



Jul 11th, 4:50 PM - 5:10 PM

Watershed and Economic Data InterOperability (WEDO) System

Rajbir Parmar

U.S. Environmental Protection Agency, parmar.rajbir@epa.gov

John M. Johnston

U.S. Environmental Protection Agency, johnston.johnm@epa.gov

Mike Galvin

U.S. Environmental Protection Agency, galvin.mike@epa.gov

Amber Ignatius

Oak Ridge Institute for Science and Education, ignatius.amber@epa.gov

Catherine Foley

CSRA Inc., catherine.foley@csra.com

Follow this and additional works at: <https://scholarsarchive.byu.edu/iemssconference>

 Part of the [Civil Engineering Commons](#), [Data Storage Systems Commons](#), [Environmental Engineering Commons](#), [Hydraulic Engineering Commons](#), and the [Other Civil and Environmental Engineering Commons](#)

Parmar, Rajbir; Johnston, John M.; Galvin, Mike; Ignatius, Amber; and Foley, Catherine, "Watershed and Economic Data InterOperability (WEDO) System" (2016). *International Congress on Environmental Modelling and Software*. 10.
<https://scholarsarchive.byu.edu/iemssconference/2016/Stream-A/10>

This Event is brought to you for free and open access by the Civil and Environmental Engineering at BYU ScholarsArchive. It has been accepted for inclusion in International Congress on Environmental Modelling and Software by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

Watershed and Economic Data InterOperability (WEDO) System

Rajbir Parmar^a, John M Johnston^a, Mike Galvin^a, Amber Ignatius^b, Catherine Foley^c

^aU.S. Environmental Protection Agency (Parmar.Rajbir@epa.gov, Johnston.Johnm@epa.gov, Galvin.Mike@epa.gov)

^bOak Ridge Institute for Science and Education (Ignatius.Amber@epa.gov)

^cCSRA Inc., Falls Church, VA USA (catherine.foley@csra.com)

Abstract: Hydrologic modeling is essential for environmental, economic, and human health decision-making. However, sharing of modeling studies is limited within the watershed modeling community. Distribution of hydrologic modeling research typically involves publishing summarized data in peer reviewed journals, producing reports, or presenting at conferences. This narrow data distribution process limits interoperability, decreases transparency, wastes leveraging potential, and discourages detailed evaluation by other modelers. These limitations are also experienced beyond hydrologic modeling. For example, an economic analysis for nutrient trading may not be possible without access to detailed hydrologic modeling data. To address these issues, the Watershed and Economic Data InterOperability (WEDO) system provides web services, a database, and web pages to disseminate model data and promote interoperability within the hydrologic modeling community. WEDO lets researchers publish modeling studies with the full complement of input data, calibration parameters, time series output, and associated metadata for easy duplication, integration, and evaluation by others. WEDO greatly increases transparency by making detailed data available to the scientific community. Transparency and documentation are paramount issues within regulatory environments and are increasingly necessary to address data archival policies. WEDO includes a Publishing Utilities desktop application to aid researchers in standardizing their modeling files and assist with data upload. This encourages consistent, organized database management to aid in searchability and data discovery. WEDO data is also presented on the USEPA EnviroAtlas interactive map to allow for spatial query of the WEDO database. The WEDO database allows data download by all users to promote access and ease data integration. WEDO provides a flexible, centralized, and interoperable database for multiple models, as well as web services and web pages for storing, retrieving, evaluating and publishing modeling output as time series.

Keywords: Watershed Modeling; Hydrology; Interoperability; Web Services; USEPA; Data Sharing

1 INTRODUCTION

Discovering and accessing hydrologic and climate data for use in research or water management can be a difficult task that consumes valuable time and personnel resources (Ames et al., 2012). Development of standards for publishing data and models in forms suitable for automated discovery, access, and integration are among the highest priorities of the integrated modeling community (Laniak et al., 2013). A typical watershed modeling study involves a researcher running a model or models on their computer and publishing the results in a paper, conference presentation, and/or report. This paradigm of peer review and publishing (often in restricted-access journals) limits reuse, evaluation, and transparency of hydrologic research. Without direct access to model inputs, including initial conditions and parameters, as well as study metadata for all input data utilized and expert knowledge used in calibration, others experience barriers to reproducing the science. These barriers are also experienced beyond watershed modeling. For example, an analyst who wants to assess market feasibility for nutrient trading in a watershed of interest will likely have to repeat flow and water quality modeling because the analyst may not even know that a study for the area of interest exists. However, even if the analyst is aware of the report, access to the detailed data beyond the published summaries may be unavailable. Another example is that of calibration studies, a vast majority of which are

published in state and federal government reports. These and other grey literature may not be searchable on the Internet. Many journals now require as a best practice that research data for the study be made available online. In addition, managing, archiving, and sharing large amounts of data are essential tasks in ecological laboratories, and detailed data management plans are now required by major funding agencies (Conner et al., 2013). **WEDO (Watershed and Economic Database for interOperability)** is a solution that enables publishing of modeling studies for evaluation, download, and leveraging by others.

Many government agencies have repositories for monitored stream flow and water quality data. EPA's STorage and RETrieval (STORET) and the U.S. Geological Survey (USGS) National Water Information System (NWIS) are well known examples. Although monitoring data are commonly made publicly available via repositories, currently there are no government watershed modeling study data repositories. WEDO provides a flexible, centralized, and interoperable database for multiple models, as well as web services for storing, retrieving, evaluating and publishing modeling output as time series. In addition to model output and results, model inputs and study metadata are also stored in the WEDO database. WEDO achieves the goal of increased transparency of modeling study publishing by facilitating the upload of all data and metadata required to duplicate a study for a waterbody of interest. To publish a completed study, all model input and output are uploaded via WEDO web interface (with a brief description of the model study purpose, scope and supporting metadata), and the study is tagged for inclusion in the EnviroAtlas national stream layer for the lower 48 states. Currently HSPF and SWAT watershed models are supported, but other models are planned including economic models. WEDO is designed as an interoperable database to accommodate multiple models rather than being targeted to one or a few. WEDO is available as beta at <https://www.wedo.epa>. Modeling studies are uploaded to EnviroAtlas every three months for discovery, review and download as a data link to WEDO.

2 IMPLEMENTATION

2.1 WEDO Architecture

The architecture of the WEDO software (Figure 1) has three logical layers: desktop publishing utilities, the WEDO database web site, and the EnviroAtlas web site. Navigation and information flow between the layers is shown as arrows in the architecture diagram. The desktop publishing utilities layer contains utilities for publishing modeling studies on the WEDO database web site from desktop computers. Currently, publishing utilities are available for HSPF (Bicknell et al. 2001) and SWAT (Neitsch et al., 2011) models, with other models to be added as needed because WEDO architecture allows addition of publishing utilities for other watershed models. Publishing utilities have been developed using the Microsoft .NET framework 4.0 and VB.NET programming language; however, publishing utilities can be developed in any programming language. The WEDO web site itself consists of three logical groups of software: database, RESTful web services, and web forms/pages. All software for the WEDO web site has been developed using the Microsoft .NET framework 4.0, MVC5, and C# programming language. A combination of SQL/T-SQL has been used to perform database operations. Microsoft Visual Studio 2015 was used as the Integrated Development Environment (IDE) for developing the WEDO web site. The WEDO web site provides tools to publish summary data to the EPA EnviroAtlas (www.epa.gov/enviroatlas) which is not part of WEDO but makes the data discoverable to a wide audience. EnviroAtlas is part of the EPA GeoPlatform developed on the ESRI technology stack.

The WEDO architecture supports interoperability concepts of discovery, evaluation, and integration. Publishing summarized results as a GIS layer on EnviroAtlas facilitates discovery. WEDO web forms/pages support deeper evaluation of the published data. WEDO REST web services support access and integration to enable duplication of modeling results. Both the data publishers and consumers use the web services to integrate their work flows with WEDO. WEDO manages a modeling study as a reproducible unit for others to use. WEDO supports modelers who want to take advantage of work that has already been completed in their area of interest by making it easy to locate and retrieve all necessary modeling data. WEDO also supports modelers who want to share their studies with others by making it easy to publish their study online. No additional database or language proficiency is needed.

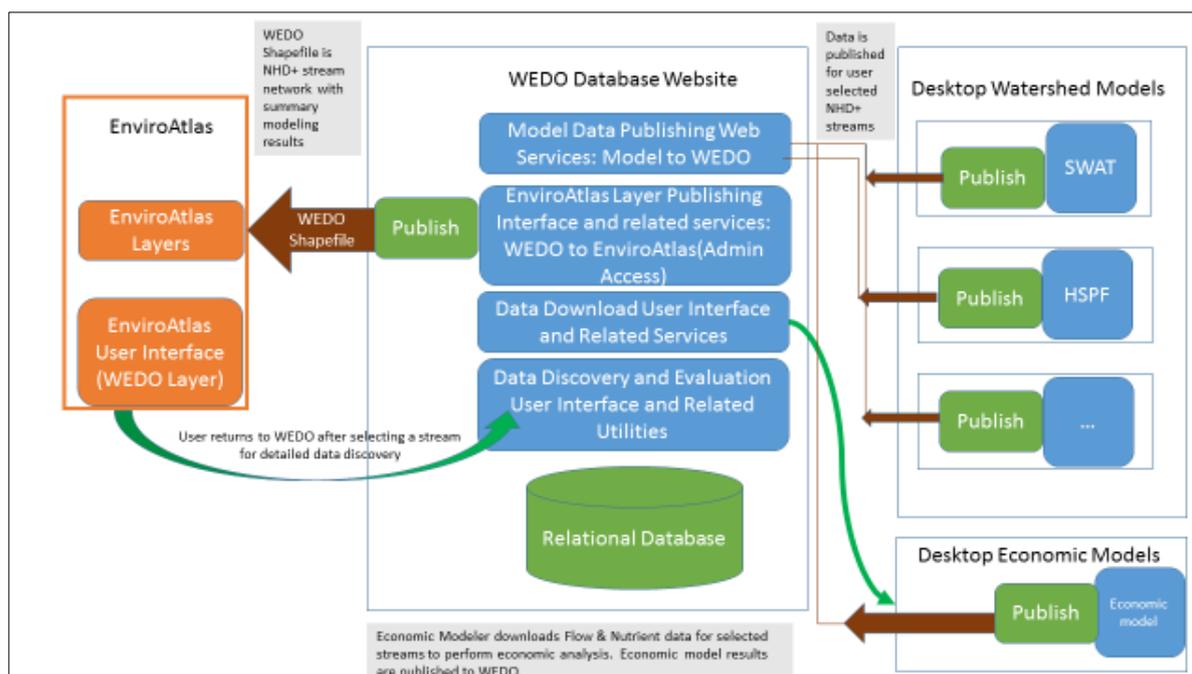


Figure 1. WEDO Architecture

2.2 WEDO Tasks

WEDO supports various user roles (Table 1) to publish, discover, evaluate, administer, and consume data. Currently, WEDO is available within the EPA firewall but we plan to make it publicly available in the near future. WEDO home page (Figure 2) is the starting point for publishing a study. The steps for a modeler to publish a study to WEDO are briefly described. First, the modeler uses a publishing utility provided by WEDO that is downloaded to the modeler's desktop. Once it is opened, the publishing utility presents a series of dialogs to collect information to more fully describe the study. A subset of the dialogs is shown in Figures 3-5. The modeler has the option to either publish the whole study or parts of the study outputs. All model inputs get published regardless of whether the modeler selects all or a subset of outputs to publish. Publishing of all inputs ensures that other modelers and researchers can reproduce the study. The publishing utility compresses all input files into a single zip file. The publishing utility presents a list of all rivers and streams in the output files. The modeler selects river(s) and stream(s) to be published. The publishing utility also presents a list of parameters or constituents. The modeler selects output parameters to be published from the list. The modeler is then instructed to standardize the stream numbers by mapping the stream identifiers to the National Hydrography Dataset (NHD) reach codes (<http://nhd.usgs.gov>). The publishing utility provides a hyperlink to a map containing NHD streams and stream search capabilities along with zooming and panning tools. Finally, all inputs along with the self-selected model outputs (water flow, dissolved, and suspended constituents for example) and metadata are compressed into a single zip file which gets uploaded to WEDO. Metadata consists of items such as author name, simulation dates, author's organization/company, calibration statistics, etc. WEDO unzips the uploaded file, saves the inputs zip file in database, extracts data from the output files, saves the output data and metadata in relational tables in WEDO database. A study on WEDO is labeled "Pending" status at the time of upload. An administrator periodically may decide to publish the studies with "Pending" status on WEDO to EPA EnviroAtlas.

To publish studies to EnviroAtlas, the site administrator selects studies on a WEDO web page, presses a button to generate a Comma Separated-Values (CSV) file, and sends the file to EnviroAtlas team by email; the EnviroAtlas team loads the CSV file to its NHD map layer. Each row in the CSV file contains a stream Reach Code, a list of corresponding water quality and flow constituents, and a hyperlink pointing to the corresponding study in WEDO. The published streams are shown in red color on EnviroAtlas map layer whereas all other streams are shown in blue color. Clicking on a published stream on EnviroAtlas results in a popup showing its Reach Code, a list of water flow and quality constituents, and a hyperlink to the corresponding study details in WEDO. Clicking on the hyperlink takes the user to the details page of the corresponding study. The status of studies included in the CSV

file changes to “Publishing” as soon as the CSV file is generated for publishing. After confirming that the file has been published on EnviroAtlas, the administrator changes studies with “Publishing” status to “Published” status. A study in “Publishing” or “Published” status cannot be directly withdrawn by a modeler because it would result in a broken link from EnviroAtlas to WEDO. Only a study with “Pending” status can be withdrawn, and this can be completed only by an administrator. A modeler can send an email to one of the contacts listed on the contact page requesting a “Pending” study be withdrawn, or to request that a published study be changed to “Pending” and then withdrawn.

Information discovery is a major aspect of interoperability. EnviroAtlas is one of the EPA’s main public-facing decision support tools for communities. Publishing data on EnviroAtlas provides an intuitive means of discovering studies of interest (Figure 6). The other two major aspects of interoperability are evaluation and integration. WEDO facilitates evaluation by providing graphical interfaces to search and query summary as well as detailed data. WEDO facilitates integration by providing data download and upload services that make it possible for anyone to duplicate the results of a published modeling study. WEDO supports various administrative tasks such as publishing a study to EnviroAtlas, withdrawing a study, changing study status, adding another administrator, and responding to user reported issues.

Table 1. WEDO Tasks

Role	Tasks
Modelers	Download publishing utility, run publishing utility and publish modeling data on WEDO, send email to WEDO administrator to withdraw a study
Analysts	Discover streams of interest on EnviroAtlas, visit WEDO web pages to view dashboard, study summaries, study details, download study data
Administrators	Log in, add administrator(s), withdraw a study, generate CSV file to publish studies on EnviroAtlas, email CSV file to EnviroAtlas team, for publishing, change status of “Publishing” studies to “Published,” respond to issues reported by researchers and modelers

3 SUMMARY AND CONCLUSIONS

The project goal was to develop a system for publishing, discovering, evaluating, and integrating watershed modeling data in order to achieve interoperability. WEDO provides a flexible, centralized, and interoperable database for multiple models, as well as web services and web pages for storing, retrieving, evaluating and publishing modeling output as time series. Model inputs and study metadata are also stored in the database. Publishing on WEDO also fulfills the requirements of transparency, especially regulatory requirements of the United States Environmental Protection Agency.

WEDO is novel for publishing watershed modeling studies and differs from other modeling data repositories in features and scope. One example, HydroShare (Tarboton et al., 2014), includes a repository for hydrologic data and models that also goes beyond watershed modeling data. However, the major difference between WEDO and the HydroShare repository is the use of a standardized data structure in WEDO through the use of publishing utilities, whereas HydroShare does not impose a standard. In addition WEDO standardizes stream numbers to NHD Reach Codes. WEDO also leverages the publicly available EnviroAtlas GIS and mapping tool to locate waterbodies of interest as NHDPlus stream segments. We believe the standardized publishing and map-based query features substantially enhance discovery, evaluation, and integration. SwatShare (Merwade et al, 2013) is a repository specific to the SWAT model, whereas WEDO is designed to accommodate modeling studies from any watershed model. Currently, WEDO is available as a beta product within EPA firewall. Readers interested in beta testing are encouraged to contact the authors of this article.

4 ACKNOWLEDGMENTS

We would like to acknowledge the support of the USEPA Office of Research and Development’s Safe and Sustainable Water Resources research program, especially Suzanne van Drunick, Michael McDonald and Blake Schaeffer (of the National Exposure Research Laboratory). Without their support, WEDO would not have come to fruition. Appreciation is also extended to Topic 1 co-leads Christopher

Nietsch and Matthew Heberling of the National Risk Management Research Laboratory for guidance on evaluating the economic feasibility of nutrient trading markets, including the interplay of traditional and non-traditional economic models and watershed modeling. The research described in this document was funded by the U.S. Environmental Protection Agency through the Office of Research and Development and was conducted at the Computational Exposure Division of the U.S. Environmental Protection Agency National Exposure Research Laboratory in Athens, Georgia. Mention of trade names or commercial products does not constitute endorsement or recommendation for use. Although this work was reviewed by EPA and approved for publication, it may not necessarily reflect official Agency policy.

5 REFERENCES

- Ames, D.P., Horsburgh, J.S., Cao, Y., Kadlec, J., Whiteaker, T., Valentine, D., 2012. HydroDesktop: Web services-based software for hydrologic data discovery, download, visualization, and analysis. *Environmental Modelling & Software* 37 146-156, <http://dx.doi.org/10.1016/j.envsoft.2012.03.013>.
- Bicknell, B.R., J.C. Imhoff, J.L. Kittle Jr., T.H. Jobes, and A.S. Donigian, Jr., 2001. Hydrological Simulation Program - Fortran (HSPF). User's Manual for Release 12. U.S. EPA National Exposure Research Laboratory, Athens, GA, in cooperation with U.S. Geological Survey, Water Resources Division, Reston, VA.
- Conner, L.G., Ames, D.P., Gill, R.A., 2013. HydroServer Lite as an open source solution for archiving and sharing environmental data for independent university labs. *Ecological Informatics* 18(0) 171-177, <http://dx.doi.org/10.1016/j.ecoinf.2013.08.006>.
- Laniak, G.F., Olchin, G., Goodall, J., Voinov A., Hill, M., Glynn, P., Whelan, G., Geller, G., Quinn, N., Blind, M., Peckham, S., Reaney, S., Gaber, N., Kennedy, R., Hughes, A. 2013. Integrated environmental modeling: A vision and roadmap for the future. *Environmental Modelling & Software* 39: 3-23.
- Merwade, V., Song, C., Zhao, L., Zhe, S., Rajib, M.A., 2013. SWATShare – A portal for sharing, publishing, and running SWAT models using XSEDE resources, Proceedings of the 2013 SWAT Users Conference and Workshops: Toulouse, France, July 2013, <http://water-hub.org/swatshare>.
- Neitsch, S., Arnold, J., Kiniry, J., Williams, J., 2011. Soil and Water Assessment Tool theoretical documentation version 2009. Texas A&M University System, College Station, TX, USA.
- Tarboton, D. G., Idaszak R., Horsburgh J. S., Heard J., Ames D., Goodall J. L., Band L., Merwade V., Couch A., Arrigo J., Hooper R., Valentine D., Maidment D. 2014, HydroShare: Advancing Collaboration through Hydrologic Data and Model Sharing. in D. P. Ames, N. W. T. Quinn and A. E. Rizzoli (eds), Proceedings of the 7th International Congress on Environmental Modelling and Software, San Diego, California, USA, International Environmental Modelling and Software Society (iEMSs), ISBN: 978-88-9035-744-2, http://www.iemss.org/sites/iemss2014/papers/iemss2014_submission_243.pdf.

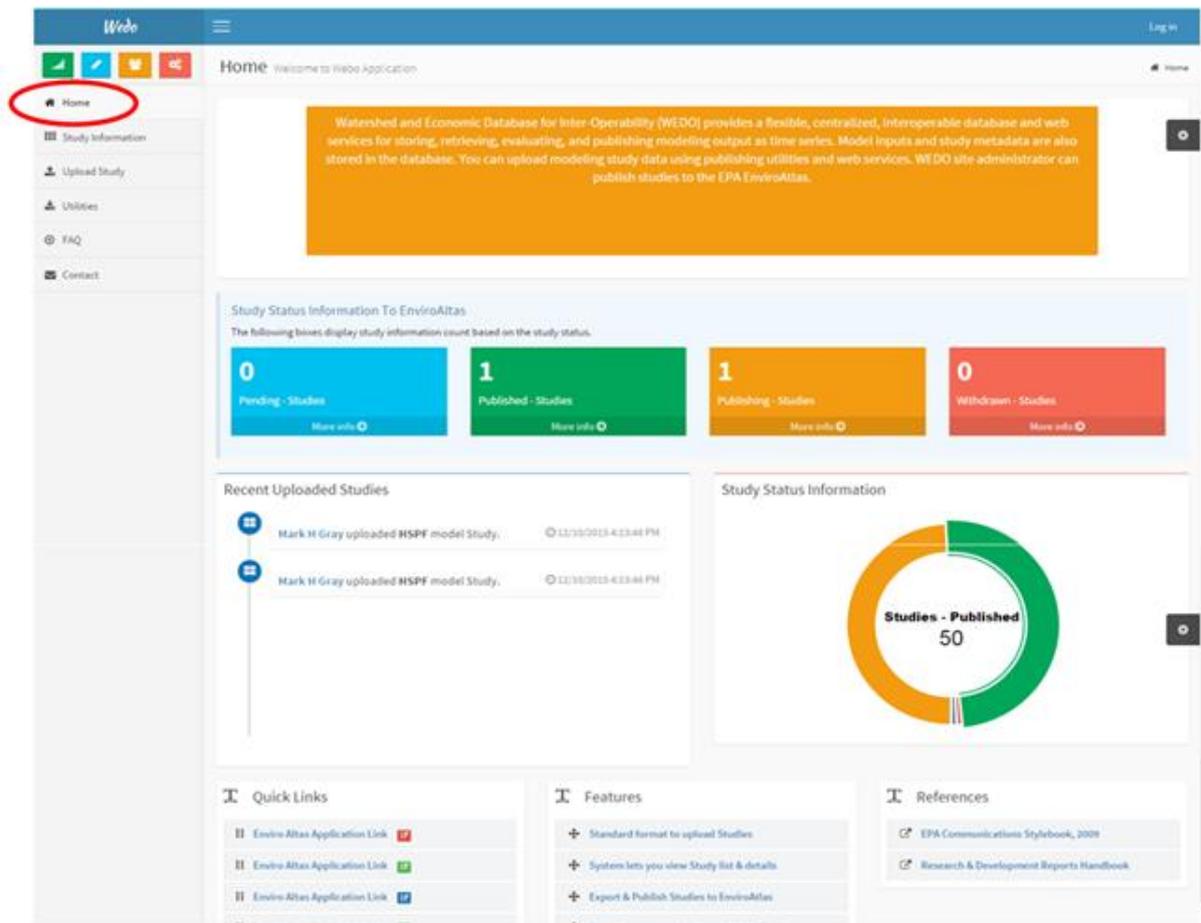


Figure 2. WEDO Home Page

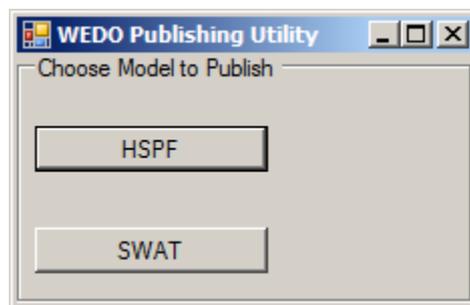


Figure 3. Choosing a model within WEDO Publishing Utility

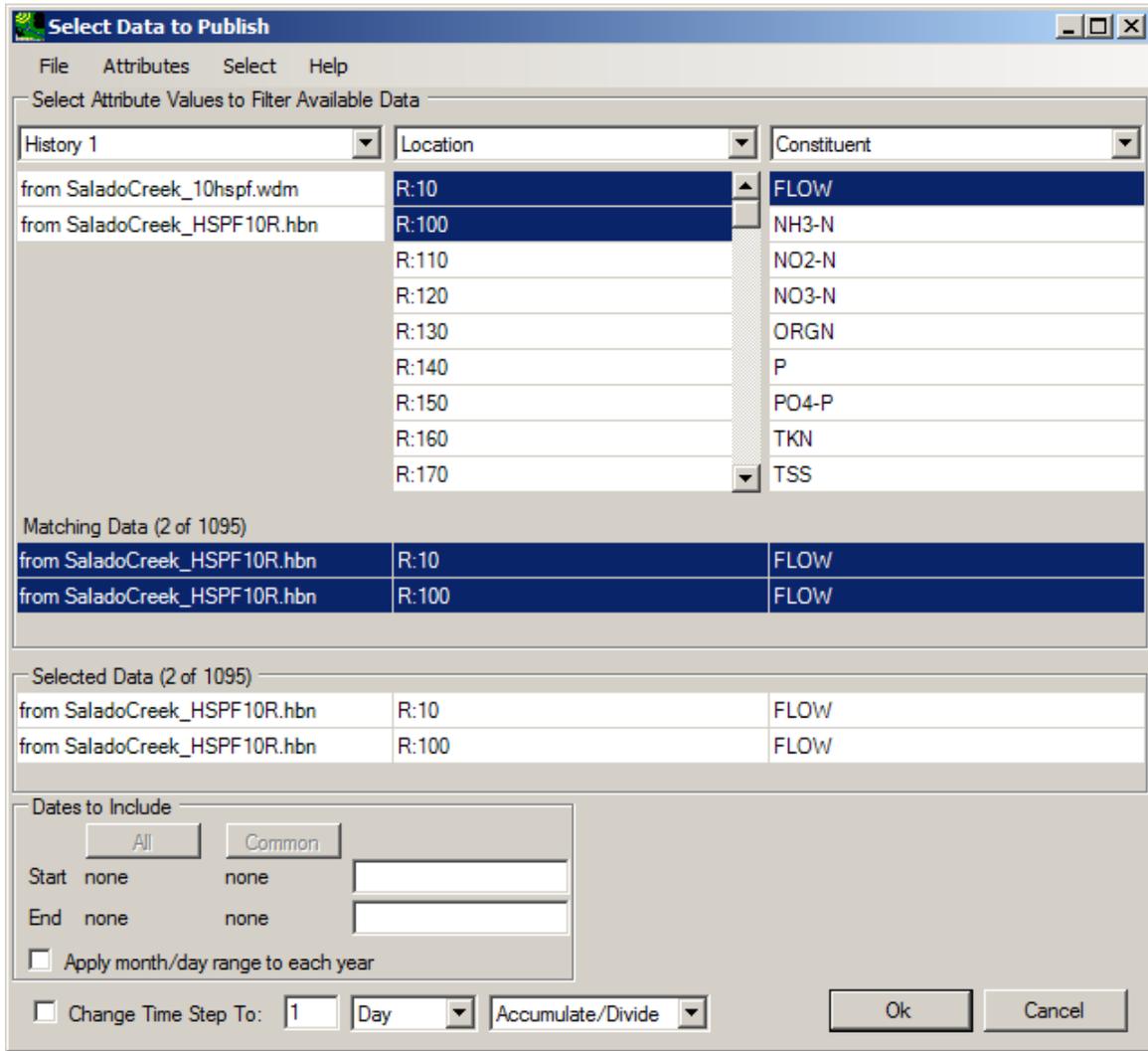


Figure 4. Selecting data to publish within WEDO Publishing Utility

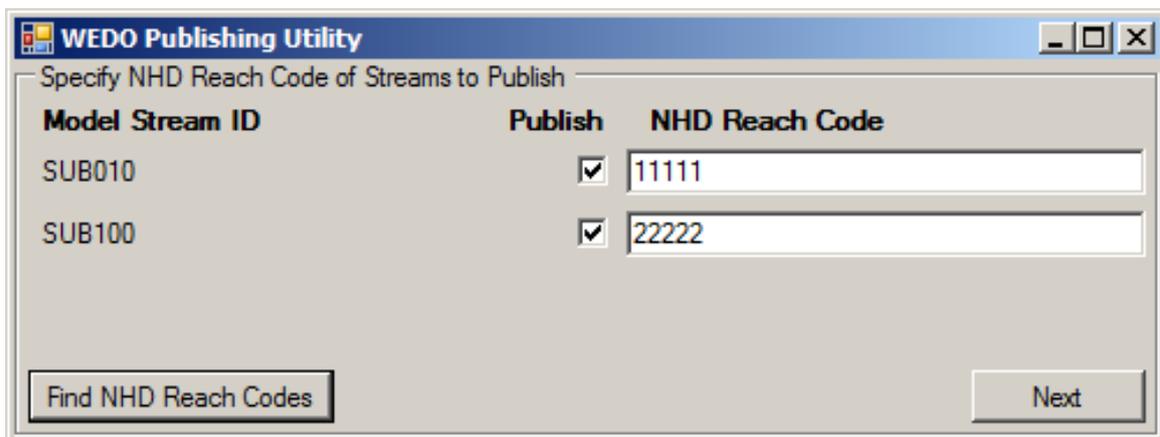


Figure 5. Specifying NHD Reach Codes within WEDO Publishing Utility

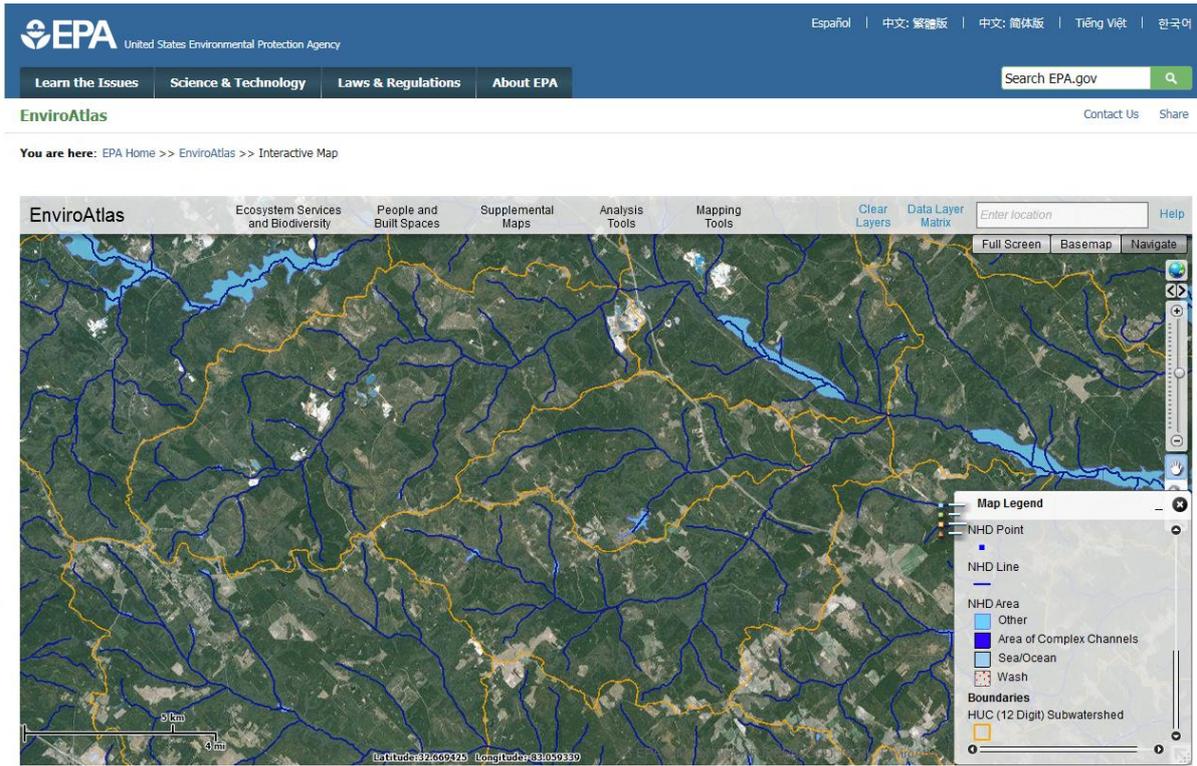


Figure 6. EPA EnviroAtlas showing stream network in blue and yellow