

Going Public with Advanced Simulations

Niclas Hjerdt, Berit Arheimer, Göran Lindström, Ylva Westman, Esa Falkenroth, Martin Hultman

► **To cite this version:**

Niclas Hjerdt, Berit Arheimer, Göran Lindström, Ylva Westman, Esa Falkenroth, et al.. Going Public with Advanced Simulations. 9th International Symposium on Environmental Software Systems (ISESS), Jun 2011, Brno, Czech Republic. pp.574-580, 10.1007/978-3-642-22285-6_62. hal-01569242

HAL Id: hal-01569242

<https://hal.inria.fr/hal-01569242>

Submitted on 26 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Going public with advanced simulations

Niclas Hjerdt, Berit Arheimer, Göran Lindström, Ylva Westman, Esa Falkenroth, and Martin Hultman

Swedish Meteorological and Hydrological Institute, Folkborgsvägen 1, SE-60176 Norrköping, Sweden

Correspondence to: niclas.hjerdt@smhi.se

Abstract. Web technology provides public access to advanced simulation models and efficiently distributes vast volumes of data in an unprecedented way. The traditional boundary between the producers and the users of data has become less distinct. Users now download data generated by themselves or by others. The role of the hydrologist has therefore shifted towards the management of systems and tools that are publicly available. A recent example of this development is the SMHI VattenWeb, a web application that makes available different databases and modelling tools managed by the Swedish Meteorological and Hydrological Institute. Since VattenWeb was primarily developed to meet the demands of the EU Water Framework Directive, one of the key elements is a database of simulated discharge and nutrient transport generated by the SMHI HYPE model. These data consist of daily, monthly and annual hydrologic simulations for all Swedish water bodies reported within the WFD. We show the present status of SMHI VattenWeb and highlight the strategies that have made this application successful among users seeking to characterize the quantity and quality of water at high spatial resolution across Sweden.

Keywords. Large scale, High-resolution, WFD, public access, web interface

1. The new challenges of water management

Hydrologic modelling has traditionally been an exclusive activity for trained hydrologists, but this has begun to change rapidly as a result of recent development in technology and legislation. The implementation of the European Union Water Framework Directive (WFD) has created a need for hydrologic information at high resolution in both time and space. In addition, web technology now provides tools to efficiently share vast volumes of data. Other recent legislation at the EU level, such as the INSPIRE Directive, has created a framework for sharing geographic information with the public in a standardized manner.

At 450,295 square kilometres, Sweden is the third largest country in the European Union by area, and Sweden's 23,418 reported water bodies make up a relatively large proportion of all European water bodies. Since Sweden is a country

2 **Niclas Hjerdt, Berit Arheimer, Göran Lindström, Ylva Westman, Esa Falkenroth, and Martin Hultman**

with vast water resources and a relatively small population, it is impossible to maintain a monitoring program that covers all water bodies. Therefore, models offer a cost-efficient option for many steps in the six-year water management cycle. Therefore, the implementation of the WFD and subsequent delineation of water bodies has created a need for models that can operate at a much finer resolution than previously needed. This could not be achieved with, for instance, the hydrological model used in the international reporting of nutrient loads to the Baltic, i.e., the HELCOM Pollution Load Compilation. The model used for this is the HBV-NP model which only divides Sweden in around 1,100 subbasins, as compared to the 23,418 water bodies reported to the EU (Figure 1)..



Fig. 1. The resolution of models used in the HELCOM pollution load compilation reports (left) and for the EU Water Framework Directive reports (right).

2. The solution, part one: A new, large-scale hydrologic model with high resolution

The challenges of the WFD clearly created a need for a new type of hydrologic models. One such model is the hydrological catchment model HYPE (HYdrological Predictions for the Environment), which simulates water flow and substances on their way from precipitation through soil, river and lakes to the river outlet [1, 2]. The catchment is divided into subbasins which in turn are divided into classes (calculation units) depending on land use, soil type and elevation.

An important feature of HYPE is that classes can not be coupled to a geographic location within the subbasin but are given as part of its area. Typical land uses are forest, lake, open land, but also different crops, e.g. cereal and potatoes, are common. Elevation can be used to get temperature variations within a subbasin to influence the snow conditions.

In contrast to the earlier type of models, such as the HBV-NP model, HYPE simulates the water flow paths and nutrient transformations in soil. The soil is modelled as several layers which may have different thickness for each class. Agricultural land classes commonly use three soil layers to simulate nutrient transformation more accurate. Model parameters may be associated with land use (e.g. evapotranspiration), soil type (e.g. water content in soil), or be common for the whole catchment. This way to couple the parameters to geographic information makes the model better suited for simulations in ungauged catchments.

HYPE simulates the nutrients nitrogen and phosphorus divided into the following fractions: inorganic nitrogen, organic nitrogen, soluble reactive phosphorus and particulate phosphorus. In addition, organic material and conservative substances like ^{18}O can be simulated. The calculations are made with a daily time step for water and concentrations, but result can also be delivered as mean values over a longer period and in the form of transported amount of nutrients and source apportioned transport. The HYPE model is set up for a growing number of basins, including Sweden [2], the Baltic Sea drainage basin, Europe [2,3], and La Plata in South America.

3. The solution, part two: Storing data in an efficient and flexible database

Running the HYPE model at high resolution for a country or a continent quickly generates vast volumes of data. For Sweden only, the present S-HYPE version with 17,313 subbasins and calculations of discharge, nitrogen- and phosphorous transport with a daily time step for the period 1990-2010 generates approximately 400 million values. For Europe, the present E-HYPE version with 47,000 subbasins generates approximately half a billion values of daily discharge values.

The best solution in our opinion was to follow standard procedure. As a first step, a database model was developed to characterise the dependencies and relationships between different attributes to be stored in the database. After the structure of the database had been determined, a Postgres 8.4 database was generated. A database loading program in Java was developed to read model output files into the SQL database.

It was crucial already at this point to assure that the database would be highly accessible through a web-based interface. Achieving short response times was a priority. Different accessibility solutions were developed for the Swedish database and the European database. For the Swedish database, data were cached in main memory (currently 68 GB main memory) for rapid delivery. For the European database, we chose to partition the database per subbasin, which resulted in a good throughput as long as you retrieve data from a limited number of subbasins. The partitioning require much less main memory on the database server but the disadvantage is sacrificed flexibility of retrieval, i.e., retrieve all subbasins for a single point in time is not practical.

4. The solution, part three: A web-based user interface

Developing a web-based user interface was a strong requirement to assure public accessibility. We have developed two services that satisfy a range of data accessibility needs.

The first service is a simple but functional map interface which allows users to first identify and select subbasins of interest. After subbasins have been selected the user fills out a short form to identify the data period, temporal resolution and variables to download. Finally, after the “download” button has been pressed, data series will be delivered in a spreadsheet that can be opened or saved to a local hard drive.

The second service is a web service that delivers data in xml-format following a REST call. This solution is suitable for machine-to-machine communication and when downloading data for a large number of subbasins. Due to network security reasons, the web service has so far been restricted to registered users, but there are plans to turn this into an open API in the near future. Technically, we have identified a series of components that together meet the requirements of effective, web-based, data-sharing (Table 1):

Table 1. List of reliable components that have been used to construct the technical basis of SMHI VattenWeb.

Purpose	Component
Web-application framework	Google Web Toolkit
Client side map viewer	Openlayers
Database solution	Postgres 8.4
Coupling to database	JDBC
Log	log4j
Map server	Geoserver and GeoWebCache with cached images for different zoom levels
Web server	Apache
Run web applications in Java	Tomcat
Web service call	Jersey
Spread sheet generation	Apache POI

The web-based user interface is constructed with the framework Google Web Toolkit (GWT) together with Openlayers as a map viewer. With the help of GWT both client and server side can be written in java-code. On the server side, JDBC is used to communicate with a Postgres database and Apache POI generates spread sheets from returned data. The web services REST interface is built with the help of Jersey. Both the web-application and web services is hosted within a Tomcat web server that runs under Apache. Geoserver and GeoWebCache is used to host to the Openlayers client with geographical metadata and pre cached maps through standard protocols WFS and WMS.

5. The solution, part four: An operational production line for WFD support services

In most cases, piecing together models, databases and web pages is carried out in project form with financial resources restricted to a limited time period. However, to meet the challenges of European water management, information must be updated and accessible more or less continuously. One solution is to incorporate supporting systems into an operational production line, but this requires a solid technical infrastructure.

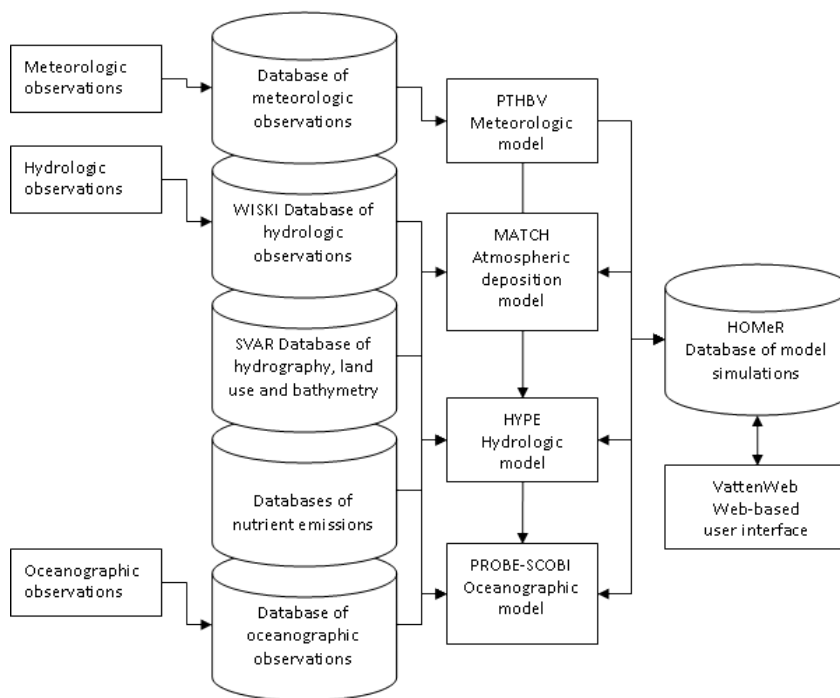


Fig. 2. A simplified illustration of the operational production line for WFD supporting services at the Swedish Meteorological and Hydrological Institute. Parts of this production line is also utilized for our pan-European services, such as E-HypeWeb [4].

This technical infrastructure is already in place at SMHI. As a government agency with national responsibility for meteorology, hydrology and oceanography, SMHI already operates a number of services on a daily basis. There is a strong IT department with high computational capacity and around-the-clock technical support. Data security issues are dealt with in a professional manner.

The production line for WFD supporting services has now been established at SMHI but is still in its infancy and subject to improvements (Figure 2). Due to Sweden's late implementation of the WFD (2004), and even later involvement of

WFD at SMHI (2008), many WFD supporting services were initially developed as individual projects with minimum overlap. As a result, these systems became very expensive to keep updated and operational. Therefore, many first-generation systems are now being merged to maximize cost-effectiveness.

The establishment of a production line is critical also for handling system and information updates efficiently. For data on river discharge and nutrient transport, both products of multiple data sets and calculation steps, it is critical to keep track of the exact combination of model versions and data sets used in the calculation. Nevertheless, new observations, or changes in the delineation of water bodies, require databases and models to be updated frequently. Thus, data sets must be carefully associated with metadata to ensure that data can be traced and reproduced at a later stage. In our opinion, it is very difficult to sustain the handling of data versions and reproducibility without a firm production line in place.

6. Conclusions

The first web application that gave public access to large-scale hydrologic simulations with high resolution was SMHI VattenWeb [5], a web interface that offers time series of river discharge and nutrient transport from 17,313 subbasins in Sweden. Since the launch in April 2010, the number of unique visitors has steadily increased and approach 3000 at the end of March 2011 (Figure 3).

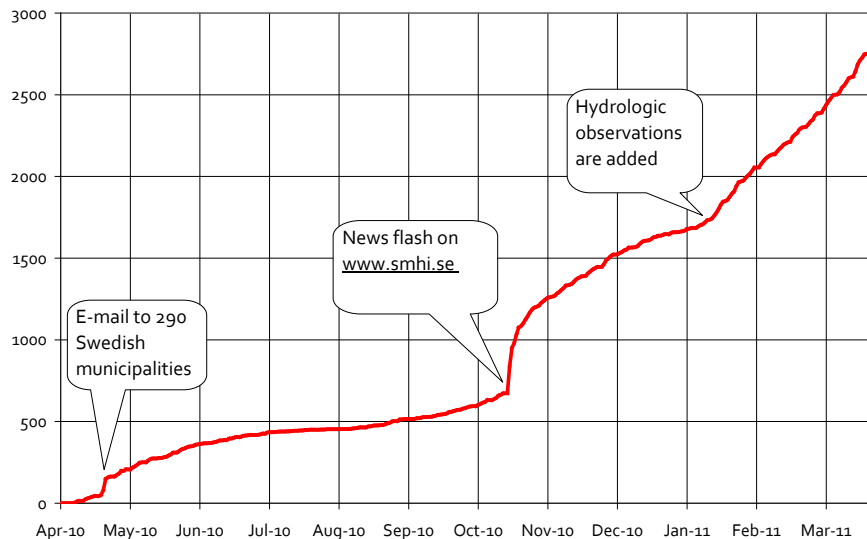


Fig. 3. Number of unique visitors at SMHI VattenWeb [5] from the start in April 2010 to March 31, 2011, according to statistics from Google Analytics.

Later during 2010 a pan-European HYPE model, E-HYPE, was set up and launched as a demonstration tool [4]. The E-HYPE model has a resolution of approximately 47,000 subbasins across the European continent. Other publicly available web services include Balt-HYPE, a HYPE model for the Baltic drainage basin [6].

We believe that strong hydrologic research and development must be placed in an operational production line to satisfy the needs of the WFD. We have demonstrated how a production line has been set up at the SMHI and described the different components. It is clear that public interest for these services is very high and we see many benefits of such systems at the European scale, not only for fresh water management related to the WFD, but also in other sectors. In the near future we will probably see the emergence of a large variety of downstream services that utilize data from publicly available core services.

References

1. Arheimer, B., Lindström, G., Pers, C., Rosberg, J. and Strömqvist, J: Development and test of a new Swedish water quality model for small-scale and large-scale applications. XXV Nordic Hydrological Conference, Reykjavik, August 11-13, 2008. NHP Report No. 50, pp. 483-492. (2008)
2. Donnelly, C., Dahne, J., Lindström, G., Rosberg, J., Strömqvist, J., Pers, C., Yang, W. and Arheimer, B: An evaluation of multi-basin hydrological modelling for predictions in ungauged basins, Proc. of Symposium HS.2 at the joint IAHS & IAH Convention, Hyderabad, India, September 2009, IAHS Publ. 333. (2009)
3. Strömqvist, J., Dahne, J., Donnelly, C., Lindström, G., Rosberg, J., Pers, C., Yang, W. and Arheimer, B: Using recently developed global data sets for hydrological predictions, Proc. of Symposium HS.2 at the joint IAHS & IAH Convention, Hyderabad, India, September 2009, IAHS Publ. 333. (2009)
4. Discharge in European rivers according to the E-HYPE model, <http://e-hype.smhi.se>
5. Discharge and nutrient transport in Swedish rivers according to the S-HYPE model, <http://vattenweb.smhi.se>
6. Discharge and nutrient transport for the Baltic drainage basin according to the Balt-HYPE model, <http://balt-hype.smhi.se>