USGS Hydrography Seminar Series Session 7

[00:00] [silence]

Steve Aichele: [00:11] Thanks for joining us. This the seventh session in our Hydrography Webinar Series. My name's Steve Aichele. Al Rea will also be helping out some of the host duties through his call. Sue Buto, our WBD coordinator will also be helping out with the Q&A session later on today.

[00:33] Sorry for the delay and for the new phone number. We actually had so many people register for this we had to change teleconference technologies late in the day yesterday, so we're getting some things sorted out here.

[00:50] The purpose of these calls is to really get the opportunity to share some success stories and share information about using the National Hydrography Dataset, the Watershed Boundary Dataset, and other products that come about from those datasets in a forum more or less like a conference or seminar setting.

[01:14] Today, we have two talks, one from Dan Isaak with the US Forest Service (sorry about that, I'm not sure where that went) about building a National Stream Internet, and a second session from Howard Reeves with the USGS Michigan Water Science Center, talking about integrating hydrography and fisheries data to assess ecological flows. That's in the context of some work on the Great Lakes Basin Compact and environmental flow frameworks.

[01:46] Between sessions, we will have a Q&A. Because of the phone bridge, the risk of somebody putting us on mute, hearing Muzak for an hour and a half, having to cancel things, and make that up, everyone is in lecture mode.

[02:01] That means we use the Q&A tab within the WebEx to convey questions to the speakers. Then, Al and Sue will read those off. That's also helpful because it gives us a written transcript of those things, which we will publish to the Web later on.

[02:19] At the end of the sessions, stick around for the poll. We have just three simple questions for you related to this session and content for future sessions.

[02:29] If you want information on the seminar series, visit the website. This seminar will be posted, there, including a video, a transcript, the questions, and the PowerPoints. All the previous sessions are also posted to the website.

[02:45] If you're interested in more NHD/WBD news, send an email to the link, there, and we'll get you on the newsletter list. We plan to do at least one more of these before going on summer break here. We plan to do one on April 28th, same time. With that, I'm going to pass the ball to Dan Isaak and we'll hear about National Stream Internet.

[03:11] [background sounds only]

Operator: [03:11] Please check your mute button.

Dan Isaak: [03:21] OK. I just clicked share my screen, and can you hear me OK?

Steve: [03:31] We can.

Dan: [03:33] OK. Let's see. Let's get this set up. OK, thank you Steve and Al for the invitation to participate in this webinar series. It's great to see that level of interest and enrollment in today's webinar. I am going to talk about building a National Stream Internet. The basic idea behind this is to try to create a generalizable and flexible analytical infrastructure that is specific to data measured on stream networks.

[04:06] Before I get too far into that I do want to also acknowledge my co-authors on this project Erin Peterson, who is a spatial stream ecologist with the Queensland University of Technology over in Australia. Dave Nagel, who is our geospatial analyst and lead here in the Boise Aquatic Sciences Lab out in Idaho that has handled many of the geotechnical aspects of this project. Jay Ver Hoef who's a biometrician with NOAA, the Marine Mammal Lab in Fairbanks, Alaska, and Jeff Kershner who is a fish biologist with USGS, recently retired.

[04:41] This is also then is a project that was funded through the US Fish and Wildlife Services Landscape Conservation Cooperative. They had a national initiative RFP a few years ago and this was one of the two projects that that particular outreach funded. The motivation or the possibility of this Stream Internet and creating new and valuable types of information for the nation's aquatic resources basically emerges from the fact that there is so much data out there already.

[05:20] Hundreds of different agencies have spent hundreds of millions of dollars collecting data on water quality attributes, biological attributes, habitat conditions, etc., etc. The fact that those data exist opens the world to just a giant array of possibilities, if we can build a system that allows us to tap the information out of those databases.

[05:46] Before we get too much further along in terms of what we're going to talk about today, I do need to give you kind of a concrete definition to this otherwise fuzzy concept of a Stream Internet. This is what we've come up with for describing this. It's a network of people, databases, digital information systems, and analytical techniques that interact synergistically to create and communicate massive amount of information efficiently.

[06:13] There's a variety of different things then that go into that. First of which is just people on landscapes from those many different agencies that are interested in aquatic resources, collecting data on all manner of different things.

[06:28] Ideally they're using standard protocols that are reparable by other groups that are also interested in those attributes. Then you want to have that data flowing into open access comprehensive databases, so that it's usable not just by the folks that collected at one point in time for one particular study, but also that it can serve as an archive that we can tap into later as additional ideas come around about how to use information.

[06:57] Thirdly, then analysis and extracting information from those databases. If you can build inner compatibility amongst all these pieces then you can complete the loop here where that new information that's been developed from those big databases, then ultimately flows back to people

making decisions about conserving and managing resources on the landscapes so that they can do their jobs more efficiently.

[07:21] You just build an iterative loop here so that the next set of data they collect in, it's going to be collected more efficiently in places where we don't already have it etc., etc. Inner compatibility amongst all those things is really important if you're going to assemble this larger system from component parts. Next I'm just going to step through what some of those key components are.

[07:45] First of which is the national hydrography data set itself. The fact that Al and Tommy Dewald and others have been working on this particular project for, my understanding is almost a couple of decades now, to build a nationally consistent stream network geodatabase, is just a huge thing, and is the foundation that everything else that we're...I'm trying to do here rests upon.

[08:11] I don't need to tell folks on this webinar more about this. It's absolutely key to be able to pull this whole thing off.

[08:19] Second part of that, and that's really important for us in terms of thinking about building models or predicting things about stream network that we want to better understand is the plus part of NHD.

[08:31] The fact that each unique reach ID is then also has associated with set of nationally consistently derived reach descriptors, so things like elevation, slope, land use, etc., etc. These are the value added attributes in NHD plus framework. That's just a huge thing.

[08:48] Those are sorts of things that we would often use as predictor variables in a model of stream temperature or species distribution. The fact that those are all done, and we can draw on those greatly accelerates other aspects of this work.

[09:06] There's also other groups now that have pulled together databases of these reached descriptors, so that Wang et al group with the National Fish Habitat Assessment a few years ago.

[09:17] Then more recently Hill et al just published the Streamcat database, an EPA funded effort to, again, build a consistent set of reach descriptors that are tied to NHD plus version 2 in that particular case.

[09:33] The third component of a national stream Internet then is just those mountains of data that already exist, whether you're talking about temperature, or discharge, or water quality parameters, or species distributions, genetic attributes, there are literally tens of thousands of unique streams sites across the country where these sorts of things have been measured.

[09:55] I would argue that we are literally only scratching the surface of the amount of information that resides within those databases. We want to build the system that allows us to tap that information out of what it already exists.

[10:09] Fourth component, and this is a key one, it is having statistical model that are actually specific to data measured on stream networks. This may seem a little bit odd at first that we're building models all the time. It's not that we don't have models.

[10:26] The vast majority of things that we're applying to stream systems at least in, and we're looking at empirical statistical models, those were developed originally for terrestrial systems. Those don't account for, or even acknowledge, that the stream network exists.

[10:44] To really optimize that the type of inference that you're going to draw from these sorts of data sets, you want to have models that are designed to work with them. That's a lot of the work that Erin Peterson and Jay Ver Hoef have done over the last decade or so, is to basically build that class of statistical model.

[11:01] Jay is a statistician and biometrician. He's published a lot of theories for these models now in the statistical literature. Erin has worked closely with him to translate that into geospatial scripts and other forms of software so that it's more accessible. There's a nice set of software that goes with this and runs in R.

[11:24] I guess too, that the origins of a lot of this work actually date back probably 15 or so years ago to an old EPA star map grant. This is work that funded Erin to do her dissertation research on the flows project. The functional linkage of watershed and streams that has just continued on in subsequent iterations to get to where we are today.

[11:47] Just a little bit more about those models. Their key innovation is having that co variant structure that is based on network topology. The models to understand the direction of flow, they understand tributary confluences, and the differential waiting that a small stream has relative...

[12:06] [phone tone]

Dan: [12:14] The other thing that these models can do is because they are spatial statistical models, they can account for auto correlation or the clustering in large aggregated databases that you might pull together from multiple sources. That sort of thing is really problematic to deal with unless you're running a spatial model that doesn't understand the spatial coordinates of where things are collected in space.

[12:37] These models can work with that, and actually that's one of the things that goes into allowing them to make better predictions than some of the traditional non spatial models. There's a website that has this information online. If you just want to start playing around and gather familiarity with running these models, just do a Google search on SSN and STARS and that will pop up. In this website are their software that you can download, example datasets and lots and lots of documentation.

[13:10] Two last things about these models. I just want to point out that these are a basic generalizable class of statistical model. There's nothing here that's inherently specific to using it for just stream temperature applications. I'm going to show here in subsequent slides.

[13:28] These are just a general branch of statistics. Like any other branch of statistics, it's just that they're tuned up to work with these stream datasets. You can model all these different attributes, and you can model them using different length functions, whether that's a Gaussian distribution, Poisson distribution, or binomial distribution.

[13:46] The fifth important component to the NSI, then, is having a network that is tuned specifically to work with these SSN models. I mentioned that we're building everything off of that NHD hydrography layer. But at the same time, the way it was developed, it wasn't done so specifically for these SSN models. One of the things that these network models need to have is connectivity throughout the full network and have unique flow paths.

[14:21] The biggest value added new thing that we did as part of this project was to go through the NHD Plus version 2 hydrography layer for the lower 48 states, and recondition it slightly such that we can maintain those unique flow paths, things like uninitiated flow, braids and diverging flow, etc. You'll see the list there.

[14:44] We basically had to go in and do some slight modifications to that. The specific modifications that we did are all well described in the user's guide that Dave Nagel developed that has kind of the metadata on how those adjustments were made.

[14:57] Key, though, we did want to make sure that we maintained back compatibility with NHD Plus, so that once someone does an SSN analysis using this NFI network, they can take the results of that and map those back onto to the NHD Plus network. You're not losing anything by going off and doing these SSN models.

[15:16] The sixth components, then, are having killer apps, the sort of things that people want to get better information about, and that these models potentially offer ways of getting much better information than we've had in the past.

[15:31] There's just a whole laundry list of things that can be done with these SSN models. Things like parameter estimation and significance testing. Lots of statistical models can do that. These SSN models, especially when applied with dense datasets, will generate more accurate parameter estimates and correct statistical significance levels.

[15:53] You can use the models then, once you have a model fit, to make predictions that un sample location throughout a stream network. If you want to make a status assessment, and just an accurate map of what things look like across a larger area, they do that very well. Efficient monitoring design that accounts for that directionality of flow and proportional weighing of different stream sizes. There's a paper that looks at that specifically and how you can use this SSN framework for that

[16:22] Block kriging to make predictions or comparisons between reference and control sites, so that you can do an apples to apples comparison to look at how maybe watershed conditions affect water quality versus some reference site. Fish population estimates at large scales can be done with block kriging,

[16:41] Then one of the really powerful applications, and I'll show an example of this here next, is mining big data databases. Because these models can account for spatial correlation among sites, and that doesn't degrade their performance, the more data you load into the models, the better they are. That allows you to aggregate data from many existing sources. A lot of these different applications are described in that Isaac et al applications paper that we provide the reference for at the bottom of that page.

[17:10] Just an example, then, of how we might go about mining a big data database. One of the flagship projects that we've been building using the NHD network and these SSN models is this NorWeST project, which is an attempt to build an integrated interagency stream temperature database across 100 different agencies in the western United States. Then, basically, use those datasets to create a set of interpolated high resolution stream temperature climate scenarios, so that different agencies have that information for planning purposes.

[17:46] We're fairly well along in this project. Actually, by the end of this year we'll have completed the entirety of the western United States. At present, this database has more than 150 million hourly temperature records in it, copies of data that already existed that people send to our database teams so we could forge it into one usable database.

[18:08] There are more than 20,000 unique streams sites represented in this database. If you want to go out and just collect that much data from scratch, it would cost something on the order of $10 million. But the fact that we can pull this data together from existing sources, at a relatively low cost using a small database team, allows us, then, to start adding significant value to what already is out there.

[18:33] On top of just that value of the raw data, then, is going to be all the information that flows out of these databases once we start setting accurate models to them. That's harder to put a price tag on, but it's got to have a multiple that is many times that $10 million figure that it would cost to collect the raw data.

[18:54] This is just comparison and these two graphs of fitting a traditional non-spatial non-stream model to a big temperature data set, for several river basins within the western United States and have a lot of data. The bottom graph there then shows fitting the spatial SSN model to that same dataset and you can see that there is a big gain in predictive accuracy and performance of the models with the same dataset once you apply these spatial models to them.

[19:27] Typically, with these aggregated data bases. We got an r-squared around 0.9 and average prediction error of about one degree Centigrade.

[19:37] The predictor variables that are used in the models are listed there on the left. Those are all things that come off of nationally available pre existing geospatial data bases things out of NHD network. The StreamCat database etc. It's basically just a mapping exercise once we get to this point of having an organized database that we can work with in reference against those geospatial layers.

[20:07] This map is just the end result of going through that process of building the database, linking it to NHD Plus, and [inaudible] it and then modeling it and using the SSN model so that we can interpret among those observation locations. This is the thermal-scape across big chunks of the western United States.

[20:27] We've got that information and a high resolution throughout that full stream network. We got the model set to make perdition that one kilometer resolution and we can do that now across more than million kilometers of rivers and streams.

[20:41] The last part of what we do in this project, but many of the others that we are doing routinely is to build a custom website to deliver that database and interpolated scenario maps out to user community. We've also got a NorWeST website were people go to get that information. That gets heavily used because the data that went into it were contributed by people that are trying to address the whole range of questions related to temperature and climate change and regulatory standards out there on the landscape.

[21:15] These are just some examples of the ways in which people use the data and temperature scenarios that they can get from the website. Just to kind of wrap up here and kind of circle back here to this broader concept or idea of a Stream Internet the project was fairly straight forward. We had three simple objectives that we were trying to achieve, first of which we was developing compatibility between this SSN spatial stream analysis tools and NHD layer. That was done and completed sometime earlier this year.

[21:54] We also then hosted a national workshop about a year ago here in Boise. Al was part of that. we also tried to get leaders from other national aquatic programs with different federal agencies that have a broad mission and vision to what they're doing as well as engaging key researchers in just a discussion about a possibility of what can come out of these existing technologies and data sets.

[22:23] Thirdly, we're working to develop a grassroots user base with the technical skills to apply these tools. I will talk a lot a little bit more about that in a second. The hope is by having done some of the on the ground geospatial work throughout different parts of the country now we can make these sorts of big data mining projects like NorWeST a lot easier for other groups in the country to do.

[22:47] There will be incentives now that it exist for aggregating smaller data sets into larger databases that we can use to extract information from once people see all the possibilities. We've got a website now that we launched in association with the National Stream Internet project.

[23:07] There are a lot of different components here things that I talked about today, and basically the website can act as a portal that pulls those different things together. We've got some of the information from the workshop that we did, the presentations that we were given from last year.

[23:24] We got the geospatial layers for that NSI hydrography network that is slightly modified version of the NHD version 2 network. If people want to download that they grab the information there. We also provide links out to databases of stream reach descriptors that different groups have developed, things like the StreamCat database, or the [inaudible] et al database was developed for the National Fish Habitat Group Assessment a few years ago.

[23:56] Because all those things have their own websites, we are just providing a way to link to those, so that other users that aren't familiar with the fact that those exist, if they come to the NSI website they then can link out to find those resources. Then also stream databases of stream measurement. There is a lot of data that is being pulled together and organized by different groups kind of as we go through this process. These are just a couple of examples of the large measurement databases now that exist for different parts of the country.

[24:29] This Western Center for Monitoring, Assessment and Freshwater Ecosystems is run out of Logan and they've got thousands and thousands of sites where macroinvertebrates have been measured in streams across the country.

[24:41] The same thing with MARIS this multi-state Aquatic Resources and Information System, basically millions of observations that are geo referenced, available online on websites where people can go to download that information and to start thinking about how you might use that with these other technologies to build better models.

[25:02] The last thing or major thing that we do on the website is just kind of keep real time up to date bibliography of different studies that have applied these SSN models. As this evolves going forward people can go to this bibliography and access those different applications of the model to get a better sense of different ways that the user community is now applying these technologies.

[25:29] Lastly, just trying to grow that user community, we launched the SSN & STARS website for the statistical software a little over three years ago. We get about 10,000 visits a year to that website now and people are in there downloading the software and example data sets, etc. It is fairly popular. We also host annual training workshops in Boise where we have Aaron come over from Australia and Jay comes down from Fairbanks. Then we have people across the country that are interested in applying these models.will travel here for a several day workshop, where they just get to work hands on with Aaron and Jay to better understand the statistical code, and link these different components to their data sets, so they can start to model them for a variety of purposes. We are to the point where a lot of the heavy lifting has been done to make using these models and linking that to the NHD network as easy as possible to do.

[26:35] The last step is really just for people that are interested to build those databases that they want to model for anything, anywhere in the country, and just have at it. Once they have the technical expertise to do this stuff, then they can take the ball and run from there is our hope. The vision motivating all of this is what we call TSI or total stream information through the NSI.

[26:58] The hope is that may be a few years down the road, as these things get to be used more and more frequently, you could literally throw a dart at any part of the country, and wherever that dart landed on a stream reach, you would have a precise prediction of the water quality attributes that are there, a statistical confidence interval on that, you would know something about the biology, and all the other things that make that particular reach of stream a cool thing for aquatic critters, and something that we are interested in as nation in terms of managing and conserving our aquatic resources. I think that is all have got.

Steve: [27:36] OK, thanks, Dan. We did start a couple of minutes late, so we may just do a question or two for Dan while we switch presentations to Howard Reeves. If you do have a question for Dan, use the Q&A button at the top of WebEx window. You should be able to enter questions in through there, and then Al and Sue will take care of that.

Al: [28:04] The Q&A panel, it moves around. Right now we are not seeing any question in there. Dan, I just wanted to see if you could clarify a little bit about the network and how...You had to make some modifications to the NHDPlus stream network to make it work with the

statistics that you have. Just wanted you too comment on maybe how you link back to the NHD network from that, from the modified network.

Dan: [28:53] Probably the best thing to do, the protocols are all documented in the NSI users guide. As far as the technical details that's really something that Dave Nagel, here in our shop, he is the person to talk to about that. I could take a swing at it but I would, I'm sure, flounder and not be as accurate as I probably should be. I would direct you to those resources.

Sue: [29:17] We have time for one more question. Dan, the NHDPlus, the questioner believes is 1:100K resolution. Has any work been done with high res NHD, the 1 to 24k data?

Dan: [29:38] A person can use these SSN models with a stream network that comes from lots of different sources. It can be the 1 to 24k. Can be a synthetic network that's derived from a digital elevation model. We chose the NHDPlus network in the 1 to 100k because of the fact that there are those reach codes in it that are already built and linked to every reach in the country.

[30:04] That just makes the modeling and the mapping of the predictions from the model so much easier than if you have to build those covariats yourself. That's kind of the default strategy that we've taken. People aren't locked into having to use that. But it's just going to be more work to use some of these other stream networks at least until the 1 to 24k version of NHDPlus that Al and Tommy and the development team are working on becomes probably available. At that point we will then probably migrate what we are doing over to the 1 to 24k.

Al: [30:46] Dan, I'll just mention, this is Al Rea. I will mention that we are working on a high resolution NHDPlus. We're just getting started in production on that. We hope to have around half a dozen of hydro regions of data done this fiscal year. We should have some sample data for people to look at pretty soon. Nationally it's going to be a year or two at least before we have that done, but we will have something very similar for high resolution NHD soon.

Steve: [31:28] Thanks, Dan. Thanks, Al. We will move to Howard Reeves, and see his presentation. Then we'll use whatever the balance of the hour is for questions about either. Take it away, Howard.

Howard Reeves: [31:39] Thanks, Steve. Can you hear me?

Steve: [31:42] Yes.

Howard: [31:42] I want to fill in a little bit on what Dan just presented, and rather than showing the background data show some applications that we were involved with here in Michigan, and then broader in the Great Lakes Basin. One thing that's happened in the last 10 years or so was that the Great Lake states passed the Great Lakes St. Lawrence River Basin Compact, which is an interstate compact. It has a companion international agreement with the two provincial Canadian provinces.

[32:17] In that agreement, in the compact, the states and provinces pledged to, on that last bullet, prevent adverse resource impact from new or increased water withdrawals. There's some other features in terms of water conservation and efficiency, but I want to focus on this last bullet here,

adverse resource impact. Each of the jurisdictions got to define what they meant by average resource impact.

[32:54] In Michigan, what was done was to use a system similar to what Dan outlined, that combined stream information, water temperature information, and ecological information, in terms of what the fish species presence abundance was to define an adverse resource impact, in terms of would the water withdrawal that you're proposing change the characteristic fish populations at any point in the state?

[33:31] To implement this assessment process under the compact, we had to combine these three models, the stream flow model, somehow, to determine what the flow is at ungaged regions across the state. A community model, what kind of fish are present everywhere, and then, how does that community respond as the stream flows are changed? Then, a withdrawal model to relate either surface water withdrawal or ground water withdrawals to changes in stream flow.

[34:07] You can imagine that a surface water withdrawal is pretty simple, just directly removes water from that stream. A ground water withdrawal would be distributed in space, maybe among multiple streams, and delayed in time. We want to know what the impact of the pumping is on stream flow.

[34:28] In terms of the NHD and the NHDPlus, those provided the background data framework to build these withdrawal models and then, subsequently, serves as the data structure for the legislation that was ultimately passed in Michigan in 2008. The NHDPlus, as Dan outlined, brought in some information about the catchments associated with each NHD reach. That information was used in regression models to make estimates of streamflow, and they're also used to make estimates of stream temperature.

[35:13] It related individual measurements of fish to fish models that would predict fish presence and abundance across the state. It feeds, finally, into a accounting system, so that any withdrawal can be accounted for, and how it impacts the streams as they move downstream across the state.

[35:37] One of the key steps in the implementation in Michigan was to classify streams to recognize that streams are different as you move across the state. We ended up with 11 different classes based on size and temperature. This slide is illustrating the theory of Michigan is dominated by warm streams that are runoff dominated. The upper part of the lower peninsula of Michigan is dominated by base flow, groundwater fed streams, cold and cool water.

[36:19] The state legislature and the scientific advisory panel that helped recommend this legislation, wanted to make sure that we treated these systems differently, between the famous trout streams of the upper part of the lower peninsula of Michigan and runoff dominated streams in other parts of the state.

[36:51] To build the response curves, or how the fish community would change based on streamflow changes, I put this flowchart in here just to highlight the amount of data and the different kinds of data that go into this decision making. There's stream temperature data, fish surveys, information from the NHD and the WBD, in terms of what's the area associated with the streams, what's the surface geology, what are the other landscape characteristics that are associated with each stream reach.

[37:29] That fundamental data, which is very similar to the Stream Internet that Dan Isaak just outlined, to, then, be used to build a model, as he discussed, and then, finally, to provide outputs that were then used to make decisions. In Michigan, the legislature that put this legislation together decided that they would look at ecological response curves. How does the fish population, the characteristic fish population change as you remove flow from the stream?

[38:24] Instead of just saying there's one point where, if the streamflow reduction is low everything is good, and then, once you hit a hard point, we've crossed a tipping point, and it's now bad, they decided to make zones to help with the management of the resource within the state.

[38:51] We provided these response curves that were built on NHDPlus, the response model that I had in the previous slide. The legislature that decided how to break these zones together to say, "If your withdrawals are in this zone, then the user just has to register the use, and they can move on."

[39:16] If the withdrawal starts to move into this zone B, then you have to alert your neighbors, and you have to do conservation activities. As you move closer and closer to that line that we're trying to protect, there's more responsibility of the users and then the community to manage the resource.

[39:36] Here's how those lines appear for the 11 different classes that I outlined on that earlier map. The class here that's called cold transitional, they're dominated by cold fisheries, but they're very close to warming up and maybe losing that fishery.

[40:02] When the legislation was finally implemented, they only allow very little water to be removed from those systems before conservation measures need to be implemented. They built these ideas into an online screening tool to allow users to register new uses without having to go through a large formal permitting process.

[40:39] The advisory council that helped advise the legislature on this recognized that there's a lot of water in Michigan, and many uses aren't going to have an adverse resource impact.

[40:51] The screening tool embodies those different models, the stream withdrawal model, the fisheries model, and the response model, and that allows users to apply for a new or increased use without a long process. Water withdrawal tools went into effect in about 2008 and 2009.

[41:24] At about that same time, groups here, like coauthors and colleagues here in the Water Science Center were asked whether we could take those same science components from Michigan and provide the same kind of information regionally.

[41:42] This work was funded by the Great Lakes Restoration Initiative. The idea was we weren't involved in a legislative process the same way we were in Michigan, but we could provide the data, the response curves, and the classifications that would allow people to think comprehensively across the basin and build management decisions on that same consistent data. Again, we needed to use the NHD and NHDPlus. We used a different method to estimate streamflow that was even more tied to the stream network. I'll talk about that in future slides.

[42:28] We used the streamflow model, stream temperature model, and, again, the fish presence and abundance modeling. In order to estimate flow in the ungaged basins, we used a method called AFINCH, which is a constrained regression technique that's explicitly tied to the NHDPlus. It uses the routing information to route the mass of water that moves downstream, and then check the estimates at each active gage.

[43:00] We'll do a linear regression, sum the flows down the network until we get to a gage and then at the gage, we'll correct the estimate by adjusting yield upstream of that gage. The flow information, the connectivity information in NHDPlus became crucial to producing the estimates of ungaged flows across the basin. We put these out, now, in a mapper from 1950 to 2012 flows. It produces a time series of monthly flows.

[43:40] At this website, we can come in now and, built on that medium res, 1 to 100,000 NHDPlus, we can show estimates of the flow at any point. Here's an example of just one reach from that system that shows the monthly flows for this Flat River site. If there was a gage available here, we would provide a link that allows us to also download the gage data.

[44:17] Another way to look at this information is to look at the yield for the flow per drainage area, and look at what parts of the landscape are producing the high yields that lead to the streamflow in the system. What Jim McKenna and his colleagues and the Great Lakes Science Center have been doing, are taking these time series of flows and relating them to the fish presence and abundance data that they have in their labs.

[44:57] We looked at August median yields and fish data normalized to catch per unit effort to develop response curves for different species across the Great Lakes. Here's an example for brook trout, using all the records and the raw data, looking at the catch per unit effort, and the number of brook trout they found in these different stream classes across the basin.

[45:36] Then, we step back and generalize that a bit and say, what we see, in general, is for brook trout, they're dominated in the cold streams, is where we see the highest abundance of these fish, and we can see where does this population respond most quickly to changes in flow. Jim and his colleagues have labeled these the critical flow zone for each of these environments, in terms of as a fishery's manager, where would you be most concerned about changes in flow, because it would have the biggest impact in the population.

[46:15] If you have a really cold stream, very high yield, and there's some change out here, then the stream is pretty well buffered, and you may not expect to see much change in the brook trout population if there's a change in flow, as reflected by a reduction in yield. As we start moving into this zone, then the system would be stressed, and then, here, the change is very rapid. This is a critical zone, where you could lose a lot of population for a relatively small change in flow.

[46:57] These were developed for a whole number of species across the Great Lakes, and then, what the researchers did, is take this idea and map it back to a spatial distribution to point out to, in this case, for fisheries managers for brook trout, which streams appear to have the optimal flow and temperature characteristics to support that species, which ones would be stressed, and which ones are critically stressed.

[47:30] A fishery's manager can look and identify areas where, let's say, we might expect brook trout to be here, but because of the current flow and temperature measurements, we think it might be a stressed stream, and might be an area we want to go in and look at in more detail before we make management decisions there. This is the last slide, and because they're posting these PowerPoints that'll be available for you to download from that main website. Thanks, that's it, Steve. Thanks.

Steve: [48:12] All right. Thank you, Howard. We have a few minutes for questions. Again, use the Q&A button on the WebEx button bar or elsewhere. I'm not sure where else it'd be, but use that Q&A button bar to submit your questions. We will record these questions. If we don't get to your question during the session today, we'll post a written answer to the website.

Al Rea: [48:45] We haven't had any questions come in, yet. The way the thing is right now, it'd be on the right hand side. There's a tab for Q&A. Howard, the AFINCH model that you showed was just done on the US side of the Great Lakes?

Howard: [49:08] That's right, the US Great Lakes Basin.

Al: [49:13] Is there anything happening, similar, on the Canadian side?

Howard: [49:19] Not that I know of. I know that there's work on harmonization across the boundary for the WBD and NHD, so I think it would be feasible to do, if they have the NHDPlus connectivity.

Steve: [49:34] All right.

[49:37] [crosstalk]

Steve: [49:37] Go ahead.

Al: [49:40] We do have a question for Dan, I believe, on the stream Internet, if we have a little bit more time?

Sue: [49:48] Dan chatted and said that he had to move on to another meeting, so I believe he's left the call, Al.

Al: [49:55] OK.

Sue: [49:57] We can take the question, ask him offline, and post the answer on the NHD website, on the Hydrography Seminar Series.

Al: [50:03] There were a couple of other questions that we didn't get to, that we'll do that with as well.

Dave: [50:12] I've tried to submit a question, but have had difficulty activating the send button. One thing I was interested in is expanding your database to include information from bioacoustics monitoring with automatic species identification. Do you have any of those type of monitors included in your network?

Steve: [50:42] Is that a question for Dan as well?

Dave: [50:49] Two questions. Do you have access or interest in bioacoustics monitoring as an indicator of stream and ecosystem health, and did you hear me?

[50:59] [laughter]

Steve: [50:59] Fair enough. Yes, we did hear you, Dave. We will forward that question to Dan. He had to duck out to another meeting.

Dave: [51:09] Thank you. What was the temporal resolution of your data, the temperature data that was predicting stream temperatures to 0.6 degrees centigrade? Is that daily values, or annual averages, or what?

Al: [51:23] Is that still also about the NorWeST stream temperature of modeling?

Dave: [51:27] Right.

Al: [51:28] Again, we'll have to forward that onto Dan and get that back later. There was a question, I believe this is about the AFINCH model. To what extent can you link high yield area delineations to groundwater conditions, and how are those linkages made?

Howard: [52:10] In the AFINCH method, there's a underlying regression model, and that relies, usually, on some geology or soil input information. The link was made through a geology layer that had outwash sand as a characteristic that would be a high yield area, and then, the regression model would estimate high yield for those reaches that are dominated by that geology.

Al: [52:46] OK, thanks.

Steve: [52:47] Do you want to do one more question, and then, we'll wrap up?

Al: [52:55] There was another question on the National Stream Internet. We'll pass that onto Dan. There was a question about a deadline for questions that'll be answered. We'll leave the question and answer panel up here for a few minutes, probably 10 minutes, so even after the webinar is finished, we'll leave it open for a little while longer.

[53:21] Allyson has a poll for you that she'll open up now which is just a couple of questions for you. I think three questions for the attendees about what you found useful, and what we could do better next time, that kind of thing. Those will be open, also, for about 10 minutes.

Steve: [53:43] Thanks for attending everyone. I'll point out, we do have another session planned, and the theme of that session will be Elevation Hydrography Integration. With that, thanks to Dan and Howard for the presentations, and thanks for participating. Please do take a couple minutes to fill out the poll questions, so we can try to make this better the next time. Thanks everyone.