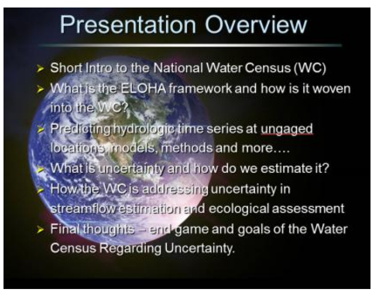
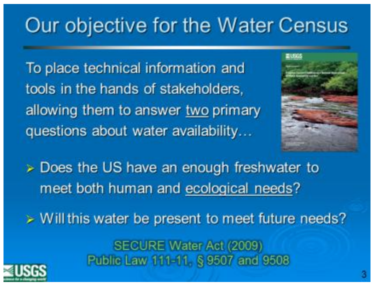
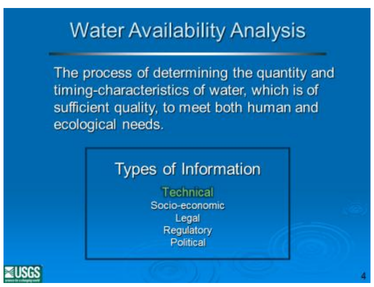
JONATHAN: Thanks, Brian. This morning I'd like to spend a little time talking about a program that I'm involved in. First, if you indulge me, the National Water Census, then second, how the ELOHA framework that Brian spoke about earlier, is integrated or woven into Water Census goals and interests. It's important to think about the Water Census in terms of one of the major tenets of the ELOHA process, which is the development of time series at ungaged locations, which we all find immanently important to the work that we're doing. And then, how does uncertainty play a role in this? The panel sessions that have gone on the last day and a half have really focusing on uncertainty, so I want to talk about how the water census is addressing uncertainties, specifically in terms of both stream flow and hydrology but also from a biological perspective.



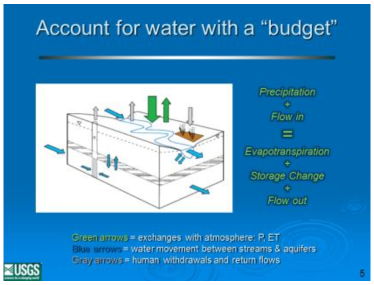
As we develop our models and as we think about flow and ecological relationships, or water and ecological relationships, we need to understand how uncertainty plays a role in our understanding of that information. And then some final thoughts, as Tom Annear would say, what is the end game or what is the "so what" story that is related to the work that we're doing. So, the primary objective of the water census really is to place technical information in the hands of the stakeholders, allowing them to answer two primary questions about water availability. One, does the US have enough fresh water to meet both human and ecological needs, and will that water be available in the future?



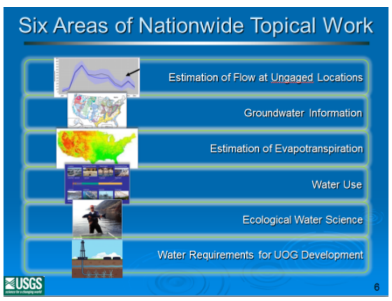
I recall that there was some discussion yesterday that water availability analysis is very important. We think of water availability analysis as the process of determining the quantity and timing characteristics of water which is of sufficient quality, to meet both human and ecological needs. But we deal primarily with the technical side. I think yesterday there was a little bit of conversation about the technical being the easiest component. That may be true, so, in essence we leave the harder stuff up to the stakeholders to utilize information that we provide in a way that helps them make good decisions about the socioeconomic, the legal, the regulatory, and the political side of the water availability.



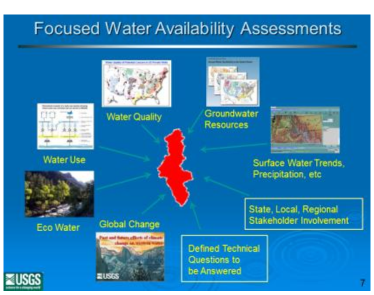
As Brian said, we deal with water accounting within the USGS, so the framework we're coming from is really one of a water budget. In the same way we think about a financial budget, we have inputs and exchanges, and flows and fluxes within the water budget process that we try to account for within the USGS. The green arrows in this case represent exchanges with the atmosphere. The blue arrows represent movement of water between streams, aquifers, lakes and ponds. And then of course, the grey arrows are the human dimension, the human withdrawals and return flows that -- and all of these components are important to us, both from the ELOHA perspective and the water accounting perspective.



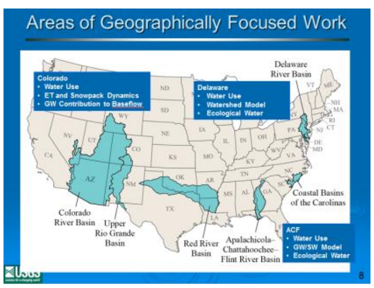
Briefly, the Water Census has six national topical areas where we are working on National applications for water availability. They include the estimation of flows at ungauged locations, the groundwater information, which is really an important component and related to water use, estimation of Evapotranspiration, ecological water science component, which is what I am involved with, and then of course, water requirements for unconventional oil and gas development.



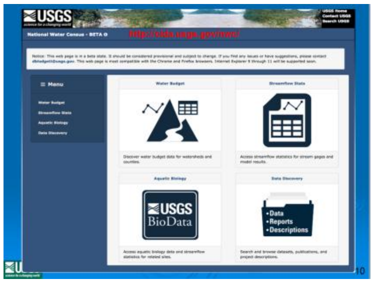
We have, what we consider to be, these focus area studies, which are basin-based studies where we bring to bear all the resources and information, and data available through the Water Census. The objectives of the Focus Area Studies are primarily driven by the state and local stakeholders interests and needs. So the questions we ask and the information we provide are really based on shared interests. This provides us with a good opportunity to use all the information that we have, water quality, grounder resources, water use, ecological flows, to answer questions that the stakeholders really have about the water availability within these basins.



This is not a cookie cutter approach because the objectives and needs differ among focus areas. Everybody has a little different perspective on what the needs might be. We are currently coming off the first round of these water-focused, geographic-focused area studies, and the first three were the Delaware River, the Apalachicola Chattahoochee Flint Basin, and then the Colorado River Basin. And you can see that the focuses in each of these studies were a little different. For example, in the Colorado River, they were very interested in water use, ET and snow pack dynamics, and the groundwater contribution to stream flow, whereas in the Delaware River Basin, the stakeholders really were interested in more information on water modeling, water availability, water use, and using that information to support a better understanding of environmental water needs. And the ACF is pretty comparable -- they were very interested in groundwater/surface water interaction and ecological water and water use.



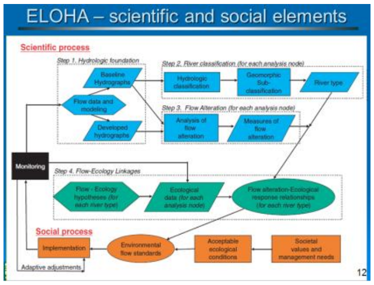
The Water Census is really trying to think about this in terms of multiple layers of information we can deliver to stakeholders at a national scale, but provide it at the HUC12 level. HUC12's are very small, 12-35 square miles, but represent an accounting based process for aggregating information on precipitation runoff, base flow, ET, recharge, et cetera. Ideally, as I said, it is our goal to put this information into the hands of the stakeholders. And so, we have developed a web portal. I'm not going to go into this in any great detail today, however, if you're interested, please visit the website you see here -- http.cida.usgs.gov, and investigate and look into some of the tools and processes that we have available, especially the delivery of data.



Again, we're trying to provide information that's useful, water accounting data, so water budget information at the HUC 12 level for the entire United States, stream flow, access to all the stream flow information, and the ability to develop statistics instantaneously based on the stream flow information, biological data, and of course, the capacity to really do some investigative evaluations of the data that we have available.

Brian mentioned the ecological limits of hydrologic alteration, the "ELOHA" process in his opening presentation. ELOHA is a very important framework that brought together a large international group of scientists. We have integrated this framework into the Water Census. ELOHA was based on, as Brian stated earlier, an extensive publication history and strong international interests in flow and water information. The goal of ELOHA was to develop a more cost-effective, scientifically defensible, and pragmatic approach to evaluating water-ecology relationships.

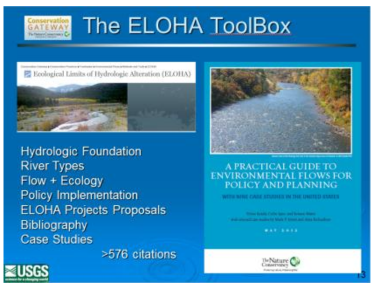
Now, in Tom Annear's session yesterday, he talked about the various steps in the ELOHA process. I'm not going to go into this in any great depth, but I think what I want to do is move the conversation a little bit further along, in the same way that Brian did, from a historical perspective, and think about water in as broad a capacity as possible. When we developed ELOHA, implicitly we knew that flow really represented all aspects of water. However, our conversation, both in the literature and amongst ELOHA practitioners, has really been driven by the idea of flow.



Now, that may be important for some, but as you heard yesterday from Tom, he asked the panels directly, "What are the implications for components of the water system, like the lakes, the ponds, the wetlands, and even the water levels that are associated with those aspects as well as groundwater?" So I modified the ELOHA framework a little to bring this conversation a forward. And the idea here is to think about it in terms of hydrology, water levels, and flow modeling, and also think about it in terms of water body type as opposed to stream flow types. I like to think of it in terms of developing water-ecology relationships not just flow-ecology relationships. I think what we found in the past three or four years using the ELOHA model is there's a lot of published papers that are really trying to link groundwater processes to changes in stream flow process that alter the complexity of the communities we see in streams. So I think we're moving in this direction already. In essence, I think it is important to represent the broader interests of the In-stream Flow Council and the many other scientists and practitioners who have influenced the way we think about these processes and the trajectory the ELOHA framework is already taking.

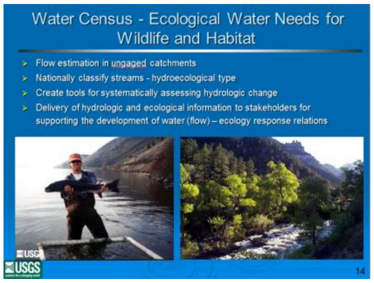
Now, again, on the technical side, the USGS and the Water Census deals mostly with water components that relate to the first three steps, maybe the first four steps of this in terms of delivering information, yet one of the larger components of ELOHA is the social process, that involves the interaction with stakeholders and their interest. So I think the ELOHA process really represents a framework that includes all of these water components, and so we have to think of that in those terms.

So, instead of going through a lot of examples of where ELOHA is being used, I want to refer you to the ELOHA toolbox because it really is a nice on-line toolbox that the Nature Conservancy has developed on the conservation gateway that gives you a really broad overview of the various information available on ELOHA, Over 576 citations and counting at this point, according to Scopus. And that's only in a very short period of time.

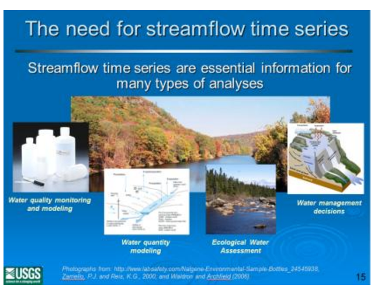


What this web site shows it that there has been a lot of work, and there are a lot of really good case studies. If you're interested in this, Eloise Kendy is doing an excellent job of maintaining this information. At this point, she is barely able to keep up with all the information and publications on ELOHA. A recent publication from the Nature Conservancy led by Eloise is the practical guide, which I think is a really good place to start with and learn about the ELOHA process.

So, how do we incorporate these attributes within the Water Census? And how do we look at this from an ELOHA perspective? The basic building blocks of the Water Census Environmental Water component is the ELOHA process. This includes the foundational elements --developing flow information a ungaged locations around the US, the classification of stream flows, the digging down into the classification structure to understanding how these streams relate to each other, and creating tools that systematically allow us to assess hydrologic change.

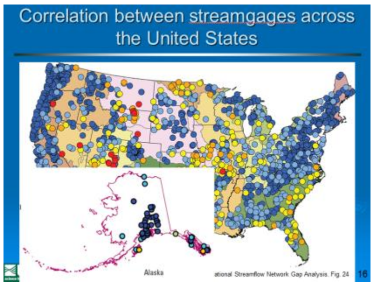


And then of course there is data delivery. As I mentioned earlier, the delivery of hydrological and ecological information to stakeholders to support the modeling and water management, and to support the understanding of the flow ecology, or water ecology relations that stakeholders are most interested in one of our main functions. Yes, developing a stream flow time series is "uber" important in everything that we do, not just the USGS, but we all need this information in order to develop any of the relations that we're talking about here. Stream flow time series are essential to this process. They are essential for water quantity monitoring, water quality monitoring, ecological water assessments, and of course, water management decisions. Without these parts of that broader ELOHA puzzle, we can't really do the science, and we can't really make and create the relationship we're looking for.

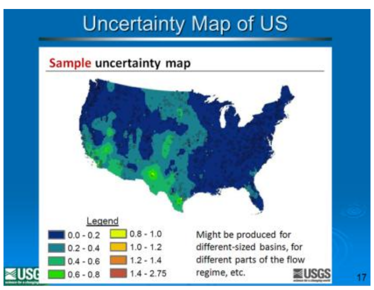


We know there is a relationship between the stream flow gauges across the US and Alaska—Christopher, I am not purposing leaving out Alaska here—we recognize that all these relations are important, and for a presentation, it's just easier to show the lower 48.

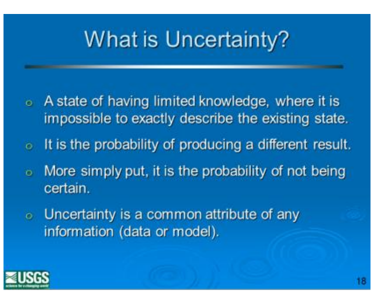
The work shown here is by Stacey Archfield and Julie Kiang from the National Research Program, but it was supported and funded by Christopher Estes. I think it is important to understand that we know that there's a capacity for models to explain some of similarity we see in stream gauges. And if you haven't seen this report, take a look because it it’s quite comprehensive. But also, the flip side of that is that there is similarity and we have the ability to transfer hydrologic information from one location to another, but as you can see, there are holes throughout the US where there is low density of gauges and –these information gaps also represent areas where there is a high amount of uncertainty.



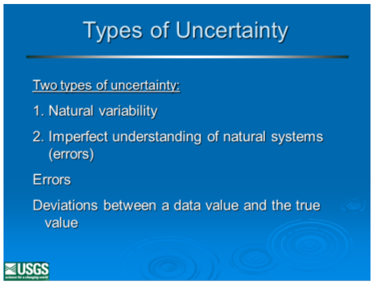
This is a sample uncertainty map. We can produce a number of these, of course, but the idea really is to show that certain areas of the country have relatively low uncertainty in terms of stream flow information, and other areas have much higher levels of uncertainty.



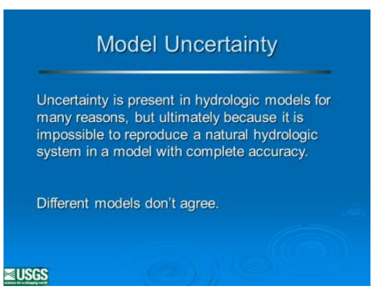
Now, what do we mean by uncertainty? I think it was defined yesterday pretty well. I’m not going to go into this in any great detail, but really, it’s a state of having limited knowledge. More simply put, it’s the probability of not being certain. And it’s a common attribute in all the data that we work with, whether its stream flow or ecological data.



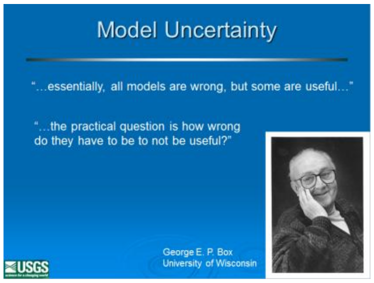
We think of it in two ways. There are two types of uncertainty. There is the natural variability that we deal with and have a hard time separating from other information, and then there is the imperfect understanding of natural systems or errors. And the deviation between the data value and the true values is what we're interested in. That's the uncertainty.



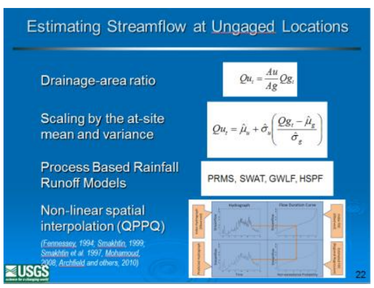
The USGS loves the models! Of course, we model flow. We model practically everything that has to do with hydrology. It's a good thing, because that gives us the ability to provide useful information to stakeholders that they can use to make informed decisions. But ultimately, because it's impossible to reproduce a natural hydrologic system, we have to understand whether we can incorporate and (or) develop an estimate of uncertainty that can be included in our models so that our models are more useful to our stakeholders.



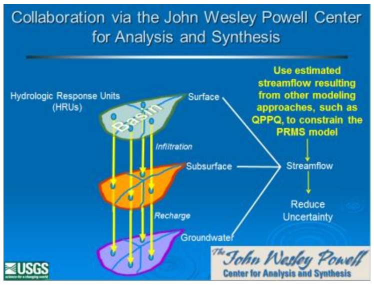
And ultimately, all models don't agree. And if you prescribe to George Box's perspective, essentially all models are wrong, but some are useful. And I think the USGS and the Water Census does not disagree with that perspective. But the practical question really is how wrong do they have to be to be not useful? And so, that question can only be answered if you understand the uncertainty that's implicit within a lot of these models themselves.



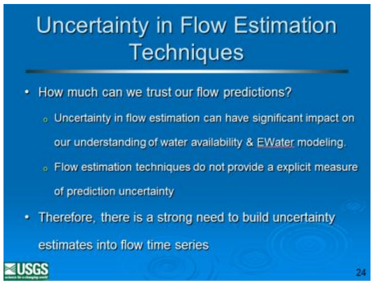
And like I said, we're in the process of working with stakeholders around the US to develop various techniques to apply stream flow estimation in various ways in various areas—using drainage area techniques, scaling techniques, process-based models that many of you of course are familiar with, PRMS, SWAT, the GWLF model, HSPF and then of course non-linear space interpolation which many know as QPPQ, this approach is really just a flow transference model that has been used throughout the Northeast and has done very well in predicting stream flow at un-gauged locations.



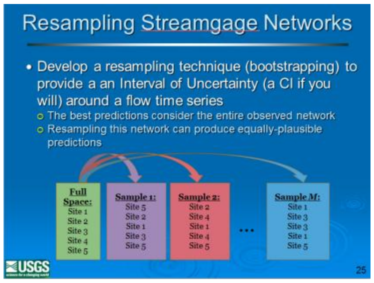
But that doesn't really get us to a better understanding of uncertainty in the prediction of stream flow. It doesn't really get us necessarily to being able to think about uncertainty in a way that we can reduce it. However, there is an effort going on right now at the Powell Center that is being headed up Stacey Archfeld, as part of an international collaboration, where they are using the best of both worlds --that is, they are combining the process-based rainfall runoff models and the regression models. The idea is to constrain the parameters of the hydrologic models using the regression-based, or transference-based models, like the QPPQ. So, we know that these transfer-based models are stationary. That is, you can't use them in a predictive fashion. What we want to be able to do is understand and control the uncertainty in flow models. By constraining the parameterization of the process models like PRMS using the flow transference models like QPPQ, we can reduce uncertainty and develop better flow estimates. This gives us the best of both worlds, and this is really cool stuff and I hope that we'll have some updates for you on this relatively soon.



The bottom line really is can we trust our flow estimates, or our flow predictions? Uncertainty definitely can have a significant impact on our understanding. I think that was brought up yesterday and was a major theme of the earlier sessions, reflecting the broad effect uncertainty has on streamflow and e-water modeling. Flow estimation techniques do not provide an explicit measure of prediction uncertainty. This is another area that we can provide water information and help move the conversation forward.

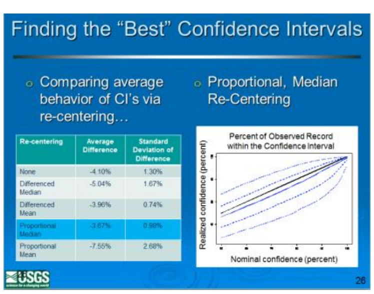


One of the research areas we are currently working on is the re-sampling networks to provide and interval of uncertainty. We have these fairly substantial hydrologic networks. And if we can develop re-sampling techniques to apply to them—such as boot strapping, or Monte Carlo permutation processes—we can use those methods to provide an interval of uncertainty or a confidence interval, if you will, around a flow time series.



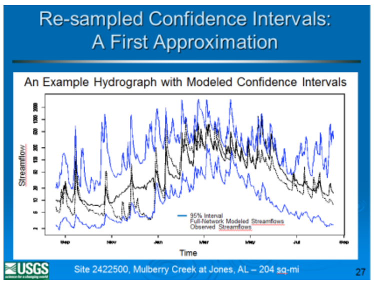
So we use the best predictions, the best information that we have available to us, and create this re-sampling network to produce a subset of equally plausible outcomes. And so, think of this in terms of your full space, your network, whatever you're working with. And we're trying to re-sample this network in a way that it gives us a number of equally plausible solutions, and thereby giving us an end number of solutions. And that way, we can develop, using quantiles, an uncertainty boundary around our stream flow information. So, ultimately, we want to find the best set of confidence intervals.

We want to provide the best information to our stakeholders which allow them to feel confident in the data that we provide. You can get a stream flow estimate, and it can be a time series, but if stakeholders do not understand the level of certainty associated with a flow time series, it's hard to use that in a management context. So we are working on developing techniques that include an estimate of uncertainty using a re-centering approach. I don't want to go too deeply into the weeds here statistically, but that the idea is really to find define —in this case, a proportional mean—that really provides us with the minimal average difference and minimizes the standard deviation.



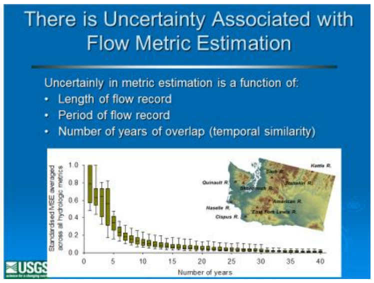
Of course, this approach it's not perfect. And right now, we're still working on refining this methodology. But if everything was perfect, you would find that the one-to-one line shown here would follow the actual flow line. And if all things were perfect within these quantiles, the blue line would directly follow the back line. But, as you can see, that's not true. But it's close. So we think we're getting close on this or are on the right track.

If we take this and superimpose it onto a hydrograph, as shown in this example hydrograph from Mulberry Creek in Alabama, which is a200 square mile basin give or take. You can see a couple of things that pop out here.

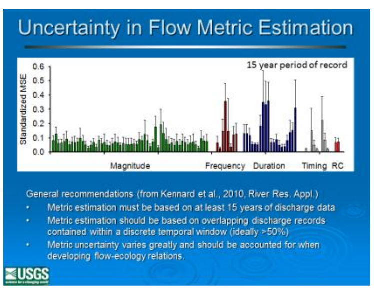


One, we were able to develop these confidence intervals around our estimated stream flow, and for comparison you can also see the observed stream flow on this graph. You will note, that there is much greater uncertainty in attributes such as peak flows and of course low flow. But during some periods of the year, you also find is that the uncertainty boundary lessens—May, June, April, May, June, March, there tends to be lower uncertainty. And that's good. So we're getting to the point where we are able to better understand what those uncertainty boundaries are. And if we can provide data where we are 90 or 94 percent confident in our ability to predict stream flow, then I think we are moving in the right direction regarding uncertainty assessment and analysis.

Now, if you take that information and create your flow metrics, you find that there is also implicit uncertainty in those. Thanks to a lot of the work that Julian Olden has done with his collaborators down under, looking at length of flow, period of record, we better understand that the longer the flow record you have, the more likely you are to reduce uncertainty, in this case expressed as mean squared area in this particular graph at this time. And what Olden and Kennard and their collaborators are trying to say, is that you want to minimize uncertainty by maximizing the length of record that you have.

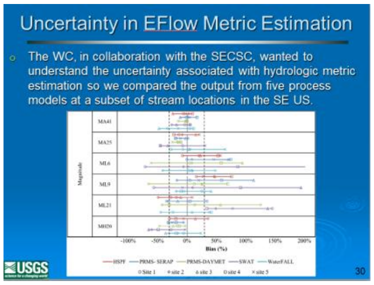


And they had a number of general recommendations which I think are really important an we, the Water Census, are incorporating into our perspectives as well. That is, according to Kennard, 2010, 15 years was a good range to minimize the uncertainty for the magnitude, frequency, duration, timing, and rate of change metrics that we all work with. Metric estimation should be based on over lapping information. And we know that the uncertainty will range and we have to be able to account for that uncertainty in the flow-ecology relations that we develop.

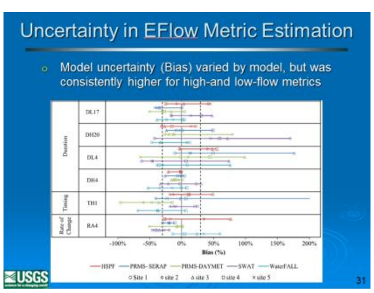


The Water Census was very interested in this uncertainty, and we are collaborating with the South East Climate Science Center to develop an analysis to look at and compare how this uncertainty ranges amongst the models that we work with on a daily basis. We, the USGS and other agencies, work with a lot of different flow models, for example, HSPF and PRMS --the precipitation runoff modeling system, of which there are a couple of parameterizations. For this effort, we concentrated on a subset of stream sites in the Southeast United States. In this graphic, the vertical dotted lines represent approximately, the plus or minus 30 percent uncertainty boundary, which is a common rule of thumb found in the literature for understanding the uncertainty associated with the hydrological data.

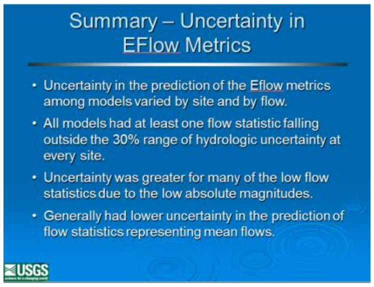
And shown are a subset of hydroecological metrics representing flow magnitude. And what you will see here—the take-home—is the there is going to be greater uncertainty, regardless of the model, in low flow metrics, and some of the high flows as well.



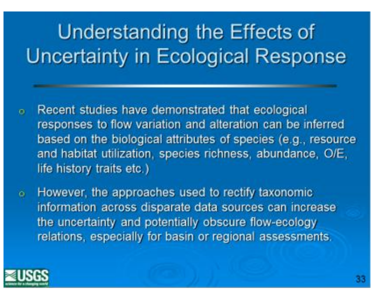
This next slide represents the duration metrics that we also looked at. And you see a similar pattern in terms of high and low flow metrics. The take home message here is we need to understand: one, how the models are parameterized and what flow attributes best support the relations we are trying to develop. In general, a lot of model parameterization focuses on central tendency attributes—i.e., median, mean flows—however, if we are interested in the low flow or the high flow, we have to make sure that we are calibrating the models to the portion of the flow that we're most interested in. So otherwise you're going to end up with fairly high uncertainty boundaries around the parameters that you're most interested in. So it is important that investigators understand the calibration strategy of the model being used for streamflow estimation and how that strategy affects estimation of the hydrologic parameters before developing flow-ecology relations.



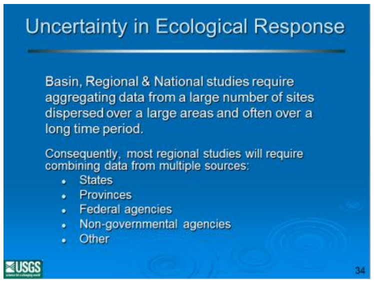
JONATHAN: Okay. So, -- no modeling method, as I mentioned, is perfect. All of the models we evaluated had the capacity to develop flow time series, but the metrics that were produced from those time series ranged in their uncertainty, and some fell outside that 30 percent range that we're trying to maintain. So ultimately, uncertainty was greater for many of the low flow and high flow attributes, and generally, the central tendency metrics tended to have lower prediction uncertainty.



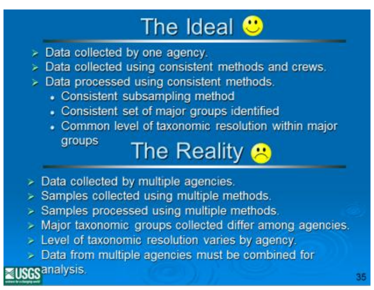
Okay. So let's come at this from a slightly different direction now. So, we talked a lot about hydrology, metric development, but what about ecology? I'm an ecologist. I deal with ecological data all the time. Is there uncertainty associated with the ecological data side of the flow-ecology equation? From my perspective, understanding how the handling of ecological data affects uncertainty in flow-ecology relations is as important as the uncertainty associated with the hydrological side of the equation. Recent studies -- there's been a lot in the past four or five years, have really expressed some fairly good relations between stream flow and ecological response. This includes relations between ecological processes, habitat, species richness abundance, O over E, and species traits.



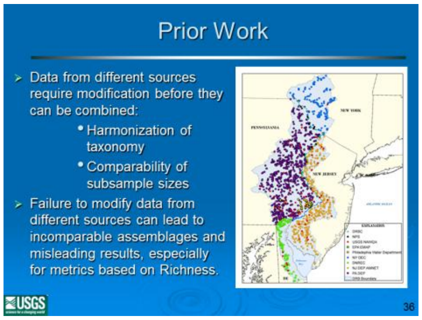
However, one of the things that really hasn't been emphasized –are the approaches we use to modify and deliver the taxonomic information equally as important in terms of us understanding the effects of uncertainty as the flow models that we are developing. So, there is a potential that we can obscure flow-ecology relations depending on how we process the ecological data. So, what we're interested in is really understanding where along the trajectory of the analytical processes do we see uncertainty creep in?



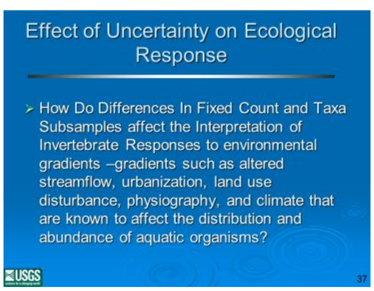
Ideally of course, it is best to be working with one cohesive data set -- however, when you are working at the basin and regional level, we are often combining different data sets from multiple agencies. And just for example, ideally, we hope that all data is collected by a single agency, but that typically is not true because the reality is the data collected in these types of studies is by multiple agencies. So samples from multiple agencies are collected, and they're utilized, and they are processed differently. And so there are some inherent discrepancies in the way in which these data can be analyzed. So, our perspective is well, let's understand the limitations of this process.



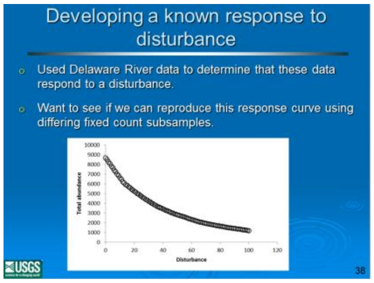
In some prior work, we were working with data from the Delaware River focused area study. We had nine different agencies that were part of this process and supplied data. We wanted to look at the comparability of the data, the harmonization of the taxonomy and then try to put all the information on an even footing –that is, make it comparable. Basically, what we found is failure to modify and make these data sets comparable could provide misleading results.



I think most of us understand that if you have a mixture of taxonomic levels, you are going to increase uncertainty. But really, what we are most interested in is whether we can extract a disturbance gradient, such as flow alteration, from the data that you are using. If the way you assess the data obscures that relationship, then you have basically lost the capacity to develop your models. And it may be why that many of us who work in this field end up with these shotgun relationships with a lot of scatter in our flow-ecology models. If all you end up with is a big spread of data and then you throw a line through it or a quantile regression, there will be very little you can interpret for those relations? So the idea here really is we need to tighten up these relationships.

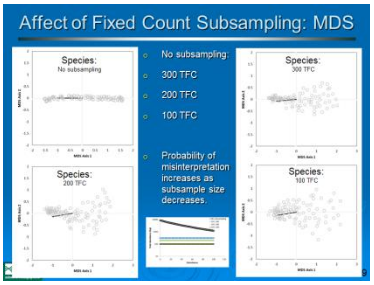


As you explore the literature, in general, you will find that many of the flow-ecology relationships are not very tight. But, most practitioners are doing the best they can with the data that they have. So what we did – and this is really cool – is, we constructed a hypothetical disturbance gradient based on the data that we had available to us within the Delaware River Basin – so we know there's a gradient there. As you can see, there is a defined disturbance gradient based on the relation between taxa abundance and the rate of disturbance response. And what we wanted to do is see if we could sub-sample the data and then reproduce this gradient with the data that we typically would use with these types of analyses, so that we could better understand the ecological side of the flow ecology equation.

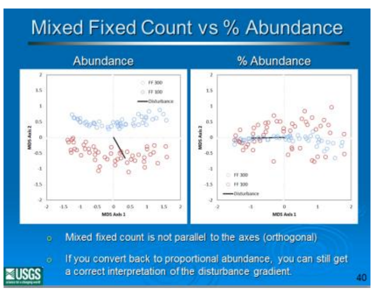


So, what we found was that as you go down in the sub-sampling structure (for example, from 300 to 100 taxa), the probability of misinterpretation increases as sample size decreases. So, what does this mean? Does it mean that everybody who uses a 100 level count data should toss out their data? No, of course not. What this indicates is that your probability of misinterpreting the disturbance gradient increases as you go to lower and lower levels because you are increasing the level of uncertainty.

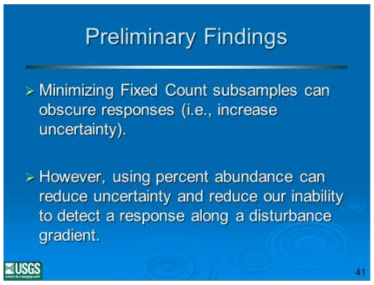
So in the top left graphic, you see a small black line which represents the hypothetical disturbance gradient, i.e., there is no sub-sampling. This is an ordination plot, a non-metric multidimensional scaling plot looking at the relationship, and it does a pretty good job of representing the hypothetical gradient, the disturbance gradient without any sub-sampling. However, as you progress from a 300 to a 100 level taxa count, you can see that the spread in data becomes a little bit broader. You also see much greater orthogonality between the actual gradient itself and the data as we move down into smaller subsamples. So what we find is that there is a greater capacity to misinterpret the actual gradient because there is an increase in noise as you work with data that has greater and greater variability.



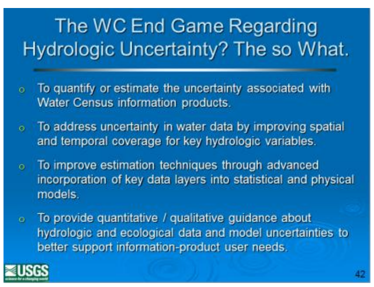
I think this finding is important to bring in to the conceptual framework of the flow-ecology models. Now, a lot of us work with just straight family level data, and I think this is good, but what I want to show here—and this is the second part of the analysis—was that if we take a 300 level or 100 level count, it really does not necessarily do a very good job using strict abundance measures of following that disturbance gradient. In fact, you have a very distinct orthogonality between the gradient and the data you are trying to use to interpret the gradient. But, the message here is we can fix that, to a certain extent. If you convert abundances to percent abundances—and that is converted back to percentiles—you find that it does a much better job of representing the hypothetical disturbance gradient. So good news / bad news is that as you move down in fix counts, you find that you increase uncertainty.



Okay, we got that, but also there are certain aspects about the way we utilize these data that will either increase or manage uncertainty from an ecological perspective. So, fixed counts can obscure response along a gradient as you go toward smaller and smaller subsamples, and then we can reverse and (or) reduce uncertainty by detecting or changing our analysis more to a proportional abundance perspective. So, I think I'm running out of time here, right, Brian? I'm good?



So really, the end game here from a Water Census perspective, and I know I covered a lot, is to quantify or estimate the uncertainty associated with a lot of the data products that we work with, the information products. Because our goal is to get this information to you, and ultimately, I know that everybody here wants to work within as minimal uncertainty boundaries as we possibly can. I think that's important. So we want to provide data that allows you to address uncertainty in water level data, in stream flow data, in lake level data, in wetland information, to improve the estimation techniques and allow you to understand how much flexibility you have in water availability management and to have confidence in moving forward with the developing flow-ecology. And then of course to provide the information, both hydrologic and ecological, that allows modeling uncertainties to better support the information product users, and that's you. Basically, it's the stakeholders.



So, I'm going to end there. I thank everybody for your time. I also want to thank the Instream Flow Council for giving me an opportunity to talk today, especially Tom Annear and Christopher Estes, thank you. I really do appreciate the opportunity. Thanks Again!