An investigation of the WaterML2.0 data standard for data sharing for the purpose of hydrological forecasting

S. Hamilton & Touraj Farahmand

*Aquatic Informatics, Vancouver, Canada*

N. Booth, Laura De Cicco & David Sibley

United States Geological Survey, Middleton, United States

J. Halquist

*National Weather Service, National Operation Hydrologic Remote Sensing Center, Chanhassen, United States*

L. Liu

*Environment Canada, Ottawa, Canada*

 J. Woodward

*Saskatchewan Wateshedr Authority, Moose Jaw, Canada*

B. Harrison

*Manitoba Water Stewardship, Winnipeg, Canada*

 R. Boals

*International Souris River Board, Regina, Canada*

A. Joseph

*Cybera – Water and Environmental Hub, Calgary, Canada*

ABSTRACT: The Open Geospatial Consortium (OGC) and the World Meteorological Organization (WMO) are collaborating in the development of a new data standard for the sharing of water information. The new data standard is referenced as WaterML2.0. The ability of the WaterML2.0 standard to meet the needs of end-users and to satisfy the concerns of data providers is evaluated and tested in a series of interoperability experiments. One such experiment is addressing the application of WaterML2.0 in hydrological forecasting. Hydrological forecasting requires timely access to a wide variety of data from multiple sources in a timely manner. The typical forecast desk dedicates a considerable amount of time on manual data search, retrieval, parsing, and ingest. These tasks are at the expense of thoughtful analysis and timely warnings. Hydrological forecasts must, in turn, be readily discoverable, searchable, and accessible. Data providers have concerns about the security of the data; user-level acknowledgement of risk; a traceable waiver of liability; and the load that web services will put on their servers. The study area selected for the hydrological forecasting interoperability experiment is the Souris River basin, which originates in Saskatchewan, crosses the international border into North Dakota, and returns north to Manitoba. Participating agencies include: the National Oceanic and Atmospheric Administration (NOAA); The U.S. Geological Survey (USGS); The Water Survey of Canada (WSC); Saskatchewan Watershed Authority (SWA); Manitoba Water Stewardship (MWS); International Souris River Board (ISRB); Cybera - Water and Environmental Hub; and Aquatic Informatics (AI). The objective of the study is to test the operational readiness of the WaterML2.0 standard; to identify gaps in the standard; and to prepare recommendations for improvements.

# InTroduction

Flood forecasting is data intensive operation. Statistical models are required for seasonal outlooks, dynamic hydrological models are needed to represent landscape-scale runoff processes and hydraulic models are required to represent in-channel and over-bank flow routing. The skill of these models can never be better than the data used to calibrate the models and to update initialization fields. Data are needed evaluate model performance in real-time. Early detection of false predictions can mean the difference between life and death. Timely validation of predictions raises confidence in forecasts improving compliance with emergency response actions. In some cases, warnings are needed which can only be based on sparse evidence with no modeling support.

Forecasters are typically responsible for large, hydrologically complex, regions. Within any such region, hydrometric data are collected for a wide variety of purposes by diverse agencies of which the forecaster will have direct access to only the primary data providers by pre-arranged agreements. Based on this model for data sharing the inputs to the river forecast system will typically be hard-coded to specified data communications protocols and data parsing algorithms.

This is an efficient system to provide timely data for specified locations and parameters that the forecaster can use with a confidence based on previous experience with these locations and the data provider. However, this also ensures that the forecaster will have access to only a sub-set of the data that could be relevant to any given forecasting problem.

Information from any private, public and academic data producer could be searchable by descriptive, meaningful, metadata queries. The results of these queries could be filtered by, and/or marked with, interpretive encoding linking to methods used, traceability to source, uncertainty, and quality management standards of the data provider. All the information need to acquire the data resulting from these queries could be provided as part of the search result. The data could be correctly integrated into any data management schema by the use of clearly defined, universal, semantic and syntactic translation protocols.

Agreement on a comprehensive framework of data interoperability standards is needed to achieve such an open and flexible data sharing environment. Forecasters should be able to react to unusual events with seamless search; discovery; and access to any data in the watershed that is relevant to the forecasting problem In the face of any quickly evolving crisis; forecaster time is best spent on analysis and dissemination of warnings.

Forecasting floods across jurisdictional boundaries can be even more complicated. There should be no loss of information about an event as it passes across a boundary. This means that forecasts as well as data should be searchable, discoverable and accessible by the same framework of inter-operability standards.

Achieving consensus on a framework of data standards is no easy task. However, the Open Geospatial Consortium (OGC) has a history of demonstrated success at developing the type of standards that allow anyone to quickly pull up a map on their cell phone of anywhere in the world regardless of whom the data provider is.

Under the auspice of an agreement between the World Meteorological Organization (WMO) and OGC to harmonize international standards for water data, the Hydrological Domain Working Group (HDWG) was established. The HDWG took on the task of integrating WaterML1.0 into existing OGC standards. The result, WaterML2.0, has been tested in two previous inter-operability experiments. The special needs of the forecasting community are considered in this, the third inter-operability experiment. Each experiment provides important feedback to the HDWG, identifying deficiencies and guiding further development.

There are many challenges in developing standards. There are performance expectations of data consumers that must be addressed. One measure of performance may be comparison against existing hard-coded, or proprietary, communications protocols. A simple race of bits of data from server to server would not be a fair contest. A better test would be more like a decathlon, where the relative performance with respect to search, discovery, filtering, and integration of data into end-user applications are all considered.

Having end-users ready to ingest data using the new standard is the lesser part of the problem. The main problem is to develop the standards so that they meet the needs of the data providers. Most data providers already have some method of making their data available to their primary end-users of the data.

There is little motivation for a data provider to make their data more widely available using a method that they do not explicitly control. At a minimum, the data standards should not expose the data providers to risks. The security of the source databases is of paramount importance. There are also risks of liability. There are legitimate concerns about the release of data without acceptance of a disclaimer and/or copyright terms and conditions.

The Souris River Project is conceived as part of an experiment to identify opportunities for improvement in the WaterML2.0 standard with respect to sharing inter-jurisdictional data for the purpose of hydrological forecasting. Central to the experiment is to identify potential barriers to data sharing that can be alleviated by improvements in the standards.

# Method

**Stu, Nate and John** to provide details of OGC IE method (basically executive summaries of: 105\_Hydrological\_forecasting\_Interoperability\_Experiment\_1.doc; and OGC® Engineering Report: Water Information Services Concept Development Study) and the specifics of the Souris River Project

# discussion

**Lingling, Jeff, Bob, Nate, Laura and Dave** to provide a data provider perspective.

**John, Jeff, Bob, Russell, Alex** to provide a data consumer perspective.

# conclusion aNd NEXT STEPS

**Nate, John** to provide lead-in to pilot project

# References