Chesapeake Bay Program
Science. Restoration. Partnership.

Nontidal Network Workgroup Monthly meeting

Wednesday, February 16th, 2022
1:00 PM – 2:30 PM

Meeting Materials:
https://www.chesapeakebay.net/what/event/nontidal_network_workgroup_february_2022_meeting

This meeting was recorded for internal use to assure the accuracy of meeting notes.

ACTIONS

- ✓ Contact Bob Hirsch (USGS) with any further questions on WRTDS-K at rhirsch@usgs.gov
- ✓ Topics to prioritize for March:
  - Station optimization of the NTN (Qian's progress)
  - Ideas/approaches for merging historical discrete data with new continuous data (Doug Moyer, Data Integrity WG involvement)
- ✓ Topics to prioritize for April:
  - Quarterly round robin on sampling efforts, program challenges in field/lab, field audit needs, data delivery needs.
  - What is the status and what work is needed to improve historical data integrity with nontidal stations
  - Ideas/approaches for merging historical discrete data with new continuous data (Doug Moyer, Data Integrity WG involvement)

AGENDA

1:00 PM Welcome and announcements

Peter Tango (USGS) welcomed the meeting attendees and presenters and underscored the importance of the continued improvement and use of tools and techniques such as those being presented on at the meeting for telling the story of watershed conditions that nontidal network data gives us. Doug Moyer (USGS) explained that Chris Mason (USGS) has taken on the computation of RIM loads and trends, collaborating with the PA Water Science Center and MD Science Center for that annual computation.

1:05 PM WRTDS-K – WRTDS enhancement for improved load estimates – Bob Hirsch (USGS)

Bob Hirsch presented on what is Weighted Regressions on Time Discharge and Season – Kalman filter approach (WRTDS-K), and why one would want to use it. He explained that Kalman filter is a common approach in time system analysis and linear systems thinking, and this takes its concept from Kalman filtering although it's not really a Kalman filter. WRTDS was designed to describe trends, but sometimes the question people have is not about trends. For example, you may want a time series of inputs to a water body, such as Chesapeake Bay, to test a deterministic model to test the conditions in the Chesapeake Bay. There may be an ecological and water quality model of the Bay, and the people who run those models want to know the actual inputs as best as possible on an annual, monthly or daily basis. The flow normalized product of WRTDS is not useful for their purposes of seeing how well their model reproduces the actual history that's unfolded for water quality in the estuary. Another reason one might want it is you might want a time series of outputs from a watershed to test a deterministic model of pollution outputs from a watershed – seeing how well the model does on estimating actual outputs/loads coming out of any basin. WRTDS would be the right tool for evaluating the model of the watershed and how well it works. Or, may want to do mass-balance calculations for a reach or reservoir (for example, Conowingo reservoir). Conowingo is like a major Best Management Practice (BMP) – it's a trap. The characteristics of that trap have changed a lot over the past 2 decades. To understand changing conditions, one needs to understand how that trap efficiency is changing over time. To do that, one needs to model at a daily timestep the inputs and outputs of the reservoir reach, and WRTDS is the perfect tool to do that. Another use of WRTDS is to produce now-casts of operational use and be able to say what the loads are today, or the last few days.

Bob said that we don't want to use the flow-normalized flux for these other kinds of questions, we want the best estimate of flux that analysts can give you for a specific timescale. The main idea is that when you think about estimating the concentration, and from the concentration a load, on a day for which we have the sample the logical thing is to use the data to say that's our best estimate of the load today. The data are always better than any model. For a day that's just before or after a sample is taken, particularly if the flow conditions didn't change much around that sample, we want to use a load number close to the sample number because we expect concentrations will only change a little day before or after (barring major events such as storms). We want to make a lot of use of sample data. We can learn a lot about what's going on by looking at the residuals (errors) of the WRTDS model and use that to help fill in the record for days that we don't have a water quality sample.

Bob went over a case study using WRTDS – the case study was total phosphorus (P) at the Potomac River at Chain Bridge, Washington, D.C. from July 1996 to January 1997. In September 1996, there was a 20-year flood event. The line on the graph is the WRTDS model of what the concentration was, and the dots are the observations. One would expect that the dots would be equally above and below the curve because the WRTDS model is an unbiased model and in the long run the positive and negative errors balance each other. Particularly after the flood, though, there were a lot of negative residuals – in other words, the observations were lower than what the model said they should be. This can show how wrong the model can be for any given slice of time. There's a mechanistic reason this error should be there – the flood of

September 1996 probably washed off a large amount of the readily available phosphorus in the watershed, and on subsequent days, there was less phosphorus than we normally expect at that time of year with that flow. However, over time new phosphorus from decaying plant material becomes available.

How can we better our estimates of what happens? We can look at all of the residuals from this entire data set. The 1996 piece is hardly noticeable. Overall, the residuals are symmetrically distributed around the 0 line, as they should be. There are departures, including from a period of drought. Why might they clump the way they are? Because of high flow and low flow events. In a large watershed it could arise because one particular high flow event gets a lot of water from one subbasin which may be a small or large deliverer of phosphorus and some other storms that were coming out of another basis. That's only an issue with large basins like Potomac or Susquehanna. The errors tend to be persistent.

How to construct the most accurate daily time series? If you have an observation on a given day, use that observation, not the model. For non-sampled days, use 3 pieces of information: the preceding sample (the residual calculated), the residual from the following sample, and the WRTDS model. All 3 make up the WRTDS-K estimate for any given date. Let's say the sampling is far apart in time. In that situation the preceding and succeeding samples are going to be uninformative and you'd want to use the WRTDS model heavily. However, if the preceding and succeeding samples are close in time, you'd want to use those samples predominantly and not the model as much.

The WRTDS gives us a model of the mean and variance for each day. Between observations, the residuals are modeled as an autoregressive lag-1 process with correlation coefficient typically 0.9. Bob showed the results from the same time period (the case study) showing the difference between the samples, the WRTDS, and the WRTDS-K, showing P flux. He showed the year 1991 as well, showing how the WRTDS-K stayed close to the sample observations. The model can also handle where the data is reported as less than some reporting levels. Bob showed publication titles which provide the mathematics for the WRTDS-K model. He also said that in some cases WRTDS-K makes a big difference in flux estimates compared to WRTDS and in some cases it only makes a minor difference, but experimental studies show that is overall better than regular WRTDS estimates.

Bob ended the presentation by reiterating that this is not about trend studies, but about having the best possible record of what happened for any given day, month or year.

Peter Tango asked if we have available comparisons like Bob showed with the Potomac with WRTDS vs WRTDS-K. Bob replied that that information is all there to be done, it's just a question of how much detail one wants to look at it. Doug Moyer added that Chris has a couple of comparison plots and has performed the comparison this past run, but also if you want to dig into a particular site or constituent they have that information. There's some that have significant change as a result of application of Kalman, while some don't have as much change. Qian Zhang commented he recalled there was comparison of WRTDS and WRTDS-K in the paper that Bob mentioned. In general, the difference is minor, though it depends. When they look at annual scale or long-term scale the difference is minor, but looking at the short term scale the difference can be larger. Joel Blomquist (USGS) commented that we have already adopted this technique for the hypoxia forecasting that we do on an annual basis and it was also part of the development of the new hypoxia modeling system that was published in Scavia and others a year and a half ago. The WRTDS-K is being adopted as a standard technique for looking at the true condition fluxes in RIM and other programs.

1:30 PM RIM Load and Trend Results through Water Year 2020 – Chris Mason (USGS)

Link to report - Nitrogen, phosphorus, and suspended-sediment loads and trends measured at the Chesapeake Bay River Input Monitoring stations: Water years 1985-2020

Chris Mason presented on RIM load and trend results with an application of the Kalman method. Doug Moyer introduced Chris as a team member within the load and trend Chesapeake Bay work within the VA/WV water science center. Chris has taken on RIM analyses across tidal and nontidal work.

The total nontidal network is 123 stations across five states. The RIM stations are 9 representing 78% of the total Chesapeake Bay watershed. We try to target up to 2400 samples a year especially focused on storm impacted sampling and high flow conditions. Chris gave an update on what's new for 2020. The nontidal network as a whole is in progress. They implemented the Kalman method for all loads and that includes 5- and 10-year yields. They're remaining with flow normalized WRTDS method for short and long term (20 and 36 years respectively). These are published on a USGS science base data release.

Chris showed Kalman method. They see varying shifts based on the constituent. For total nitrogen, there wasn't much of a shift, although this will vary across sites. He showed suspended sediment for James River at Cartersville. For 1998, there's a large difference between WRTDS and WRTDS-K methods, but for 1999 there is less of a difference between methods. The grouping of sample visits affects the estimate from one point to the next. In the graph you can view how the sample visit days are pulling down the model when you use the Kalman filter. They underwent an arduous QA process when they started implementing it last year. These aren't published, but readily available and they're looking at it prior to data release. For the first time they were publishing the Kalman loads they were looking at the 90% confidence band between the Kalman loads and what it would be if they didn't use the Kalman filter. Going forward for the 2021 RIM publication they'll do the 2020 RIM Kalman load vs 2020 Kalman load which is more in line with how they QA their flow-normalized loads.

These are the flow-normalized fluxes, long term 36-year and short term 10-year trends. They have percent reduction and percent increase and non-detectable trend based on a likelihood estimate based on 90% confidence intervals that they're confident there is or is not a trend. It's another way to look at a P value; 33-60% probability of a trend. They decided to label as likely that there is a trend as there isn't a trend. There may be a trend; there almost certainly is, but they're not as confident in saying. One asterisk is likely, two is very likely, three is extremely likely. After all the re-samples and the bootstrap estimates and confidence bands throughout the time series, they're confident in the upward trend.

Chris showed some key points from the data release. Overall, they're seeing long term and short-term trend positive results in the Susquehanna, Potomac and James. They're seeing almost the same for short term for all three in Phosphorus. The Potomac doesn't show an improvement; it may or may not but they don't have enough confidence to label it as improving. For the Susquehanna-Conowingo, they've observed the same pattern for the last 3 data releases. The short-term trend result is 23.6% reduction with extremely likely confidence, and the long-term trend result is 3.2% reduction with likely confidence. Chris explained the internal QA process they do for all sites/constituents at 90% confidence intervals and showed a comparison of the 2020 results to the 2019 published results using the Susquehanna at Conowingo flow-normalized flux as an example. They try to figure out problems with the input data sets, or issues with the decision tree, or errors with the scripts. If they see something egregious this will tell them. With any new data, things will change and so if they see something suspect they can go back and figure out if it's real or an error. They have this QA process for every constituent and site if anyone is interested. Chris then showed a couple more examples of comparisons at different sites. His presentation has slides from these sites.

The Patuxent was a well-researched site and well-known point source issue. There was a treatment upgrade and reduction has continued ever since. The same with total phosphorus and the same trend direction for both constituents for both time periods for both sides. Everything is extremely likely, as well. For the Potomac at James Bridge it follows the same trend, short and long term, and same with phosphorus. Chris noted that visually it looks like a reduction and it is close to the line on the probability scale, and we're not confident to call it one way or the other.

Chris showed the Virginia sites. They show a long-term reduction similar to the others. However, there is more an increase in the short-term. For the Pamunkey, you can see it's close for the long-term. Any shift up or down could put it in another bracket. Most of the estimates, particularly for short-term, are very or extremely likely in their estimates. The same thing is seen with TP. There are more increases in the short-term in Virginia. Chris showed freshwater flow and 2020 nitrogen and phosphorus loads. 2020 was a very middling year as far as stream flow to the Bay – it was 2% below the mean. He emphasized the separation and difference between the loads and the flow-normalized trends, and how in the past they used the loads to inform the trends but now they're separating them completely. When he gets the trend, he's not using the Kalman load, he's using the load originally derived from the WRTDS model and not applying a Kalman filter on that.

Chris showed the landing page for water-quality loads and trends at nontidal monitoring stations in the Chesapeake Bay Watershed, which can be found at the web address cbrim.er.usgs.gov. The most updated trends and loads can be found here. RIM will be quicker this year than in past years. Chris will start on RIM 2021 in a couple months. All this will be updated. As they get new data, the values are retrofitted to the new data and they archive the previous results – they'll still be available but just archived.

Qian Zhang (UMCES) asked, will there be a similar update this year for 2021? Chris replied yes. Qian asked, and this will be separate from NTN update? Chris responded that he is behind on NTN 2020, the NTN 2020 will be published in 2022. The RIM 2021 will probably be published close to when the NTN 2020 is published. NTN in the spring and RIM maybe a month or two behind. Qian also commented on the slide showing the 2% below the average flow – he commented that he noticed the flow is 2% below the mean but the load for TN TP is much smaller than the mean. That means we have less load on a per flow basis than previous years – which is good!

Doug commented that they are working hard to get the 2020 NTN loads out and hoping for a March time frame to release them. Then they'll start right away on the 2021 RIM loads and trends and that should come out in May or June. Chris added that it is a temporary delay for the NTN 2020.

Bob Hirsch commented that how the flow normalized trends results change with the acquisition of new data. You can see that 2011 was an enormous year for phosphorus especially in the Susquehanna, and 2011 particularly in Susquehanna but elsewhere as well was followed by a series of relatively dry years. The way that WRTDS operates is that it doesn't know much about how the system is behaving at the higher flow levels subsequent to 2011 because there's been non opportunity to observe it. When 2018 and 2019 rolled around, they were wet years, and at that point the model becomes much more grounded in reality because now we have high flows in 2011, 2018 and 2019 and that will cause those curves to shift towards the most recent part. At this junction because we've now seen a very wet year, a series of dry years and then relatively dry years, and we can have a fair amount of confidence those are meaningful numbers because they are pinned down between a number of dry and wet years. When it acquires new data, it can result in significant changes in the result and we get more confident in the results.

Doug commented that they had a lot of re-evaluation of their whole process for the NTN, working with James Colgin at the PA Water Science Center starting with observed data. What Mike Langlin had initially was a SAS database, and that whole code was re-written in R. That caused them to re-evaluate how they handle data, QA, etc. There was stability in the dataset but also areas they needed to tighten up. They also switched to the Kalman method that Bob showed and had modifications to the database. The integrity of their project has increased with all of this, but it took some time. Now that they have the process down, it will be faster in the future.

1:55 PM **Final adjustments to PSC monitoring report – Peter Tango (USGS)**

Peter Tango gave an update on the PSC monitoring report with a focus on the nontidal network specific needs and thanked everyone for their contributions on monitoring needs. Peter emphasized this is not a one-time exercise but an on-going process.

Discussion:

Tom Parham (MD DNR) asked when you meet with the PSC and lay this plan out, are you planning on meeting with them every year as new things pop up? Peter responded that this was a first report back to them based on their questions and interest, they anticipate this going to the PSC and back to the MB and expecting to update the PSC on how the partnerships are forming and updating on outcomes that are still figuring out the exact infrastructure, etc, that they need. So this will be a conversation that goes on over the course of years.

2:15 PM **Survey on key topics for future meetings**

What topics should we prioritize for March?

Topic	# of votes
BMP monitoring discussion with Chesapeake Monitoring Cooperative (NFWF-based workshop findings)	5
BMP monitoring update on USGS-EPA-NRCS projects plus climate influence on BMP efficiency	6
Station optimization of the NTN (Qian's progress)	9
Station optimization: Priorities for station reduction (if needed) identified by each jurisdiction.	4
What is the status and what work is needed to improve historical data integrity with nontidal stations	4
Ideas/approaches for merging historical discrete data with new continuous data (Doug Moyer, Data Integrity WG involvement)	9
Quarterly round robin on sampling efforts, program challenges in field/lab, field audit needs, data delivery needs.	3

Ideas for GIT funding proposals supporting the program's assessment & reporting needs.	4
Other	NTN Sample collection training needs

What topics should we prioritize for April?

Topic	# of votes
BMP monitoring discussion with Chesapeake Monitoring Cooperative (NFWF-based workshop findings)	3
BMP monitoring update on USGS-EPA-NRCS projects plus climate influence on BMP efficiency	5
Station optimization of the NTN (Qian's progress)	3
Station optimization: Priorities for station reduction (if needed) identified by each jurisdiction.	3
What is the status and what work is needed to improve historical data integrity with nontidal stations	6
Ideas/approaches for merging historical discrete data with new continuous data (Doug Moyer, Data Integrity WG involvement)	6
Quarterly round robin on sampling efforts, program challenges in field/lab, field audit needs, data delivery needs.	7
Ideas for GIT funding proposals supporting the program's assessment & reporting needs.	2
Other	NTN sample training and auditing needs

Attendance:
Amy Goldfischer (CRC), Doug Moyer (USGS), Peter Tango (USGS), Cindy Johnson (VADEQ), Lucretia Brown (DOEE), Tom Parham (MD DNR), Chris Mason (USGS), Mark Brickner (PA GOV), Tyler Shenk (SRBC), Breck Sullivan (USGS), Curtis Schreffler (USGS), Kristen Heyer (MD DNR), Qian Zhang (USGS), Tammy Zimmerman (USGS), Mark Nardi (USGS), Jamie Shallenberger (SRBC), Brianna Hutchison (SRBC), Carl Friedrichs (VIMS), Alexander Gunnerson (CRC), Robert Hirsch (USGS), Ellyn Campbell (SRBC), Jon Dillow (USGS), John Wirts (WV DEP), Durga Ghosh (USGS), Joel Blomquist (USGS)