting twitchy, but the funky spikes that show up that we all kind of have to figure out how to deal with that information, as to how it came to where it did, or the zeroes that show up and what does really the zero mean, and I think that's a bigger -- The outlier thing needs to be addressed at the higher level, to say it across all data sources and not just the catch.

DR. SEDBERRY: Agreed. Does anybody disagree with that? Okay. Do we want to say something here about TOR 2, something quick? I think what we wanted to say about it -- Let me see. I might even have some notes here.

DR. WILLIAMS: Just so we're clear, when I come rolling in in March with two stock assessments in hand, and somebody says, hey, what about that spike in red porgy in -- I am going to turn around and walk out the door.

DR. ERRIGO: Hopefully, with the intense detail that we got of the species that are undergoing assessments now, we should have a good understanding of where the spikes came from.

DR. SEDBERRY: TOR 2 is about the use of new estimates for unassessed species, and 2A was to compare current ABC values to updated values based on revised estimates and determine if any further information or analysis is needed for the SSC to provide updated ABC recommendations for unassessed stocks, and so Mike has prepared that spreadsheet that makes all those comparisons, and I believe, or we believe, that the ABC working group is going to take up this issue between now and the fall SSC meeting and that we can leave this in their capable hands to come up with something.

DR. ERRIGO: I think what we'll do is we'll have that working group meet via webinar once before the October meeting, to go over some of those graphics, the landings and the estimated ABCs that are in there and look at the time series trends. Then, at the October meeting, we can address the ABCs for unassessed species.

DR. SEDBERRY: Very good, and so the only thing we need to do right now is just make a statement to that effect in our report and in our notes here.

DR. SCHUELLER: I am fine with that. Can you remind us who is on the ABC working group?

DR. SEDBERRY: That's a good question.

DR. REICHERT: And put that list in the notes.

DR. SEDBERRY: That way, there will be many places we can find it. While Mike is looking that up, I will take this opportunity to ask for public comment. This is our second public comment period. We had one at the beginning of the workshop, and we're getting close to the end here, and so I would like to see if there are any members of the public who would like to make a comment at this time.

PUBLIC COMMENT

MR. ESTES: Thank you. My name is Jim Estes, and I'm with the Florida Fish and Wildlife Conservation Commission, and I am the Deputy Director for the Division of Marine Fisheries Management. The first thing is I had a hand in writing all the letters that came from Florida, and I think that you all have done exactly what we wanted you to do, and I thank you.

Also, I would like to thank the Office of Science and Technology for making excellent presentations, including the infinite detail that I enjoyed, although I will tell you that my comments here are probably going to be a little bit less technical than the things I've heard over the last three days, and so expect it.

As an aside, a little bit of an aside, and I will try to hurry my comments, and I know we're getting late, but I have to tell a little bit of an analogy. Thirty years ago, I developed, or helped develop, creel surveys in freshwater, and, as I listened to John describe what you all do and how you do it,

I thought, geez, I think I probably was like living in a cave and eating raw birds, because I didn't know anything about fire, at the same time that the Office of Science and Technology were flying starships. It's very complicated, and I appreciate the work that you do, and I also appreciate something -- I have never been to an SSC meeting, and I really appreciate the seriousness that you all take in the important work that we do. I honestly appreciate it. Either you are very serious or good scientists or very dour, and I'm not sure which, but I think you're serious about what you do.

I was going to talk a little bit about the mismatch between the tool that we have for estimating recreational harvest in Magnuson, but I'm going to skip that. That's something for a beer. The decisions that I think that you made -- Number one, as I said the other day, I am a skeptic. I am firmly convinced now, although I was somewhat before, that the Coastal Household Telephone Survey should be replaced, and I do believe that we have replaced it with something much better. I am completely convinced of that, and I also agree that, because of that, I don't think it would be -- It would be kind of embarrassing to say that we're going to go forward with these stock assessments, but we're going to use the information we know is faulty. I think that would be a little bit hypocritical, and so I appreciate what you did.

I do have a request, one request, and I will make this fairly quick, if it can. It probably won't be met with open arms, but I have a request anyway. I know Genny mentioned that she would like to see some simulations done, and I think that you added that, and so I appreciate that, but I would like to go a little bit further.

I still have questions, and that's the reason that I was a little bit provocative earlier and asked if, more or less, if you all believe that these estimates of effort reflect reality is because -- I know we don't know, but, frankly, I am not confident that they do. I am not confident that they don't, but I will tell you that we have stakeholders that expect that that's what we're doing.

I understand why it's sometimes difficult, or maybe impossible to do, but now we, as managers, we have to make some decisions, based on what we get from you all, and we have to explain those decisions to the stakeholders, and, unless John has lots of time to travel around and make presentations, I think that's going to be difficult for us to do, and so I think something that would resonate with our stakeholders, and it certainly would with the managers, is if we could do a study, or a series of pilots, that would try to look at some independent way of looking at this other than asking the fishermen what do they remember about going fishing, and I would like to take the fishermen out of it.

I know I heard the other day that this has been talked about, and it is very difficult, and I can't fathom what that means, based on the difficulty that I think that you all have done, the work that you have done, to bring us up from MRFSS, even starting MRFSS, to this, and so that's my request, that either the SSC could recommend that or, frankly, from Florida, we are going to try to do that, and I know that money is an issue, but I do believe --

Talking to the other states, I believe that we have a strong enough lobby that, if the states are going to Congress to ask about this, to give -- I would suggest that we give it to the same folks that are doing the other work and that we could find the money to do this, and I'm fairly confident of that. That's all my comments. Thank you.

DISCUSSION OF RESEARCH RECOMMENDATIONS

DR. SEDBERRY: Thanks for your comments, and we appreciate you being here and participating in the workshop, and we particularly appreciate your comments, and so thank you very much. Well, the only thing I have still remaining on my list, and it's probably not the only thing that's still on my list, but it's the one thing that pops out at me, because I don't really want to look at the list, but is some research recommendations, and I don't think we had that as a term of reference, but there's been a couple of things that have popped up in our discussions that, to me, might fit under something called research recommendations, and the SSC, and all the advisory panels for the council and all the committees come up with research recommendations, and it's good to have these things, because they actually sometimes make it into requests for proposals and funding opportunities.

We have heard about doing some simulations that might help us better understand the discrepancies between the surveys, and there might be other things that we want to list as research topics, research recommendations, that NOAA Fisheries or some other agency might want to consider funding in the future, and so, if anybody has any ideas about any of those right now, we could list some of those right here in our notes, while we've got a few minutes left.

MS. SAULS: I think you just got one from Jim Estes.

DR. SEDBERRY: Yes. Thank you.

DR. ERRIGO: By the way, here's the list of the ABC workgroup members, as it currently stands. If anyone else wants to jump on the bandwagon, we're more than happy to have you.

DR. SEDBERRY: Yes, and working group members do not need to be SSC members, and isn't that correct, and so anybody is welcome to join the ABC fun.

DR. DUMAS: I just wanted to remind folks of two items that were mentioned in the last few days that might be research topics. One was I think Fred Serchuk, I think, mentioned exogenous events and how do those affect sampling, like hurricane strikes or red tides and that type of thing and how does that affect the sampling, and that might be something worth investigating.

Another one that came up was the definition of a fishing trip and looking at consistency of the definition of a fishing trip across state data collection efforts and the MRIP and other efforts and looking at how the definition of a fishing trip varies across different data collection programs and does that affect estimates or differences among estimates.

DR. SCHARF: I would agree with those that Chris mentioned and then also echo the comment, the public comment, that we got, because we did have some discussion about asking OST to at least consider some potential study designs where they could groundtruth some of the FES estimates, even if it's on a very small, limited spatial and temporal extent, but to try to get some ground-truthing to some of the effort estimates, whatever technologies might be available, but just to at least consider some potential study designs.

UNIDENTIFIED: Mr. Chairman, just a quick question, I guess, and I don't think that was the public comment. I think the public comment was a recommendation to seek an alternative methodology for effort.

MR. ESTES: I'm sorry if that's what you took out of that, and I probably was not clear. No, I want to have some ground-truthing, so that we can have -- I can be more confident, and certainly I don't want to do -- I don't want to throw away all the work that's been done and do satellite imagery to estimate. That is not my intent. My intent was to do some ground-truthing, and it can be done on whatever scale you all think is appropriate.

DR. ERRIGO: Chris, I didn't get the first research recommendation that you had. I got the explore the definition of a trip, and I didn't get the --

DR. DUMAS: The first one was look at how exogenous events affect the survey sampling, exogenous events like hurricanes and red tides.

DR. SEDBERRY: Okay. Are there any other notes or recommendations or consensus statements that we need to review or edit or add to?

DR. SHERTZER: One other research recommendation that I think Erik brought up was to take a data-rich species and scale back the data in ways to mimic rare-event species, to examine features like bias and precision.

DR. SEDBERRY: I think, somewhere in our notes, we have in fact stated that the SSC is satisfied and agrees that the MRIP is the best scientific information available, and so we don't need to come up with alternatives for it, and we probably would need another day for that, I think. I don't know what else we need to cover, but I just want to check with Marcel and Mike and make sure we've kind of covered everything that we need to cover, and then, the SSC and all of our invited experts, do you all have anything to add here at the last minute, before we adjourn?

DR. DUMAS: One more possible research recommendation, because we spent some time discussing seasonal addresses, seasonal households, and how they are not currently sampled, and, in some areas, a large proportion of the households might be seasonal, and so we might want to look at how omitting seasonal households affects the sampling and affects the estimates and are there some areas where those should be included and what would be the implications of including those.

DR. SEDBERRY: All right. Thanks. Anyway, I would like to thank the members of the steering committee that helped organize this, and I would like to thank the SSC for all of your suggestions along the way and for participating and then all of the invited experts that have helped out in so many ways. I mean, we just covered a tremendous amount of material and heard a lot of detail that I really think we needed to hear.

I know it was a lot, and probably more detail than some of us wanted to hear, but it was really -- We really needed to hear that, and we really needed to understand the surveys and the calibrations a lot better than we did, and I think we do now, and so I want to thank everybody for coming and participating and providing your brains and your materials and everything that you provided to help us move this along.

As I have mentioned a few times during the meeting, I have to give a report on this at the next meeting of the South Atlantic Council, which is scheduled for too close to now, three weeks, here in Charleston. Anyway, we have kind of a short fuse on this report, and I know some of you have already submitted your notes to me, and so, if any of you have taken any notes, those of you that did take notes, if you scribbled down anything, if you've got things written on a cocktail napkin, anything that might be helpful to Mike and I as we summarize the meeting and prepare a report and presentation for the South Atlantic Council meeting, that will be really helpful, and so send that to me just as soon as you can.

We will compile it by the end of next week and get it back out to you to edit and review and make sure that everything that you wanted said is in it and that there's no errors and other problems with it, and then, at some point down the line from that, I will prepare a PowerPoint presentation that I will give to the council, and I will send that out for review too, and it probably won't be in time for the briefing book, but we'll have it. I think that's it. Is there anything else that I need to think about, Marcel?

DR. REICHERT: No, and I was just looking at the schedule. The end of next week is the 30th, and then I assume you will get that to us by the 5th or the 6th, and that gives us a little time to make some edits, and so the council meeting is the 16th, I believe, right?

DR. SEDBERRY: It is the 16th, yes.

DR. REICHERT: Okay, and so that would mean that you and Mike will have the report available about a week before the council meeting, and I just want to make sure that we're all on the same page, in terms of when you guys need the information, and so the first round is the 30th, and then the second round you need to have back by the 6th?

DR. SEDBERRY: Yes.

DR. REICHERT: Okay.

DR. SEDBERRY: Again, any notes you might have, get them to me today or tomorrow, and it would be great if you could do that, and some of you have already done it, and so thanks for that. Anything else for the good of the cause? All right. Safe travels, and we will see you when we do this again in October, if not before then. Thanks.

(Whereupon, the meeting adjourned on August 21, 2019.)

- - -

Certified By: _____ Date: _____

Transcribed By:
Amanda Thomas
September 16, 2019

Aug 19, 2019

SOUTH ATLANTIC FISHERY MANAGEMENT COUNCIL
SCIENTIFIC AND STATISTICAL COMMITTEE

✓Dr. George Sedberry, **Chair**
6 Sorghum Lane
Savannah, GA 31411
George.sedberry@gmail.com
6/10*, 6/13*, 6/16*

Dr. Robert Ahrens, **Vice Chair**
University of Florida
135 Newins-Ziegler Hall
Gainesville, FL 32611
(352)273-3630 (ph)
rahrens@ufl.edu
6/16*

Dr. Luiz Barbieri
FL FWCC/FMRI
100 Eighth Avenue SE
St. Petersburg, FL 33701
(727)896-8626 ext. 4920 (ph)
luiz.barbieri@fwc.state.fl.us
3/08*, 6/10*, 6/13*, 6/16*

✓Dr. Carolyn Belcher
GA DNR-Coastal Resources Division
One Conservation Way Suite 300
Brunswick, GA 31520
(912)264-7218 (ph); (912)262-3143
Carolyn_Belcher@dnr.state.ga.us
12/01*, 3/08*,6/09*,6/12*, 6/15*

✓Dr. Jeffrey Buckel
North Carolina State University-Center for Marine
Science and Technology
Department of Zoology
303 College Circle
Morehead City, NC 28557
(252)222-6341(ph); (252)222-6311(f)
jeffrey_buckel@ncsu.edu
9/05*,3/08*,6/10*, 6/13*,6/16*

✓Dr. Scott Crosson
NMFS SEFSC
75 Virginia Beach Drive
Miami, FL 33149
(305)361-4468
Scott.Crosson@noaa.gov
6/09*, 6/17*

✓Dr. Chris Dumas
University of NC-Wilmington
Dept. of Environmental Studies
601 South College Rd.
Wilmington, NC 28403
(910)962-4026(ph)
dumasc@uncw.edu
6/18*

✓Dr. Churchill Grimes
211 Moriah Creek Road
Crawfordville, FL 32327
(831)713-6543
churchill.grimes@gmail.com
6/10*, 6/13*, 6/16*

Dr. Eric Johnson
Department of Biology University of North Florida
1 UNF Drive
Jacksonville, FL 32224
(904)620-5764
eric.johnson@unf.edu
6/10*, 6/11*, 6/14*, 6/17*

✓Anne Lange
1493 Diamond Blvd
Mt Pleasant, SC 29466
(843)971-0628 (ph)
AMLange@aol.com
3/08*, 6/11*,6/14*, 6/17*

✓Dr. Yan Li
NC Division of Marine Fisheries
Dept. of Environmental Quality
P.O. Box 769
3441 Arendell St.
Morehead City, NC 28557
(540)808-8373 (ph)
yan.li@ncdenr.gov
6/18*

✓Dr. Genny Nesslage
Chesapeake Biological Laboratory
P.O. Box 38
146 Williams Street
Solomons, MD 20688
(410)326-7223 (ph)
nesslage@umces.edu
6/16*

SOUTH ATLANTIC FISHERY MANAGEMENT COUNCIL
SCIENTIFIC AND STATISTICAL COMMITTEE continued

✓ Dr. Marcel Reichert
SC DNR/Marine Resources Division
PO Box 12559
217 Ft. Johnson Road
Charleston, SC 29412
(843)953-5778 (ph)
ReichertM@dnr.sc.gov
3/08*, 6/11, 6/14*, 6/17*

✓ Dr. Frederick Scharf
University of North Carolina Wilmington
Department of Biology and Marine Biology
601 South College Rd
Wilmington, NC 28403
scharff@uncw.edu
6/17*

Dr. Amy Schueller
NOAA Fisheries
SEFSC Beaufort Laboratory
101 Pivers Island Road
Beaufort, NC 28516
(252)838-0815 (ph)
Amy.Schueller@noaa.gov
6/14*, 6/17*

Dr. Fred Serchuk
Casa Bonita 1,
26000 Hickory Blvd. Unit 704
Bonita Springs, FL 34134
(774)392-4975 (ph)
fred.serchuk@gmail.com
6/15*

Dr. Alexei Sharov
Maryland Dept. of Natural Resources
508 Taylor Avenue
Annapolis, MD 21401
(410)260-8288 (ph)
ASHAROV@dnr.state.md.us
6/15*

✓ Dr. Tracy Yandle
Emory University-Mathematics and Science Center
Dept. of Environmental Studies
400 Dowman Dr.
Atlanta, GA 30322
(404)727-6314 (ph); (404)727-4448 (f)
tyandle@emory.edu
6/11*, 6/14*, 6/17*

OPEN SEAT

Drew Cathey
Steve Poland
Mel Bell
Brad Floyd
Chris Dumas
Chris Wilson
Dr. Wilson Laney
Dave VanVoorhees
Michael Larkin
Beverly Sauls
Shep Grimes
Dawn Franco
John Foster
Lauren Dolinger Few
Rob Andrews
Vivian Matter
Ryan Rindone
Kyle Dettloff
Chris McDonough
Richard Cody
David Deltor
Kyle Shertzer
Jim Estes

SOUTH ATLANTIC FISHERY MANAGEMENT COUNCIL
COUNCIL STAFF

Executive Director
✓ Gregg T. Waugh
gregg.waugh@safmc.net

Deputy Director – Science & Statistics
✓ John Carmichael
john.carmichael@safmc.net

Deputy Director - Management
Dr. Brian Cheuvront
brian.cheuvront@safmc.net

Fishery Scientist
✓ Myra Brouwer
myra.brouwer@safmc.net

Senior Fishery Biologist
Roger Pugliese
roger.pugliese@safmc.net

Citizen Science Program Manager
✓ Julia Byrd
julia.byrd@safmc.net

Outreach Specialist
Cameron Rhodes
cameron.rhodes@safmc.net

Admin. Secretary/Travel Coordinator
Cindy Chaya
cindy.chaya@safmc.net

Staff Accountant
Suzanna Thomas
suzanna.thomas@safmc.net

Fishery Scientist
Dr. Chip Collier
chip.collier@safmc.net

Fisheries Social Scientist
✓ Christina Wiegand
christina.wiegand@safmc.net

Purchasing & Grants
✓ Cierra Graham
cierra.graham@safmc.net

Fishery Biologist
✓ Dr. Mike Errigo
mike.errigo@safmc.net

SEDAR Coordinators

Dr. Julie Neer
julie.neer@safmc.net

Fishery Economist
John Hadley
john.hadley@safmc.net

✓ Kathleen Howington
kathleen.howington@safmc.net

Public Information Officer
Kim Iverson
kim.iverson@safmc.net

Administrative Officer
Kelly Klasnick
kelly.klasnick@safmc.net

SOUTH ATLANTIC FISHERY MANAGEMENT COUNCIL
2019 COUNCIL MEMBERS

Jessica McCawley, **Chair**
Florida Fish and Wildlife Conservation Commission
2590 Executive Center Circle E. Suite 201
Tallahassee, FL 32301
(850)487-0554 (ph); (850)487-4847 (f)
Jessica.mccawley@myfwc.com

✓Mel Bell, **Vice Chair**
SCDNR-Marine Resources Division
P.O. Box 12559
217 Ft. Johnson Road
Charleston, SC 29422
(843)953-9007 (ph); (843)953-9159 (fax)
bellm@dnr.sc.gov

Robert Beal
Executive Director
Atlantic States Marine Fisheries Commission
1050 N. Highland St. Suite 200 A-N
Arlington, VA 22201
(703)842-0740 (ph); (703)842-0741 (f)
rbeal@asmfc.org

Anna Beckwith
1907 Paulette Road
Morehead City, NC 28557
(252)671-3474 (ph)
AnnaBarriosBeckwith@gmail.com

Chester Brewer
250 Australian Ave. South Suite 1400
West Palm Beach, FL 33408
(561)655-4777 (ph)
wcbsafmc@gmail.com

Dr. Kyle Christiansen
150 Cedar St.
Richmond Hill, GA 31324
(912)756-7560 (ph)
christiansensafmc@gmail.com

Chris Conklin
P.O. Box 972
Murrells Inlet, SC 29576
(843)543-3833
conklinsafmc@gmail.com

Dr. Roy Crabtree
Regional Administrator
NOAA Fisheries, Southeast Region
263 13th Avenue South
St. Petersburg, FL 33701
(727)824-5301 (ph); (727)824-5320 (f)
roy.crabtree@noaa.gov

Tim Griner
4446 Woodlark Lane
Charlotte, NC 28211
(980)722-0918 (ph)
timgrinersafmc@gmail.com

Doug Haymans
Coastal Resources Division
GA Dept. of Natural Resources
One Conservation Way Suite 300
Brunswick, GA 31520
(912)264-7218 (ph); (912)262-2318 (f)
haymanssafmc@gmail.com

U.S. Fish and Wildlife Service
Representative
TBD

LCDR Jeremy Montes
U.S. Coast Guard
909 SE 1st Ave.
Miami, FL 33131
(305)415-6788(ph); (305)710-4569(c)
Jeremy.J.Montes@uscg.mil

✓Stephen Poland
NC Division of Marine Fisheries
PO Box 769
3441 Arendell Street
Morehead City, NC 28557
(252)808-8011 (direct); (252)726-7021 (main)
Steve.Poland@ncdenr.gov

Art Sapp
2270 NE 25th St.
Lighthouse Pointe, FL 33064
(954)444-0820 (ph)
artsappsafmc@gmail.com

SOUTH ATLANTIC FISHERY MANAGEMENT COUNCIL
2019 COUNCIL MEMBERS continued

Deirdre Warner-Kramer
Office of Marine Conservation OES/OMC
2201 C Street, N.W. Department of State, Room 5806
Washington, DC 20520
(202)647-3228 (ph)
Warner-KramerDM@state.gov

David Whitaker
720 Tennent Street
Charleston, SC 29412
(843)953-9392
david.whitakersafmc@gmail.com

Spud Woodward
860 Buck Swamp Road
Brunswick, GA 31523
(912)258-8970 (ph)
swoodwardsafmc@gmail.com

August 20, 2019

SOUTH ATLANTIC FISHERY MANAGEMENT COUNCIL
SCIENTIFIC AND STATISTICAL COMMITTEE

Dr. George Sedberry, **Chair**
6 Sorghum Lane
Savannah, GA 31411
George.sedberry@gmail.com
6/10*, 6/13*, 6/16*

Dr. Robert Ahrens, **Vice Chair**
University of Florida
135 Newins-Ziegler Hall
Gainesville, FL 32611
(352)273-3630 (ph)
rahrens@ufl.edu
6/16*

Dr. Luiz Barbieri
FL FWCC/FMRI
100 Eighth Avenue SE
St. Petersburg, FL 33701
(727)896-8626 ext. 4920 (ph)
luiz.barbieri@fwc.state.fl.us
3/08*, 6/10*, 6/13*, 6/16*

Dr. Carolyn Belcher
GA DNR-Coastal Resources Division
One Conservation Way Suite 300
Brunswick, GA 31520
(912)264-7218 (ph); (912)262-3143
Carolyn_Belcher@dnr.state.ga.us
12/01*, 3/08*,6/09*,6/12*, 6/15*

Dr. Jeffrey Buckel
North Carolina State University-Center for Marine
Science and Technology
Department of Zoology
303 College Circle
Morehead City, NC 28557
(252)222-6341(ph); (252)222-6311(f)
jeffrey_buckel@ncsu.edu
9/05*,3/08*,6/10*, 6/13*,6/16*

Dr. Scott Crosson
NMFS SEFSC
75 Virginia Beach Drive
Miami, FL 33149
(305)361-4468
Scott.Crosson@noaa.gov
6/09*, 6/17*

Dr. Chris Dumas
University of NC-Wilmington
Dept. of Environmental Studies
601 South College Rd.
Wilmington, NC 28403
(910)962-4026(ph)
dumasc@uncw.edu
6/18*

Dr. Churchill Grimes
211 Moriah Creek Road
Crawfordville, FL 32327
(831)713-6543
churchill.grimes@gmail.com
6/10*, 6/13*, 6/16*

Dr. Eric Johnson
Department of Biology University of North Florida
1 UNF Drive
Jacksonville, FL 32224
(904)620-5764
eric.johnson@unf.edu
6/10*, 6/11*, 6/14*, 6/17*

Anne Lange
1493 Diamond Blvd
Mt Pleasant, SC 29466
(843)971-0628 (ph)
AMLange@aol.com
3/08*, 6/11*,6/14*, 6/17*

Dr. Yan Li
NC Division of Marine Fisheries
Dept. of Environmental Quality
P.O. Box 769
3441 Arendell St.
Morehead City, NC 28557
(540)808-8373 (ph)
yan.li@ncdenr.gov
6/18*

Dr. Genny Nesslage
Chesapeake Biological Laboratory
P.O. Box 38
146 Williams Street
Solomons, MD 20688
(410)326-7223 (ph)
nesslage@umces.edu
6/16*

11

SOUTH ATLANTIC FISHERY MANAGEMENT COUNCIL
SCIENTIFIC AND STATISTICAL COMMITTEE continued

✓Dr. Marcel Reichert
SC DNR/Marine Resources Division
PO Box 12559
217 Ft. Johnson Road
Charleston, SC 29412
(843)953-5778 (ph)
ReichertM@dnr.sc.gov
3/08*, 6/11, 6/14*, 6/17*

✓Dr. Frederick Scharf
University of North Carolina Wilmington
Department of Biology and Marine Biology
601 South College Rd
Wilmington, NC 28403
scharff@uncw.edu
6/17*

✓Dr. Amy Schueller
NOAA Fisheries
SEFSC Beaufort Laboratory
101 Pivers Island Road
Beaufort, NC 28516
(252)838-0815 (ph)
Amy.Schueller@noaa.gov
6/14*, 6/17*

Dr. Fred Serchuk
Casa Bonita 1,
26000 Hickory Blvd. Unit 704
Bonita Springs, FL 34134
(774)392-4975 (ph)
fred.serchuk@gmail.com
6/15*

Dr. Alexei Sharov
Maryland Dept. of Natural Resources
508 Taylor Avenue
Annapolis, MD 21401
(410)260-8288 (ph)
ASHAROV@dnr.state.md.us
6/15*

✓Dr. Tracy Yandle
Emory University-Mathematics and Science Center
Dept. of Environmental Studies
400 Dowman Dr.
Atlanta, GA 30322
(404)727-6314 (ph); (404)727-4448 (f)
tyandle@emory.edu
6/11*, 6/14*, 6/17*

OPEN SEAT

Brad Floyd

Chris Wilson

Drew Cathey

Beverly Sauls

Dr. Dave Van Voorhees

Vivian Matter

Rob Andrews

Dawn Defranco

Kathy Knowlton

Shep Grimes

Michael Larkin

John Foster

Amy Dukes

Dr. Katherine Drew

Chris McDonough

Dr. Erik Williams

Dr. Kyle Shertzer

Ryan Rindone

Kyle Dettloff

David Deltor

Lauren Dolinger Few

Jim Estes

Aug 20, 2019

SOUTH ATLANTIC FISHERY MANAGEMENT COUNCIL
2019 COUNCIL MEMBERS

Jessica McCawley, **Chair**
Florida Fish and Wildlife Conservation Commission
2590 Executive Center Circle E. Suite 201
Tallahassee, FL 32301
(850)487-0554 (ph); (850)487-4847 (f)
Jessica.mccawley@myfwc.com

✓Mel Bell, **Vice Chair**
SCDNR-Marine Resources Division
P.O. Box 12559
217 Ft. Johnson Road
Charleston, SC 29422
(843)953-9007 (ph); (843)953-9159 (fax)
bellm@dnr.sc.gov

Robert Beal
Executive Director
Atlantic States Marine Fisheries Commission
1050 N. Highland St. Suite 200 A-N
Arlington, VA 22201
(703)842-0740 (ph); (703)842-0741 (f)
rbeal@asmfc.org

Anna Beckwith
1907 Paulette Road
Morehead City, NC 28557
(252)671-3474 (ph)
AnnaBarriosBeckwith@gmail.com

Chester Brewer
250 Australian Ave. South Suite 1400
West Palm Beach, FL 33408
(561)655-4777 (ph)
wcbsafmc@gmail.com

Dr. Kyle Christiansen
150 Cedar St.
Richmond Hill, GA 31324
(912)756-7560 (ph)
christiansensafmc@gmail.com

Chris Conklin
P.O. Box 972
Murrells Inlet, SC 29576
(843)543-3833
conklinsafmc@gmail.com

Dr. Roy Crabtree
Regional Administrator
NOAA Fisheries, Southeast Region
263 13th Avenue South
St. Petersburg, FL 33701
(727)824-5301 (ph); (727)824-5320 (f)
roy.crabtree@noaa.gov

Tim Griner
4446 Woodlark Lane
Charlotte, NC 28211
(980)722-0918 (ph)
timgrinersafmc@gmail.com

Doug Haymans
Coastal Resources Division
GA Dept. of Natural Resources
One Conservation Way Suite 300
Brunswick, GA 31520
(912)264-7218 (ph); (912)262-2318 (f)
haymanssafmc@gmail.com

U.S. Fish and Wildlife Service
Representative
TBD

LCDR Jeremy Montes
U.S. Coast Guard
909 SE 1st Ave.
Miami, FL 33131
(305)415-6788(ph); (305)710-4569(c)
Jeremy.J.Montes@uscg.mil

✓Stephen Poland
NC Division of Marine Fisheries
PO Box 769
3441 Arendell Street
Morehead City, NC 28557
(252)808-8011 (direct); (252)726-7021 (main)
Steve.Poland@ncdenr.gov

Art Sapp
2270 NE 25th St.
Lighthouse Pointe, FL 33064
(954)444-0820 (ph)
artsappsafmc@gmail.com

5

SOUTH ATLANTIC FISHERY MANAGEMENT COUNCIL
2019 COUNCIL MEMBERS continued

Deirdre Warner-Kramer
Office of Marine Conservation OES/OMC
2201 C Street, N.W. Department of State, Room 5806
Washington, DC 20520
(202)647-3228 (ph)
Warner-KramerDM@state.gov

David Whitaker
720 Tennent Street
Charleston, SC 29412
(843)953-9392
david.whitakersafmc@gmail.com

Spud Woodward
860 Buck Swamp Road
Brunswick, GA 31523
(912)258-8970 (ph)
swoodwardsafmc@gmail.com

SOUTH ATLANTIC FISHERY MANAGEMENT COUNCIL
COUNCIL STAFF

Executive Director
✓ Gregg T. Waugh
gregg.waugh@safmc.net

Deputy Director – Science & Statistics
✓ John Carmichael
john.carmichael@safmc.net

Deputy Director - Management
Dr. Brian Cheuvront
brian.cheuvront@safmc.net

Fishery Scientist
✓ Myra Brouwer
myra.brouwer@safmc.net

Senior Fishery Biologist
Roger Pugliese
roger.pugliese@safmc.net

Citizen Science Program Manager
✓ Julia Byrd
julia.byrd@safmc.net

Outreach Specialist
Cameron Rhodes
cameron.rhodes@safmc.net

Admin. Secretary/Travel Coordinator
Cindy Chaya
cindy.chaya@safmc.net

Staff Accountant
Suzanna Thomas
suzanna.thomas@safmc.net

Fishery Scientist
Dr. Chip Collier
chip.collier@safmc.net

Fisheries Social Scientist
✓ Christina Wiegand
christina.wiegand@safmc.net

Purchasing & Grants
✓ Cierra Graham
cierra.graham@safmc.net

Fishery Biologist
✓ Dr. Mike Errigo
mike.errigo@safmc.net

SEDAR Coordinators

Dr. Julie Neer
julie.neer@safmc.net

Fishery Economist
John Hadley
john.hadley@safmc.net

✓ Kathleen Howington
kathleen.howington@safmc.net

Public Information Officer
Kim Iverson
kim.iverson@safmc.net

Administrative Officer
Kelly Klasnick
kelly.klasnick@safmc.net

August 21, 2019

SOUTH ATLANTIC FISHERY MANAGEMENT COUNCIL
SCIENTIFIC AND STATISTICAL COMMITTEE

√Dr. George Sedberry, **Chair**
6 Sorghum Lane
Savannah, GA 31411
George.sedberry@gmail.com
6/10*, 6/13*, 6/16*

Dr. Robert Ahrens, **Vice Chair**
University of Florida
135 Newins-Ziegler Hall
Gainesville, FL 32611
(352)273-3630 (ph)
rahrens@ufl.edu
6/16*

Dr. Luiz Barbieri
FL FWCC/FMRI
100 Eighth Avenue SE
St. Petersburg, FL 33701
(727)896-8626 ext. 4920 (ph)
luiz.barbieri@fwc.state.fl.us
3/08*, 6/10*, 6/13*, 6/16*

√ Dr. Carolyn Belcher
GA DNR-Coastal Resources Division
One Conservation Way Suite 300
Brunswick, GA 31520
(912)264-7218 (ph); (912)262-3143
Carolyn_Belcher@dnr.state.ga.us
12/01*, 3/08*,6/09*,6/12*, 6/15*

√Dr. Jeffrey Buckel
North Carolina State University-Center for Marine
Science and Technology
Department of Zoology
303 College Circle
Morehead City, NC 28557
(252)222-6341(ph); (252)222-6311(f)
jeffrey_buckel@ncsu.edu
9/05*,3/08*,6/10*, 6/13*,6/16*

√ Dr. Scott Crosson
NMFS SEFSC
75 Virginia Beach Drive
Miami, FL 33149
(305)361-4468
Scott.Crosson@noaa.gov
6/09*, 6/17*

√Dr. Chris Dumas
University of NC-Wilmington
Dept. of Environmental Studies
601 South College Rd.
Wilmington, NC 28403
(910)962-4026(ph)
dumasc@uncw.edu
6/18*

√Dr. Churchill Grimes
211 Moriah Creek Road
Crawfordville, FL 32327
(831)713-6543
churchill.grimes@gmail.com
6/10*, 6/13*, 6/16*

Dr. Eric Johnson
Department of Biology University of North Florida
1 UNF Drive
Jacksonville, FL 32224
(904)620-5764
eric.johnson@unf.edu
6/10*, 6/11*, 6/14*, 6/17*

✓Anne Lange
1493 Diamond Blvd
Mt Pleasant, SC 29466
(843)971-0628 (ph)
AMLange@aol.com
3/08*, 6/11*,6/14*, 6/17*

√Dr. Yan Li
NC Division of Marine Fisheries
Dept. of Environmental Quality
P.O. Box 769
3441 Arendell St.
Morehead City, NC 28557
(540)808-8373 (ph)
yan.li@ncdenr.gov
6/18*

√Dr. Genny Nesslage
Chesapeake Biological Laboratory
P.O. Box 38
146 Williams Street
Solomons, MD 20688
(410)326-7223 (ph)
nesslage@umces.edu
6/16*

SOUTH ATLANTIC FISHERY MANAGEMENT COUNCIL
SCIENTIFIC AND STATISTICAL COMMITTEE continued

✓Dr. Marcel Reichert
SC DNR/Marine Resources Division
PO Box 12559
217 Ft. Johnson Road
Charleston, SC 29412
(843)953-5778 (ph)
ReichertM@dnr.sc.gov
3/08*, 6/11, 6/14*, 6/17*

✓Dr. Frederick Scharf
University of North Carolina Wilmington
Department of Biology and Marine Biology
601 South College Rd
Wilmington, NC 28403
scharff@uncw.edu
6/17*

✓ Dr. Amy Schueller
NOAA Fisheries
SEFSC Beaufort Laboratory
101 Pivers Island Road
Beaufort, NC 28516
(252)838-0815 (ph)
Amy.Schueller@noaa.gov
6/14*, 6/17*

Dr. Fred Serchuk
Casa Bonita 1,
26000 Hickory Blvd. Unit 704
Bonita Springs, FL 34134
(774)392-4975 (ph)
fred.serchuk@gmail.com
6/15*

Dr. Alexei Sharov
Maryland Dept. of Natural Resources
508 Taylor Avenue
Annapolis, MD 21401
(410)260-8288 (ph)
ASHAROV@dnr.state.md.us
6/15*

✓Dr. Tracy Yandle
Emory University-Mathematics and Science Center
Dept. of Environmental Studies
400 Dowman Dr.
Atlanta, GA 30322
(404)727-6314 (ph); (404)727-4448 (f)
tyandle@emory.edu
6/11*, 6/14*, 6/17*

OPEN SEAT

Michael Larkin
Dr. Wilson Laney
Katherine Drew
John Foster
Dawn Franco
Lauren Dolinger Few
Kathy Knowlton
Dr. Erik Williams
Shep Grimes
Chris McDonough
Brad Floyd
Ryan Rindone
Beverly Sauls
Chris Wilson
Dr. Richard Cody
Dr. Kyle Shertzer
Kyle Dethloff
Drew Cathey
Vivian Matter
David Deltor
Dr. Dave Van Voorhees
Jim Estes
Amy Dukes

12

SOUTH ATLANTIC FISHERY MANAGEMENT COUNCIL
2019 COUNCIL MEMBERS

Jessica McCawley, **Chair**
Florida Fish and Wildlife Conservation Commission
2590 Executive Center Circle E. Suite 201
Tallahassee, FL 32301
(850)487-0554 (ph); (850)487-4847 (f)
Jessica.mccawley@myfwc.com

✓Mel Bell, **Vice Chair**
SCDNR-Marine Resources Division
P.O. Box 12559
217 Ft. Johnson Road
Charleston, SC 29422
(843)953-9007 (ph); (843)953-9159 (fax)
bellm@dnr.sc.gov

Robert Beal
Executive Director
Atlantic States Marine Fisheries Commission
1050 N. Highland St. Suite 200 A-N
Arlington, VA 22201
(703)842-0740 (ph); (703)842-0741 (f)
rbeal@asmfc.org

Anna Beckwith
1907 Paulette Road
Morehead City, NC 28557
(252)671-3474 (ph)
AnnaBarriosBeckwith@gmail.com

Chester Brewer
250 Australian Ave. South Suite 1400
West Palm Beach, FL 33408
(561)655-4777 (ph)
wcbsafmc@gmail.com

Dr. Kyle Christiansen
150 Cedar St.
Richmond Hill, GA 31324
(912)756-7560 (ph)
christiansensafmc@gmail.com

Chris Conklin
P.O. Box 972
Murrells Inlet, SC 29576
(843)543-3833
conklinsafmc@gmail.com

Dr. Roy Crabtree
Regional Administrator
NOAA Fisheries, Southeast Region
263 13th Avenue South
St. Petersburg, FL 33701
(727)824-5301 (ph); (727)824-5320 (f)
roy.crabtree@noaa.gov

Tim Griner
4446 Woodlark Lane
Charlotte, NC 28211
(980)722-0918 (ph)
timgrinersafmc@gmail.com

Doug Haymans
Coastal Resources Division
GA Dept. of Natural Resources
One Conservation Way Suite 300
Brunswick, GA 31520
(912)264-7218 (ph); (912)262-2318 (f)
haymanssafmc@gmail.com

U.S. Fish and Wildlife Service
Representative
TBD

LCDR Jeremy Montes
U.S. Coast Guard
909 SE 1st Ave.
Miami, FL 33131
(305)415-6788(ph); (305)710-4569(c)
Jeremy.J.Montes@uscg.mil

✓Stephen Poland
NC Division of Marine Fisheries
PO Box 769
3441 Arendell Street
Morehead City, NC 28557
(252)808-8011 (direct); (252)726-7021 (main)
Steve.Poland@ncdenr.gov

Art Sapp
2270 NE 25th St.
Lighthouse Pointe, FL 33064
(954)444-0820 (ph)
artsappsafmc@gmail.com

SOUTH ATLANTIC FISHERY MANAGEMENT COUNCIL
2019 COUNCIL MEMBERS continued

Deirdre Warner-Kramer
Office of Marine Conservation OES/OMC
2201 C Street, N.W. Department of State, Room 5806
Washington, DC 20520
(202)647-3228 (ph)
Warner-KramerDM@state.gov

David Whitaker
720 Tennent Street
Charleston, SC 29412
(843)953-9392
david.whitakersafmc@gmail.com

Spud Woodward
860 Buck Swamp Road
Brunswick, GA 31523
(912)258-8970 (ph)
swoodwardsafmc@gmail.com

SOUTH ATLANTIC FISHERY MANAGEMENT COUNCIL
COUNCIL STAFF

Executive Director
✓Gregg T. Waugh
gregg.waugh@safmc.net

Deputy Director – Science & Statistics
✓John Carmichael
john.carmichael@safmc.net

Deputy Director - Management
Dr. Brian Cheuvront
brian.cheuvront@safmc.net

Fishery Scientist
Myra Brouwer
myra.brouwer@safmc.net

Citizen Science Program Manager
✓Julia Byrd
julia.byrd@safmc.net

Admin. Secretary/Travel Coordinator
Cindy Chaya
cindy.chaya@safmc.net

Fishery Scientist
✓Dr. Chip Collier
chip.collier@safmc.net

Purchasing & Grants
✓Cierra Graham
cierra.graham@safmc.net

Fishery Biologist
✓Dr. Mike Errigo
mike.errigo@safmc.net

Fishery Economist
✓John Hadley
john.hadley@safmc.net

Public Information Officer
Kim Iverson
kim.iverson@safmc.net

Administrative Officer
Kelly Klasnick
kelly.klasnick@safmc.net

Senior Fishery Biologist
Roger Pugliese
roger.pugliese@safmc.net

Outreach Specialist
Cameron Rhodes
cameron.rhodes@safmc.net

Staff Accountant
Suzanna Thomas
suzanna.thomas@safmc.net

Fisheries Social Scientist
Christina Wiegand
christina.wiegand@safmc.net

SEDAR Coordinators

Dr. Julie Neer
julie.neer@safmc.net

✓Kathleen Howington
kathleen.howington@safmc.net

SCIENTIFIC AND STATISTICAL COMMITTEE: MRIP WORKSHOP

Town and Country Inn

Charleston, South Carolina

19-Aug-19

Last Name	First Name
Griffin	Aimee
Bianchi	Alan
DiJohnson	Alex
Sharov	Alexei
Tong	Amanda
Markwith	Anne
Wood	Anthony
Salmon	Brandi
Bruger	Catherine
Hanson	Chad
collier	chip
wilson	chris
Graham	Cierra
Myers	craig

Glasgow	Dawn
Lazarre	Dominique
Addis	Dustin
Burgess	Erika
Helies	Frank
Serchuk	Fred
Shepherd	Gary
Exley	Gary
White	Geoff
Sedberry	George
power	greg
Bray	Gregg
Laks	Ira
Pulver	Jeff
Bissette	Jesse
McGovern	John
Hadley	John
montañez	josé
Simpson	Julie
Pianka	Karen
Aman	Kevin

Iverson	Kim
Lee	Laura
Barbieri	Luiz
Terceiro	Mark
Vara	Mary
Seward	McLean
Schmidtke	Michael
Travis	Michael
Jepson	Michael
Celestino	Michael
Drexler	Michael
Norberg	Mike
BROUWER	MYRA
Mehta	Nikhil
Nitschke	Paul
DeVictor	Rick
Ahrens	Robert
Muler	Robert
Hudson	Rusty
Lovell	Sabrina
Stephenson	Sarah

Steinback	Scott
Allen	Shanae
McInerny	Stephanie
poland	steve
Teears	Thom
sminkey	thomas
Cross	Tiffanie
Williams	Travis
Scheffel	Trevor

SCIENTIFIC AND STATISTICAL COMMITTEE: MRIP WORKSHOP
Town and Country Inn
Charleston, South Carolina
20-Aug-19

Last Name	First Name
Addis	Dustin
Ahrens	Robert
Allen	Shanae
BROUWER	MYRA
Barbieri	Luiz
Bianchi	Alan
Bissette	Jesse
Bray	Gregg
Bruger	Catherine
Byrd	Julia
Celestino	Michael
Chagaris	David
Cheshire	Rob
Cross	Tiffanie
DeVictor	Rick
Drexler	Michael
Ellis	steven
Fitzpatrick	Eric
Glasgow	Dawn
Graham	Cierra
Griffin	Aimee
Hadley	John
Hanson	Chad
Helies	Frank
Howington	Kathleen
Hudson	Rusty
Iverson	Kim
Jepson	Michael
Johansen	Lucas
Joseph	Munyandorero
Laks	Ira
Lamberte	Tony
Lazarre	Dominique
Lee	Laura
Lovell	Sabrina
Markwith	Anne
McGovern	John
Mehta	Nikhil

Muler	Robert
Pianka	Karen
Pickens	Christopher
Pulver	Jeff
Records	David
Salmon	Brandi
Schmidtke	Michael
Sedberry	George
Serchuk	Fred
Seward	McLean
Sharov	Alexei
Shepherd	Gary
Siegfried	Katie
Simpson	Julie
Smart	Tracey
Steinback	Scott
Terceiro	Mark
Tong	Amanda
Travis	Michael
Vara	Mary
Waugh	Gregg
White	geoff
White	Geoff
Wiegand	Christina
Williams	Travis
Wood	Anthony
collier	chip
montañez	josé
poland	steve
sminkey	thomas
Exley	Gary
Krikstan	Catherine
McInerny	Stephanie

SCIENTIFIC AND STATISTICAL COMMITTEE: MRIP WORKSHOP

Town and Country Inn

Charleston, South Carolina

21-Aug-19   **Webinar Attendees**

Last Name	First Name
Rettig	Adam
Griffin	Aimee
Bianchi	Alan
DiJohnson	Alex
Sharov	Alexei
Tong	Amanda
Kaufman	Anthony
Wood	Anthony
Krikstan	Catherine
Bruger	Catherine
Hanson	Chad
Wiegand	Christina
Graham	Cierra
Chagaris	David
Glasgow	Dawn
Lazarre	Dominique

Addis	Dustin
Fitzpatrick	Eric
Helies	Frank
Shepherd	Gary
Exley	Gary
White	Geoff
Sedberry	George
Bray	Gregg
Waugh	Gregg
Laks	Ira
Breidt	Jay
Pulver	Jeff
Bissette	Jesse
McGovern	John
Foster	John
montañez	josé
Pianka	Karen
Howington	Kathleen
Siegfried	Katie
Klasnick	Kelly
Iverson	Kim

Lee	Laura
Johansen	Lucas
Barbieri	Luiz
Terceiro	Mark
Vara	Mary
Seward	McLean
Jepson	Michael
Schmidtke	Michael
Travis	Michael
Celestino	Michael
Larkin	Michael
Drexler	Michael
Joseph	Munyandorero
BROUWER	MYRA
Mehta	Nikhil
Germeroth	Rachel
Brame	Richen
Cheshire	Rob
Ahrens	Robert
Hudson	Rusty
Lovell	Sabrina

Steinback	Scott
Crosson	Scott
Haukebo	Sepp
Allen	Shanae
McInerny	Stephanie
Teears	Thom
sminkey	thomas
Cross	Tiffanie
Lamberte	Tony
Smart	Tracey