Bill Miller

BILL: Thanks, Thom. Good afternoon, everyone. **(Slide 1)** I'm going to go through a population ecosystem model that I developed with my Co-PI, Vince Lamara, from Ecosystems Research, to address a problem in San Juan River and also the upper Colorado basin for recovery of endangered species. **(Slide 2)** First acknowledgements, we had a lot of help on this with contributions from the Southern Ute Indian Tribe for funding, the Recovery Program in the San Juan River, and all the other folks that were doing the monitoring for the data we used in this recent update. **(Slide 3)** So why did we go with an ecosystem model that uses an approach, a systems approach to endangered species recovery?

Endangered species recovery includes a lot of unknowns. We have low populations. The particular species we're working with, Colorado Pikeminnow and Razorback Sucker in the San Juan River, live to be 50 to 75 years old.
So if I were a life history student and I want to understand life history, I'd like to study that animal for its entire life span, so I can start when I was born and I would be dead when it died. So it's long time to do this and a lot of uncertainty in which way you're going to go to try to get those numbers up. So one method to address the uncertainty associated with management actions of recovery of long-lived fish is to come up with a model. It integrates data and expert opinion into a single explicit framework. It integrates the physical and biological data into one model, and it provides a means to simulate these multiple management scenarios in a relatively short timeframe.

So when actions are desired such as augmenting fish populations, changing habitats, changing flows, how do you evaluate that on something that lives for 75 years? It's something where you have to project outward to see if those populations are going to be stable.

(Slide 4) Here are some background and objectives for how we built it, we needed a method for estimating populations for the long-lived species. The management actions include flow manipulations, habitat modification, non-native removal—that's another factor in these systems where non-natives have been introduced—and augmenting populations. We needed to develop carrying capacity estimates for endangered fish, mainly to determine and validate recovery goals that were coming up.

There were recovery goals being produced and estimates made of how many fish in each river system would we need for recovery. So we used this model to validate that there was enough food resource, space, and no competitors that would not allow recovery. We incorporated bioenergetics to represent food dynamics and tropic interactions, which is very important in some of these systems to get to that population level.

The model also provides a tool to critically evaluate those management alternatives and population response over long time periods. **(Slide 5)** Our study area is the San Juan River, Four Corners area of the United States, and goes from Navajo Dam all the way down to Lake Powell, it's about 225 miles for the whole system.

We have a couple of main tributaries. It's divided into geomorphic reaches. So we have geomorphology included within our reach breaks, within the system. We have ten reaches that we have developed for the model. In the original model sequence, we only had six that went from the Animas River down to Lake Powell, which was critical habitat destination. We have done an update—I'll get into that—of what we have done for the expansion, but now we have the whole system within the model.

(Slide 6) Model development chronology. We developed a conceptual model in 1998. This is during the time when flow recommendations were being derived for the San Juan. We had analyzed, as the program, all of the various monitoring that went on. We had seven years of monitoring data, experimental flows, and we were looking at response of those flows from the native fish system. We still didn't have many endangered fish. We had no razorbacks to speak of and about a hundred pike minnow in the system at that time. So we then developed the data for the model, the population productivity data from 1998 to 2001. During that time we also started developing the mechanistic model and bioenergetics models that went into the system.

In 2000, we used the bioenergetics pieces of the model to calculate San Juan River recovery goals for the Colorado Pikeminnow that the Fish and Wildlife Service then put into the recovery goals that came out in 2002. In 2001 to 2005, we went through calibration of the original model—testing, maintenance, and evaluating some initial management actions. At that point, the model was recommended by the Biology Committee and their peer reviewers that the model be used in the program to look at management alternatives, but it was recommended that we update the software because of limitations we had hit with the complexity.

So, 2012 to 2014, we updated from Stella 8—Stella's our computational platform—to Stella 9. The main advantage was going from a one-dimensional array to two-dimensional arrays within the model framework. **(Slide 7)** For the conceptual framework we got for the system. **(Slide 8)** We have physical factors in it. We have bioenergetics, and we have fish populations, those three main components of the model. For the physical factors, we look at habitat area. Run and riffle habitat are used for benthic and vertebrate productivity. That all is developed from discharge data and flows released from dam. We use a weekly time step in this model, and develop habitat area functions that have been measured over time for monitoring.

We use water temperature that we've monitored for the last 25 years on hourly time step in multiple locations within the system. That goes into the growth rates for the bioenergetics in the model. We have turbidity and storm events, which affect benthic productivity, which feed up through the trophic system up to those higher trophic levels.

In this river system, we know that we have these arid systems. We get a monsoon in August. We get a lot of sediment that comes in, very little change in discharge, maybe a 150 CFS change, but we'll get a sheen of silt, fine silt over these productive areas for invertebrates. We've taken data from research on the Colorado River that we've done in about a 5-year research study where we know that the benthic productivity drops to about 80 percent and then over six to eight weeks will come back up if there's no further storm to the full productivity level it was before the storm. So that reset is in the system.

(Slide 9) The bioenergetics trophic structure and data needs were used to do the food web analysis and to grow the fish and look at what kind of populations we'd get over time. We have producers and we have consumers. We validated those relationships with stable isotope analysis to look at the flow of energy through the system, and then we needed energetic demands for each species. Some of that comes from the basic Wisconsin model for bioenergetics that's incorporated. We've also used surrogates for similar species in this model where we didn't have this information. **(Slide 10)** So conceptually, we have the predators on top. We start with periphyton in the bottom and go through macro-invertebrates. We have detritus that comes in to the system, and we know that it goes in o these other species in their food habits through macro-invertebrates foraging, and then up to these other species. All of those food web dynamics are incorporated into the bioenergetics piece of the model.

(Slide 11) Fish population data that we gathered was needed for the bio-energetic feedback—how much biomass is out there, number per mile, length-weight relationships, how fast do they grow, how old are they, what age, total biomass, what's the prey availability. Prey in the system are based on availability of each of the species that is in the proportion that they exist within the system. We also needed fecundity and survival rates. So all that was derived. **(Slide 12)** And what I've talked about are the two species that we're really keying in on, but we have nine fish in the system that we're modeling. We also have four macro-invertebrates, and we have those physical factors of discharge, water temperature, storm events, and habitat. All of that feeds into the bioenergetic model, and all that goes back out and feeds into the population model.

For each of the species we've got in our update, these were the original species we had. We now have placeholders for four more species in case we have other non-natives that have come in. We also have a placeholder for other macro-invertebrates. Right now these are the main macro-invertebrates. If we do some temperature changes to Navajo, we may get a more robust invertebrate community within the system. So the temperatures out of the dam really controls some of that invertebrate community in the upper reach river.

(Slide 13) As I said the computational platform that we've got is for the mechanistic model. We're using Stella modeling software. And this is a graphical interface software with mechanistic relationships. We're also using Excel spreadsheets to transfer data in and out of the model, so it gives us the ability to almost do a stochastic analysis with a mechanistic model because we can have multiple spreadsheets set up with different survival rates, different fecundity rates, because those are unknowns, uncertainties we've got, and we can produce bands of data with that.

(Slide 14) Here's a typical example of the Stella input, and it has sources and sinks. We have the source goes into the stock, sinks on the side. Out here, we have rate limiters. And each one of these, when they're built with Stella, each of these little components, will put a question mark in there, and you have to then put in the equation or the number for that relationship. Each of these in Stella terms is called an entity. To give you an idea of what it takes for a model like this, we have over 280,000 entities in the model. It's something that if you're going to do this is a consideration going from that simple conceptual model that we started out with to something this complex. I think you really need to think about, as Thom said earlier, where are you on that time and money time scale, and what's your objective in the end.

(Slide 15) To show you the interface that we've got with the San Juan model itself; this is what pops up with Stella. We can choose which modules to simulate as far as the reaches. We can choose the length of the simulation. We can choose what kind of mechanical removal we have, what kind of stocking for any of the fish that are in there. We can select which reaches we want information. It gives us the ability to really manipulate pieces of the system. And something I'll point out with some of our slides later is that when we were going through this the first time with our six reaches, we had made a finite system and said, "Oh yeah, no problem. We can do this, and we'll get a certain recovery back with these fish."

What we really realized when we expanded the system because we include Lake Powell—and Lake Powell is not always connected to the San Juan River—you can go downstream but you can't come back because there's a waterfall there now since the lake is down. So we've got a disconnection. In this new edition of the model, we have all the connectivity in the system explicitly placed in here. So it gives us the ability to look at how these fish move around the system, and we actually kind of -- I'll show you in the configuration how we work this to move the fish through the system. **(Slide 16)** We're set up right now for a weekly time step, and we are capable of simulating a hundred years with this model itself. So it's very similar to some of the population viability analysis software that's out there that simulate for years and years to look at extinction probabilities for endangered fish.

We have the sub-model for bioenergetics. We have an individual based model for population, and we expand that to total population. The biomass is used for prey consumption and availability in growth, the growth feedback loop for fish and macro-invertebrates for prey density and consumption—so if they start eating too much, the prey goes down, their populations drop in response. We have a feedback mechanism in here, just like it would be in a natural system. And we're hoping it's close to that natural system. Again, another uncertainty because we don't have the monitoring data on some of those relationships. Then we have dynamic upstream and downstream movement for all species at all life stages, the Pikeminnow and the Razorback, their life history. Their eggs go into fairly clean gravels. Two or three days later, they emerge, they drift downstream. This drift takes them down into Lake Powell, then they move upstream as juveniles, and the adults move upstream and downstream. **(Slide 17)** We have those movement factors within the model itself.

The model has linkages for each of these modules, so with this new version we can simulate the entire river at one time. The older version that we had, we had to do a single reach, we had to go out to Excel, we had to then manually change the data, go into the next reach. We'd go all the way downstream, then we'd change everything, go all the way upstream. So it took about two days to do a 10-year simulation. And a 10-year simulation was a limit because at that point, we had maxed out Excel with the number of data points we had, and we had maxed out Stella because that was all that it would do. So now we've got it to where we can actually simulate these 10--year simulations with all these modules, and it takes about five minutes to run them all. So it's much more efficient as far as the computational platform.

(Slide 18) To drill down into those modules, each module is made up of invertebrates, fish, prey availability, and physical processes. If you drill down within each of these—and I'm not going to do that here because it gets too complex—you start expanding into all the invertebrates, all the fish, and all the values that link together with bioenergetics. **(Slide 19)** The spreadsheets that we have contain all of the input data, and there are also spreadsheets that are produced for all of the output data. So everything is within Excel back and forth when the model executes. **(Slide 20)** For model calibration it is very important in some of these population models to know where you're at. It was an iterative process that we used for multiple model runs. We started in the '98 version of modeling with our '98 data, up and tried to calibrate to the 2001 population as we got for the river. In this update, we now have 15 years more of monitoring data, and the program monitors all fish species in all habitats so we get a complete system. We now have 23 or 24 years of monitoring data on the San Juan, a very rich database, which as Thom and others have said, that's pretty rare to have. So we're very fortunate to have it. In the new calibration we adjusted to match the management that had occurred in 2002 to 2013, we've changed mortality rates, hatching success, and downstream and upstream migration and input the yearly changes in management as we went along to calibrate them all.

In the San Juan, one thing we don't have are river-wide population estimates. We have population estimates for portions of the river itself. So we have to calibrate by reach with our model to life stages of our species. We don't have many adults in the system yet. We still have juveniles. It takes seven to ten years for this fish to become adults, so we're now at the phase where we're starting to get adults in the system. **(Slide 21)** We calibrated to juveniles and we're trying to get general trends so you can see the up and down line is the data from Utah Division of Wildlife in their lower reach where they had some data collections, and we were trying to be within their confidence intervals on the population aspects. We matched most of those, especially in the out years we're getting pretty good. For a model with a lot of assumptions in it, we thought we're not doing too bad.

(Slide 22) We then went through validation, where we took that calibrated model, and then we have another data set that we can look at for validating our calibration, so another independent data set another reach of the river. **(Slide 23)** And again, we used iterative runs through this to come up with the validation, but in our initial runs through it, we were hitting the numbers still pretty well within their confidence limits. And they have wide confidence intervals because we have a lot of movement in these fish in and out of river reaches. From a population estimate standpoint, we're probably not meeting all the assumptions of immigration and those sorts of things.

(Slide 24) We're also looking at validation against channel catfish population estimates. They do complete the estimates to determine how well their non-native removal is working. It gave us some information where we could start with initial catfish populations in 2002 and let the model run and see where we were in 2010 through 2013 with their population estimates. And again, we're still in the same range as they are with these population estimates. So at that point, we said, "Well, we can now look at some preliminary management scenarios."

(Slide 25) We looked at mechanical removal with a hypothesis that non-natives were limiting endangered species and augmentation, how many fish do we need to put in? What's the result of putting a certain number of fish in? And how long do you stock? And then with the river reaches, we can test longitudinal connectivity. So all those things were looked at when we went through this analysis. And I'll run through now some results of those runs so I can just show you what the applicability of the model is.

(Slide 26) With the Pikeminnow, we have four curves that look like two. The top two or the upper one are stocking, continued stocking and either removal of catfish or no removal of catfish. The one where it drops off is no stocking and removal and no stocking and no removal. So what we've determined, again, looking at this, is the catfish were not impacting the Pikeminnow in the system. There've been other studies independent of our model that are coming to that same conclusion that there's enough resource out that we don't have that interaction to any extent.

And one thing I'll point out—and I'll get to it later too with sort of the lessons learned from this modeling—is that when you pick a software for some of these systems software, you have to live with their inherent displays or limitations. So some of the iterations we get, we get some kind of strange-looking results because of the way they pulse fish through the system. They're used mainly for systems analysis for manufacturing. They're not used that much for the natural resources life history type studies.

(Slide 27) We then looked at what kind of mechanical removal could we do and make a difference with catfish. The top line is our current removal, and so we're taking approximately 40,000 catfish out of the system a year and basically managing for maximum sustained yield. So, 40,000 come out, 40,000 go in.

We asked "what if we increase the removals up to 10 times of what we're doing right now? Right now they're doing about nine trips with four to six electro-fishing boats each year. Those trips are a week long to remove catfish. So think of that going up 10 times. We can really depress the population. That's the lower one. But when we stop, where the vertical bar is, then all those populations come back up. So it's an exercise to determine if we are making a difference. The previous model was useful to determine if are there problems, and this one will hopefully guide some of the future activities with the program to see if we need to continue this or increase it or not.

So, just an example of what Stella does. **(Slide 28)** This is Stella output from the weekly model, so you see this whole jagged edge of where it's pulsing fish in. These are adult Colorado Pikeminnow. We had no stocking and we're not letting fish return from Lake Powell. So the lower line is labeled number one. That's for all the reaches upstream of Lake Powell. And that's approximately what we're guessing the population is in the system before we ever started the augmentation. So it's right around 100 fish.

The top line—that's number two—those are the number of fish that are produced by the fish in the upper river but the larvae drift down into Lake Powell and never come back out. **(Slide 29)** So we have this big source of fish in Lake Powell, if they're not eaten by stripers, walleyes, smallmouth bass, and everything else that's in Lake Powell. So we don't know what type of mortality is from the lake. If we allow fish to come back from Lake Powell, with no stocking—but we return them, there's a pathway for them to get back, we get the blue line, which is all the reaches upstream of Powell again, so the lower one shows we're about 400 fish river wide, 300 adults river wide for the long term, and we've simulated this out to about 75 years, and that line stays stable without decreasing. The upper line is a combination of what's in the upper reaches as well as Powell. The numbers drop down but we're getting more fish out of Powell, we're getting to what we had predicted originally for the recovery numbers, the recovery goals with the model itself. The model is showing that we need that connectivity within the system—again emphasizing some of the components from the Instream Flow Council, and connectivity is an important thing to understand within your system.

(Slide 30) Some lessons learned from doing this. Selection of existing software packages may limit model flexibility, and there are a variety of these systems modeling software out there. You have to think hard of what I want to do. When we went into this we thought we're just going to do this once and we'll move on, it's not going to be a long-term project. Well, this is 16 years later and we're still trying to build things, refine things, and use the model productively within the program. As to our original thought with using the software we picked, we probably would have either programmed it ourselves with Visual Basic or something like that, so that we can manipulate a little more to get specifically what we wanted, or we might have looked at some other programs. At the time we were doing this, Stella was one of the main programs that was out there that was available off the shelf to build this stuff. It's data intensive but we have a system that's data rich. We have many years of monitoring data of all species. We have long term fish monitoring that's been done, we have habitat modeling and monitoring that's been done. So, in our system, it was an easy choice to try to do this. We had the data that was available. If you have to go out and collect it all, and that's a big effort, you'll need a lot of help. It's going to require multi-year data sets to reduce the modeling uncertainty, and then you're still going to have unknowns.

Dealing with some of the more common species might be easier, but with endangered species, we have a lot of things we still don't know. We think from all this effort that we really concluded that we can use this as a tool to assist in the evaluation of management actions. So the model does not replace professional judgment, interpretation and integration. You still need to do all that with a group of experts with the species and understand what's going on. Where there's lower confidence in input data or in the relationships that lower confidence in that data, it increases the uncertainty of accuracy of the long-term population projections. So if you don't know what the survival rate is from egg to larvae to age-0 to adult, you're going to have a wide range of what may happen out there. We've been lucky we've got a good database and a good literature history of investigations in a similar life stage for the populations or species we've studied.

(Slide 31) Refined input data sets for the fish populations would provide a higher confidence in our model output. The one thing we don't have yet are river-wide population estimates for any of the species that are out there. So we're doing it piecemeal and trying to hit those pieces where they've shown up. And at this point with our calibration validation value, we'd be much happier if we had a river-wide number to calibrate against.

What are some the data needs that we still need for this to really move it forward? It's a large, complex system. Endangered species require the cooperation from many agencies and other groups for the data collection. So in ours, we've got 200 plus miles of river. There are 10 entities collecting data. We have a monitoring program with monitoring protocols that tell everyone how to collect it but you're still trying to assimilate all the data. River-wide population estimates would be needed so that we can better refine our model and to make sure it's predicting in the long term what's going to happen. We need data for retention of larvae by reach. So as these larvae drift in the Colorado on the other side of Lake Powell, they have a lot of low-velocity areas that these larvae can eddy out into and grow. The San Juan has the steepest gradient of all the rivers in the Upper Colorado Basin. In two days, they're all the way down into Lake Powell during the time of year they spawn, so we have to either make more habitat or figure out how much to retain within each of these reaches.

We know that juveniles and adults move. We've got tagging studies on that. But we don't have good information on how far they move, what times a year. We know a little bit, but more would be better. And again, population as a function of habitat for key life stages. One thing that I think everybody in the in-stream flow arena has struggled with for years, if I say I have X amount of habitat area, what does that translate to populations? And with that, that's all. Thank you.

THOM: Okay. We'll take a break. I just want you to think about -- we're going to come back and have a contest. Come back in, and I want you to tell me how many dollars you think was spent over this time to do the Stella model. And the one that gets closest, we're going to go to your system and develop this model. So we're running about 30 minutes behind because of lunch or about 15. So can we compress this down to 15 to 20 to get us closer back on schedule? So let's try to be quick on the turnaround and get back in. I don't want to keep people later than what you need.