Transcript for Environmental Compliance History Online (ECHO) Webinar: Introduction to the Modernized Water Pollutant Loading Tool

November 28, 2017

Good afternoon and welcome, everyone. Thank you for joining us today for a presentation on the introduction to the modernized Water Pollutant Loading Tool.

Before we get started, I just wanted to review a few housekeeping items. Audio is available for this presentation through your computer's mic and speakers or by telephone. Your call-in number as well as your access code is in the "Control Panel" box on the right-hand side of your screen. All attendees have been muted to minimize background noise, but if you have a question during the presentation, please type it in to the "Questions" box on the upper right-hand side of your screen. We'll have a dedicated time to answer questions at the end of the presentation. If you're experiencing any technical difficulties, please contact us and we will try to troubleshoot the issues.

Note that this presentation will include a survey that will appear on your screen once today's presentation concludes, as well as some brief polls later on in the presentation. In both instances, your feedback is very important to us, so do please make sure to fill these out when they display on your screen. Thank you.

Great—hello, everyone. This is Carey Johnston, here in Washington, D.C., in the Office of Compliance. Today we'll be talking about the new, improved features of the Water Pollutant Loading Tool. What I'll do is I'll be giving a demo and showing you some of the improvements that we've made. As noted, we'll ask a couple questions to get your feedback, and we'll also provide some time to answer any questions you might have.

But before I do any demos or take questions, what I'd like to do is just give a quick, high-level overview of the tool and the data that supports the tool. I'll start with a quick overview of how EPA and states collect and manage wastewater sampling data from sewage treatment plans and industrial facilities; I'll give a demo; and then, as I noted, we'll take questions and answers.

So, what are we talking about? We're talking about water pollution. Under the federal Clean Water Act, industrial facilities and sewage treatment plants are required to get a permit in order to discharge their wastewater.

These permits, called National Pollutant Discharge Elimination System permits (or NPDES), require facilities to sample their wastewater and report the pollutants to either the state or EPA, whoever is their permit writer.

These wastewater pollution data are submitted on forms called Discharge Monitoring Reports (or DMRs). EPA policy and actual regulation require states to enter these data into their data systems and to share these data with EPA's national data system. EPA and states user DMR calculate permit effluent limits, identify and target effluent limit violations, and also to improve watershed modeling for better protection.

An important consideration in assessing the potential impact of pollutant discharges on human health and the environment is identifying the amount of pollution discharge: pounds per day. This can be either reported directly on the DMR or calculated from the data on the DMR. For example, if the facility's reporting concentration and flow, you can multiply those together with a conversion factor to get pounds per day.

Separate and apart from the Clean Water Act, EPA also collects wastewater pollution data from industrial facilities as part of the Toxics Release Inventory program (TRI). The current TRI toxic chemical list contains over 593 individually listed chemicals in 30 chemical categories. This is unique in that the data comes directly to EPA; it only focuses on toxics and only comes from industrial facilities. The tool that we have and which I'll be showing allows you to search and sort through this TRI data as well.

We created a tool, back in 2010, where we allowed people to very easily search on who is the discharger of interest, what is the pollutant of interest, when was the discharge taking places, where was the discharge, and how is being discharged. We have a URL listed there; in the demo I'll show how you can get to that site without having to type in that long URL—we'll work straight from ECHO.

The tool that we developed is designed for two main audiences: members of the general public (concerned citizens, researchers) who like quick and easy access to wastewater pollution data, as well as technical users (permit writers, watershed modelers, regulatory agencies) that like to do more in-depth analysis.

As I noted, the Loading Tool was published on EPA's 40th birthday, December 2, 2010. It has steadily grown into one of the Agency's most popular tools. We've modernized the Loading Tool recently to increase user access to data, improve search speed and performance, and streamline our site maintenance. We've simplified the navigation so Water Pollution Search, which I'll show you, combines the legacy site's EZ Search, TRI Search, and Facility Search; Get Data lists tools that generate data files for downloading; and we have Everyday Search for specialized features. So we've simplified the navigation to make it easier to get access to the data.

I noted we have tabs—and this is similar to the legacy site that we had—that allow people to focus on the search that best meets their needs. The first tab, the default tab, Water Pollution Search, is for the general public, or people who want quick access to data—I'll give a demo of that. That's how you navigate through the site, these tabs right here.

This is an example of the EZ Search. When you click that first tab, or the default view, all of your selections are in box 1, 2, 3, or 4. As you can see, this allows you to search on who (that's boxes 3 and 4) is discharging what (so you can search by pollutant) and then 1 is where—where is the discharge taking place? You can also search by when and the source of the data: DMR data or TRI data. That concludes the overview portion of today's demo.

What I'll be doing is giving you a quick overview. What we're going to do is we're going to start from Google. Go up to the URL bar in the very top; you're going to want to go to echo.epa.gov, hit Enter, and that takes you to ECHO. Previously, the loading tool was on a different site; now what we've done is we've integrated it into ECHO.

What you're going to want to do is, on this Analyze Trends button, focus area, click on that, and then you'll see the Water Pollutant Loading Tool search right there. So that long URL you saw in the

presentation, that's going to show up here in the URL . . . hopefully soon. There it is; okay. Here's that long URL—no need to write that down, you can just start from ECHO, use the Analyze focus area, and then you'll come to the Loading Tool.

This is the new home of the Loading Tool—you can bookmark this if you want to get here straight away. We kept some of the popular features, in particular, the EZ Search. The former EZ Search, is replicated here, so most of your search criteria are all on one page.

Let me do a quick search to show you how it works. Sewage treatment plants, also known as publicly owned treatment works—you can search on the "who"—is discharging "what," so let's say we're looking at sewage treatment plants discharging nitrogen. In the "where," let's say we want to focus in on the Chesapeake Bay basin. So you want to focus on sewage treatment plants discharging nitrogen in the Chesapeake Bay basin.

Let's select a time period—let's say 2016. So this is: as of last year, sewage treatment plants are discharging nitrogen in the Chesapeake Bay. And down here, we click on Search.

What we'll get is a list of facilities that discharge nitrogen—POTWs discharging nitrogen in the Chesapeake Bay basin, in decreasing amounts. So we can see a "Search Statistics" table; this gives you some basic information about the facilities included in your search. And then you see a listing of the facilities showing you the facilities that discharge the most nitrogen in the Chesapeake Bay basin in descending order. (And, not surprisingly, you see a lot of the large sewage treatment plants in the basin.)

I just want to point out: we have your search criteria, so if you recall, we selected 2016, we selected the DMR search, the Chesapeake Bay basin. Our plan is—not now, if you see these little "X"s, you can click on it but nothing will happen—our plan is to enable you to click these "X"s and then this page will re-sort. So you'll be able to modify the search or add search pretty easily; that's something you could not do in the past.

Let's do another search. So let's click on New Search. I'm going to walk you through a demo showing you some of the outlier flags that we built into the system. These are very helpful in flagging potential errors.

So why don't we do a search again on DMR. I'll keep the year 2016; I'm going to select West Virginia, and for my pollutant name, instead of one of these categories that I selected earlier I'm just going to select aluminum, right there.

So I've selected aluminum. And I am interested in a particular kind of industry, and I know the Standard Industrial Classification code: this is a refuse system, landfills, so 4953. And I'm looking for what are called "minor facilities": those are facilities that discharge, roughly speaking, under a million gallons per day.

Okay, so that's my search. This is a pretty specific search, but I want to show you some outliers and show you how that works, so I did this search earlier. You can see, again, you get your table showing information about your search, how many majors and minors were included; if you recall I limited my search to only non-majors, and you can see there are zero majors returned.

What I wanted to show you are these blue flags. These blue flags are our indication that something might be going on with the data. It might be that there was an error—that is, someone shared the

wrong data or keypunched in the wrong data, the wrong units—or that there was an extreme event, like a spill, so the correct data was shared but there was a big spike in the concentration or the amount, pounds discharged.

What I'm going to do is I'm going to select this one right here, Mount Storm Power Station. And what you get is a nice little summary of that particular permitted facility; you can see everything that it's discharging, and you can see —I've got this aluminum, it's got this blue flag, so I'm going to click on Facility Loading Calculations, that little link right there. And then, right on the page without having to leave anywhere, we've shown you all of the pounds reported by that facility in 2016. This is an improvement: previously you had to go off to another page.

Because I'm interested in aluminum, I'm going to select this little drop-down arrow. You can see the outfall—so this facility has multiple pipes. For the largest pounds, I'm going to click on View Details.

Now you can see quite easily where that big spike is. If you go down to the table, you can see we've got that blue flag again; you can see the average daily value is usually under 1, and in this one case it's 37, so probably someone forgot a decimal place. Now, the nice thing about this is that you can report an error: click on this, "Report an Error," and that error correction request will go the state water data steward. (That is a person—a state employee—who's in charge of making sure the data's of good quality.) It might mean that, if the data came the facility, the facility has to re-file that report or amend their data; if it's entered by the state—maybe the state keypunched it in wrong, in which case the state would enter in the data.

This is a process that we track; we try to make sure that error corrections are followed up in a timely manner. The nice thing about the Loading Tool is that we can find these errors pretty quickly.

With that, I'm going to pause and turn it over to my colleague to start the poll. Poll question number one, if you could take it away...

Thank you so much. The first poll question that we have is "How often do you use the Loading Tool?"

<Onscreen: Poll>

Please choose from one of the four options that you have up on your screen. We'll just take a brief moment here to let everyone submit their responses, and then when we're ready I'll just go over the results real quick.

Okay, results still coming in . . . okay. All right.

<Onscreen: Poll Results>

It looks like the majority of today's audience hasn't used the tool before, followed by "A few times a year," then "Once a month," followed by "All the time." Thank you all for your feedback.

Well, I'm excited that we've got some never-users on the phone. I hope that you'll take time to explore the tool; it really is a powerful way to find and sort facilities by what they're discharging, how much they're discharging. I'm biased—I think it's pretty cool.

I'm excited that some never-users are on the phone—hope to get you using it more often. And, for those that are using it all the time, thank you very much—we really appreciate it.

What I'm going to do now is go back to the Loading Tool, and back through ECHO, and I'm going to show you a custom search.

This was previously called the "Advanced Search" under the legacy system. So, on this Data Downloads tab, you're going to click on that tab, and then here you can see "Custom Search." "Custom Search" if formerly known as the "Advanced Search."

What you can do is you can select the output in two ways—well, actually three ways, but the main two ways are "Annual" and "Monitoring Period." The difference is, with "Annual" it gives you, for every facility, for every parameter, for every outfall, one row, one load; and then for "Monitoring Period," it's the most detailed level, so for every facility, for every parameter, for every pipe, for every monitoring period, it gives you a load. This is really geared toward the more technical user; we give you the data in a file that you can then open on your computer, but it's not a user interface where you get to kind of search and explore and see the results in nice pretty tables.

Let me just do a quick search to give you an example. Here's the date range. I'm going to select "Monitoring Period." Let me select a relatively small state, Delaware; how about "Major," and I'm going to select "POTWs," those are sewage treatment plants. I can select my time period of any time over a 12-month period. So I'll pick January 2016 to December 2016. So that's one year.

So basically what I'm asking for here is all the DMR data for all large sewage treatment plants in Delaware. If I wanted to focus on a particular parameter, like aluminum or BOD, I could do that here.

Okay—so let me hit "Submit."

These are my options for the data I'm going to get—so these are what you'll see in the download that we're going to get; these are all going to be new columns in a spreadsheet. I'll just take all the data; if I want to get all the data, just click "Download."

Let's save this: "Save as..."; I'll save it to my hard drive; and now we're going to go to Excel. Like I said, this is more for the technical user that feels more comfortable in spreadsheets. The advantage of using the data this way is, you can download bigger sets of data, you can have more complicated queries. So here's the file: this is a CSV, "comma-separated value," file.

Here is my custom download search results. I'll move this bar out, and you can see: here are the monitoring values—or, the monitoring periods. This is for December, February, March. This first one is for Shelbyville; it gives you all of the data—information about the facility, latitude/longitude, and then here are all the parameters that they're monitoring. They're monitoring for biochemical oxygen demand, chlorine, copper, enterococci, solids—and then it just keeps on repeating for each of the months. You can see they have different outfalls—or they have, actually, one outfall, one main outfall, but some facilities have more than one outfall. Let me see if I can find one . . .nope. Let's see—and then the load.

So we give you all of the data including the limits and the measured value; in this case it's kilograms per day, and if there are any load over limit—so if they've exceeded any of their limits, that'll be in this column, column BL.

Since I'm kind of used to this data, this makes a lot of sense. If you're a general user and this this looks scary, don't worry about it: you don't need to use this format if you don't want to. You can always hop

back to "Water Pollution Search." "Water Pollution Search" is the easy way of finding who is discharging, what are they discharging, and where they're discharging.

So I'm going to pause here, turn it back to my colleague; we'll have another poll question, and then see if there are any questions, and we'll take it from there.

Thank you. The next poll question that we have is "What is your primary use of the Loading Tool?"

<Onscreen: Poll>

For this one, you can select more than one choice if it applies for you.

We'll just pause for a second here, just to let some of the results come in, and we'll go over them in just a second here. Okay.

<Onscreen: Poll Results>

It looks like the spread was a little bit closer this time, but it looks like the majority of today's attendees review facility compliance, followed by "Evaluate facilities and discharges near me or in specific areas," and then followed by "Use wastewater discharge data for research," then "Review data quality," and then "Other" as well. If you did select "Other," you're welcome to provide or elaborate any more detail in your chat box in your control panel in your screen.

Thank you for your feedback.

Great.

Let me show you how easy it is to find facilities with violations of their limits. In each NPDES permit, there are not only requirements to monitor and report those data, but also limits: values, levels they cannot exceed lawfully. So let's say you were—I'm going to pick a search, let's say Virginia, and I wanted to find POTWs in Virginia in let's say the most current year, 2017, and only include facilities with effluent limit exceedances. So I click that little box right there, and that should give me a list of all of the sewage treatment plants with effluent limit exceedances in Virginia. So you can see: 90 majors and 217 minors; you can sort the facilities in terms of total top pollutants discharged.

So it's super-easy. My search was on Virginia; if you wanted to do another search, maybe just in your county or in your zip code, you can also search based on that as well. So we can search by zip code, state, or county, or city. You can also search by watershed—I showed you the Chesapeake Bay watershed. So it's super-easy to find facilities with effluent limit exceedances.

I want to take time and also show you that this Loading Tool is becoming more powerful because states are sharing more of their DMR data with EPA. We have a new dashboard, and it's related to a rulemaking we did, the NPDES Electronic Reporting Rule. What we did is, we are requiring states to share their DMR data with us on all DMR filers; you can see in this stacked bar chart on this dashboard how the states with lots of blue—Florida, Georgia, Iowa, Idaho, Alabama (Alabama's looking really good)—you can see all the DMR data for all the facilities in these states are coming in. The colors that are yellow means that we have some of the DMR data, some of it missing. The color pink means that there are facilities where we know there's a facility but we don't have limits and therefore we don't have DMR data.

What we can see is that, over time for each state, more and more of the facilities are in blue there: we're getting more and more of the DMR data. So right now we're looking at roughly 40,000 facilities across the country where for the most part we're getting all the DMR—we're getting the DMR data. So, for about 70 percent all the DMR data, we're missing some DMR data on about 20 percent, and then there's about 10 percent where we're still working with the states to share the data that's needed—the limits data to collect the data.

So I'm showing you this because I wanted to give you a quick view of both the improvements to the tool as well as the improvements to the data underpinning the tool and to show you that this tool, the Loading Tool, is becoming more and more useful.

I'm going to now turn back to the Loading Tool. Again, you click on "Analyze Trends" and the Loading Tool search is right there—that dashboard that I was just at is right down there. So I'm going to click on that. What I'm going to do is, I'm going to mention that we are transitioning to this new Loading Tool—that's why we call it "beta," means it's still in tests—and when we're ready, and hopefully this will be in early December, we're going to make this search official, and then we'll take down the legacy site. It'll be retired. If you have any links, you're going to want to update those links.

We have a modernization page—let's see. Oh, this is about the Loading Tool. Yeah, we've got a lot of resources, we have a lot of resources on the data and all of the services that we're using. We're looking to, as I said, update the site, get it done, retire the old site in December. We welcome any feedback that you might have for us, as well as any improvements that you would like to make. You can always contact us. There's a "Contact Us" page; you can provide your name and email, send us a message. We intend to keep making improvements to the Loading Tool and improving it, so your feedback is greatly appreciated.

So that's kind of the conclusion of the demo. I'd be happy to take any questions, if the moderator could help me out with that.

Yeah, we have a couple questions.

Mm-hmm.

So our first question is: "Is there a limit to the number of results available for download on the custom search?"

That is a good question. I do know, on the custom search, that if you set up, like, "I want all the data for 2017 for the entire country," it will likely bomb out. I forget the exact amount. But if you do come across an issue where either the tool crashes or you get a message saying you need to narrow your search, one way of breaking it up is by region. EPA has 10 different regions and these represent 10—like, Region 1 is New England, Region 2 is New York/New Jersey—so if you were to break it up that way, that will likely get you down below that limitation.

Our next question is: "How often are data refreshed in the Loading Tool?"

Oh, what a great question! The Loading Tool is refreshed every weekend. As states and facilities enter and share their data—so, Monday through Friday they key in data—over the weekend, starting on Saturday, we take the data out of our production database, do the necessary calculations, and then by Monday the data is ready to go.

"What happens if the 'Report an Error' was an actual spill? How is it caught in your system and who addresses it?"

Yeah, so the question is what happens if it's not an error? Like I showed you the facility in West Virginia: What if the value is truly 37 milligrams per liter?

Well, it is the duty of the permitting authority, in this case West Virginia DEP, to follow up and take action to address any noncompliance, like a spill. There are many ways that states can do that; it's part of their obligation as the permit writer.

"What if you are with a monitoring group that finds the data being put into ECHO by a company is inconsistent with your group's findings?"

Well, it's a tough one to answer in that the data that are supplied by the company must be true and accurate, and there is an attestation, a signature that the company makes that the data they're providing is true and accurate. So if a facility is sending in fraudulent data, that's a real concern: this tool won't uncover that. But it is something that we look for; we have a separate analysis looking for fraudulent data. I would note that any group that has concerns about a discharge can always contact their state environmental office or EPA's regional office if they think that there is some incorrect reporting or nonreporting going on.

"What is the difference between a major or minor designation?"

I'll go back to the water pollution search here and show you where that is. Right here: major and minor, or non-major.

If we take the example of a sewage treatment plant, EPA has defined a major sewage treatment plant as any facility—any POTW discharging more than a million gallons per day. That's roughly the size of a POTW for a town or a city of 15,000 people. So if you've got a town of 15,000 or more, that's facility; if you're below that, that's considered non-major, or minor.

So it's a term that's defined in EPA's regulations; it roughly corresponds to flow (not always). It can be useful if you're interested in facilities that are likely discharging the most pollution—selecting "Major" can help narrow your search. But it may also be the case that if you're just looking for everyone in your watershed—you know, if you type in a watershed and you want every major facility, you're going to bring back more facilities.

So, in general, I keep it on everyone. In this particular case, if I restricted it to only majors I would have only brought back the one facility and I would have missed these 16 facilities.

"Can you search for multiple years of DMR data?"

You can, but not at the same time—so, I guess, in a way no. On the Water Pollution Search, you can select year by year; you can do that as well in the Custom Search and Data Downloads.

The reason for that is, we want to make it pretty simple for people to search and sort. So it is a limitation; you can't search on multiple years at the same time. But if you need to do multiple searches across years, Data Downloads is a way that you can get data.

Just looking to see if we're limited—yeah, so, whether it's annual or monitoring period you're limited by year.

"How are the potential error flags created? How do they differ from our option to report an error? Is there an option to exclude flagged data when viewing data, or evaluating compliance, etc.?"

Right. In the background, as part of our weekend processing, we look for those spikes. If you recall from that West Virginia facility, there's that big spike; we look for those spikes, and when we see a spike like that, we tag that value with that blue flag. I don't think—we can limit results to effluent limit exceedances; I'm not sure on the Easy Search we can exclude the blue-flagged values, but I think you can on the Custom Search.

I'm going to Custom Search, and . . . I guess you could limit facilities that have been flagged as potential outliers, but there's no way to—so that might be a feature that we need to build in, like, we want to exclude loads that have been flagged. That's a good comment.

But in the results—this is the CSV file that I downloaded earlier—you can see whether or not it's a potential outlier, and if it's a potential outlier it's say "Yes." In this case, I don't think there are any . . . nope. But it would say "Yes" if there's a potential outlier.

Any other questions?

Yeah, we have several more.

"How could the Loading Tool benefit a user from a company with a NPDES permit?"

Right. So let's say you're an environmental compliance officer, or you otherwise are working for a company, and you want to make sure you understand what's going on. What you can do is you can come down here, under the facility, and if you have the name or the permit number . . . I'll just type in a permit number—each permit has a unique ID; we call this the NPDES ID, so that's the unique ID for this particular facility—so I'm doing a search just on that facility in the year 2017.

And then you can see—here it is, it's the Arlington County Water Pollution Control Plant—and then I can click on it and then this will give me a nice report showing if there are any effluent limit exceedances. If there were any violations you would see a yellow flag; if there were any potential outliers you would see a blue flag. You can also see (and this is really helpful): this facility has a limit on BOD, biochemical oxygen demand, and the maximum allowable load across the whole year is close to 600,000 pounds. And they're discharging about a sixth of that, showing that they're overall well below their limit. So this is a nice way of comparing what they're discharging against their limit. That's a nice feature.

That's one easy way for someone who's concerned about compliance at a particular facility of figuring that out.

"How are nondetects handled in calculated loadings?"

Another excellent question!

We have a writeup of that procedure in the resources and the Loading Tool FAQs, but in general: If a value is reported as nondetect—that is, people sampled for it and didn't find it throughout the entire year it was reported as nondetect, then we assign a value of zero. If however there was at least one

instance of a value reported above its quantitation limit, its reporting limit, then we use that value, the measured value, the measured and quantified value, in the loading calculation; and then for all other nondetects we assume half the detection in the loading calculation, the thinking being that the value was detected, it was quantified, so it's present, so when it's for all other values below detection we assume half the detection level.

"Does EPA plan to expand the watershed categories filter options so that one can study pollutant loads in smaller watersheds?"

Let's take a look at this real quick.

You can search, as I showed you, by major watersheds, or the USGS has watersheds; we also have two programs that focus on large watersheds. I selected the Chesapeake Bay, that was the example—but you can also search by ZIP code, any ZIP code will work, and then that'll do an overlay on one or more watersheds that are included in the ZIP code. Or, if you know your watershed ID—and every watershed has a unique ID; it goes up to 12 digits—you could search that here.

Let me do a quick search showing you how that works. Let's say . . . let me get back to this. So let me do another search; this time I'm going to search by watershed, I'm going to type in our local ZIP code, 22203, I'm going to hit "Search." So with this search, the reason this ZIP code here is different than this ZIP code up here—the top ZIP code only finds facilities uniquely in that ZIP code boundary. The second one that I used, the one below, which I'm showing you here, is all the facilities and watersheds that touch ZIP code 22203. So we've got Blue Plains—there's my Arlington County wastewater pollutant treatment plant—and Blue Plains is being picked up through the unique situation in that the Potomac River is actually from shore to shore owned at this stretch by, or is part of D.C. So, fun fact there.

Next question?

"Currently, entering and improving NPDES data occurs monthly via CDX. Is this taking the place of CDX?"

Oh no. No, this is a way of analyzing the data, of getting data out. What the questioner is talking about is how to get data in.

So, for those people who are familiar with sharing data, we have a system called the Central Data Exchange. You need a username and a password to log in. This is about—what you see here on the screen—this is about getting data into the system. What I'm showing you here, this ECHO tool, is how to get data out of the system. You do not need to log in at all to access the data.

Those are all the questions we have time for today. We can follow up with any remaining questions via email.

Yeah. Already we got some good ideas for future enhancements, and I definitely encourage you to share your thoughts. Whether it's ideas for improving it or if you needed specific help, you can contact us. I really appreciate everyone using the tool and participating in today's webinar.

We know that finding and sorting through the data can be challenging, which is why we create tools to make it easy. We hope it's easy for you, and we hope the data is of value to you and you can use it to see what's going on in your watershed.

With that, I'll turn it back over to the main organizer and he'll take us out.

Yes, thank you all so much for participating in today's presentation. And I just wanted to give you all a quick reminder that there will be a survey that'll appear on your screen in a moment; please feel free to fill that out and give us a little bit more feedback, which is very valuable. Thank you all and have a great afternoon.