

# Hydroinformatics Support to Flood Forecasting and Flood Management

**ADRI VERWEY**

WL | Delft Hydraulics, Delft, The Netherlands  
*Adri.Verwey@wldelft.nl*

**Abstract** This keynote lecture describes state-of-the-art hydroinformatics support to the water sector. A few examples are worked out in some detail, whereas for other examples the reader is guided to recent literature. The focus is on flood forecasting and flood management, with a brief description of the potential of changing technologies to support studies and facilities in this area. Examples are: new data collection methods; data mining from these extensive new sources of information, *e.g.* the use of genetic programming; data driven modelling techniques, *e.g.* artificial neural networks; decision support systems; and the provision of a hydroinformatics platform for flood forecasting. Particular attention is given to advances in numerical flood modelling. Over the past years the robustness of numerical models has increased substantially, solving for example, the flooding and drying problem of flood plains and the computation of supercritical flows. In addition, the emergence of hybrid 1D2D models is discussed with their different options of linking model components of flood prone areas.

**Key words** Hydroinformatics; flood management, flood forecasting, unsteady flow modelling; flood simulation modelling; flow resistance; data mining; Open MI; open source; proprietary software.

## INTRODUCTION

Hydroinformatics covers the application of information technology applied to the water sector in the widest sense. The continuously increasing speed of computers and increased density of information storage, the increased communication potential through internet and the creative power of scientists have brought us rapidly forward in the way in which water related studies can be executed, currently based upon a much better understanding of underlying processes and descriptive means than some decades ago. For a better awareness of what is being opened up with these developments it is recommended to participate in the two-yearly Hydroinformatics Conferences (*e.g.* Nice, 2006; [www.hic06.org](http://www.hic06.org)). This paper also gives a number of references to contributions published in the proceedings of the Hydroinformatics Conference Singapore, 2004. Another recent source of information is a number of



contributions related to hydroinformatics, published in the Encyclopaedia of Hydrological Sciences, *e.g.* Werner *et al.* (2005), Minns and Hall (2005), Stelling and Verwey (2005) etc. In this introduction we will give an outline of various interesting developments and treat in subsequent paragraphs, on a selective basis and in more detail, some more specific issues in relation to flood forecasting and flood management.

## **HYDROINFORMATICS TOOLS**

The focus on hydroinformatics emerged from the field of computational hydraulics, when it was understood and felt desirable that around the modelling systems developed in the last three decades of the last century, a complete infrastructure of informatics support existed and that its potential had to be explored and expanded to improve service to society. It comprises data acquisition and data management techniques; new simulation techniques based upon cognitive sciences and pattern recognition, such as artificial neural networks; data mining and knowledge discovery techniques; evolutionary algorithms; decision support and management systems; forecasting and data assimilation methods; fuzzy logic; cellular automata; integration of systems and technologies; and emerging internet based technologies. A state-of-the-art description of new technologies applied in the area of ecohydraulics was given by Mynett (2004). Historically, balance equations or empirical relationships were developed by scientists, *e.g.* Newton, Navier, Poisson, de Saint Venant, Stokes, Darcy, de Chézy, Strickler, Manning etc., by trying to define fundamental relationships between various system state variables on the basis of observations and by setting up balance principles, partly based upon the empirical relationships found. A thorough presentation of the role of various scientists is included in the work of Chanson (1999). Hydroinformatics can now be seen as providing an extension to these developments, partly by using the computer power to guide and process new and massive data collection techniques and partly by leaving it to computational power to



establish best fitting new relationships, often in areas which could not be explored before. The change that hydroinformatics really brings is a change in the role of scientists, from those who establish laws to those who guide the establishment of laws and relationships by mobilizing computer power.

A good example is the use of artificial neural networks (ANN), which can be seen as an extension of the traditional use of regression techniques, *e.g.* Minns and Hall (2005). However, whereas in regression techniques formulae have to be prescribed and parameters calibrated, ANN provides the additional flexibility that the relationships between state variables, or the formulae, are left open, in fact never defined, as all relationships established are based upon signals passed on through a sequence of (neural) cells with weightings established by the so-called learning process through numerous trials. Compared with the human mind, only one thing is missing: the ability to extrapolate the knowledge outside the range where the learning process took place and even many human beings have difficulties with such extrapolations. Although ANN's have opened up the way to new simulation techniques, the learning process shows clearly their limitation, for example in rainfall-runoff modelling, where physically based balance models, equipped with appropriate limiters, can be used more trustfully in extreme situations that go beyond earlier observations.

A step beyond ANN is the employment of sets of data to establish empirical relationships in the form of mathematical expressions by using evolutionary algorithms. Unlike ANN's, where the development process of knowledge in the human brains serves as an example for computer based knowledge development, evolutionary algorithms take the biological reproduction process as the blueprint for the derivation of mathematical relationships. By seeing state variables, operators and parameters as components of DNA strings, recombinations of sub strings, together with the process of mutation leads, in a learning process, to continuously better matching of mathematical relationships in a "survival of the



“fittest” process. A recent example that serves flood modelling is given by Baptist *et al.* (2005). For a number of years, Baptist (2005) has been working on the development of empirical relationships defining the resistance of flow, expressed as a de Chézy value as a function of flow depth and height and type of vegetation. For submerged vegetation, Baptist derived the relationship

$$C_r = \sqrt{\frac{1}{C_b^{-2} + \frac{C_D m D k}{2g}}} + \frac{(h-k)^{3/2} \sqrt{g} \ln\left(\frac{h-k}{e z_0}\right)}{h^{3/2}} \quad (1)$$

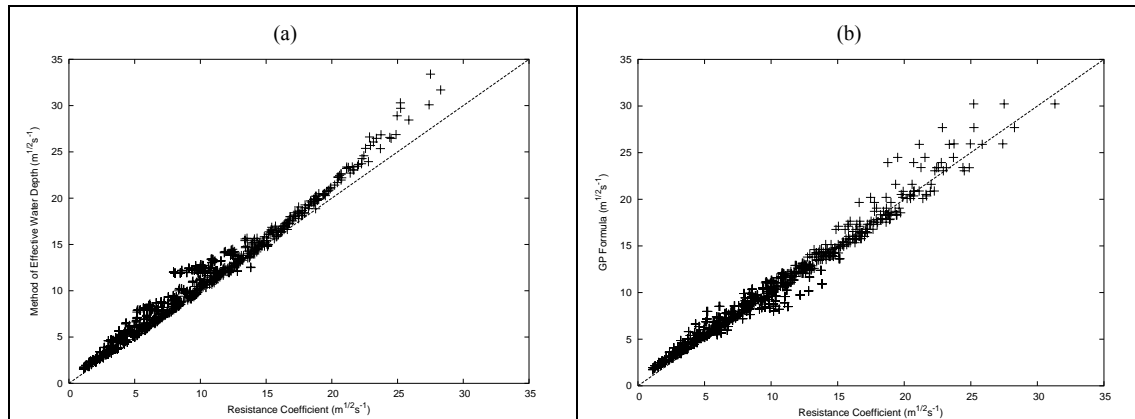
where  $C_r$  = depth dependent Chézy coefficient [ $\text{m}^{1/2} \text{s}^{-1}$ ],  $C_b$  = Chézy value for the bottom friction alone [ $\text{m}^{1/2} \text{s}^{-1}$ ];  $C_D$  = drag coefficient for flow around the vegetation stems [-];  $m$  = vegetation density [ $\text{m}^{-2}$ ];  $D$  = representative stem diameter [m];  $g$  = acceleration due to gravity [ $\text{m s}^{-2}$ ];  $k$  = representative vegetation height [m];  $\kappa$  = von Kármán’s constant [-];  $e$  = base of the natural logarithm [-] and  $z_0$  = roughness height of the top of the vegetation [m]. The equation was checked on a set of 990 results obtained with a 1-DV (1-dimensional in the vertical) model based upon the Delft3D code, including a description of turbulence developed around the stems of vegetation.

Rodriguez Uthurburu (2004, and to be published further in Baptist *et al.*, 2005) used the same set of data to train a genetic programming code and came up with the equation

$$\frac{C_r}{\sqrt{g}} = \sqrt{\frac{2}{c_D m D k}} + \ln\left\{\left(\frac{h}{k}\right)^2\right\} \quad (2)$$

Figure (1) shows the scatter plots for both equations. Also based upon RMSE values for both sets: 1.30 for Equation (1) and 0.97 for Equation (2), respectively, it can be concluded that the data-driven discovery process has led to a better fitting equation than the equation derived on the basis of existing theoretical knowledge.





**Fig. 1** Scatter plots for the two equations compared with 1-DV data: (a) Method of effective water depth, Equation (1); (b) Original Genetic Programming formula, Equation (2).

In a next step, and further analyzing the descriptive nature of Equations (2) and (1), it was decided to impose the first term of the right hand side of Equation (1) and add in the data-mining process the von Kármán's constant  $\kappa$  to the set of parameters. Genetic programming then came up with the equation

$$C_r = \sqrt{\frac{1}{\frac{1}{C_b^2} + \frac{1}{2g} C_D m D k}} + \frac{\sqrt{g}}{\kappa} \ln\left(\frac{h}{k}\right) \quad (3)$$

which had an RMSE value of 1.21, between those for Equations (1) and (2), but believed to have a wider range of applicability than Equation (2). Results were also tested against 177 experimental runs based on laboratory flume experiments from 10 independent sources. In this case the RMSE values are higher, as can be expected. However, still good enough to confirm the usefulness of the data-driven equation development.

Above experiments show that computer-based technologies provide us with better means of exploring relationships in nature, either based upon large sets of measured data or upon data generated with very fine grid numerical simulations based upon existing theories. The experiment also shows that the application of these technologies merely provides an extension



of our scientific minds and certainly not a replacement. With the current floods of data collected with remote sensing, for example, there must be many ways in which data-mining will assist us in deriving relationships in such complex areas as environmental process description, ecology and unsaturated ground-water zone analysis.

## **OPEN SOURCE VERSUS PROPRIETARY SOFTWARE**

Dealing with hydroinformatics, a continuous discussion point is the question of open source against proprietary software, which is generally provided in the form of compiled executables. This form also implies that the software codes cannot be modified by staff other than that of the software vendors. Discussions on open source come down to the question of how water resources agencies are best served and how providers of hydroinformatics services are stimulated to provide the best tools. The point of view of the agency is extensively described by Khatibi *et al.* (2004) with inputs from hydroinformatics tool providers. A number of aspects will be presented briefly here. For water resources agencies, interested only in the use of software, the following aspects are of interest:

- a) economic overall solutions for their hydroinformatics infrastructure, *e.g.* the overall costs for the development of a flood forecasting system, including the forecasting platform, costs for underlying software, costs for model development and calibration etc.;
- b) reliability of the software products, through good development practices and maintenance of the codes, based upon extensive testing procedures;
- c) openness of the software, enabling the coupling of various components;
- d) reproducibility of results through good version management;
- e) state-of-the-art scientific basis for the methods implemented in the software;
- f) quick response to development needs.

Most of these requirements also apply to vendors of hydroinformatics software, as it is in their interest to provide good service to clients and this in a competitive environment. Open source



is definitely not in the interest of this group, unless special factors play, such as market penetration for other services (*e.g.* consultancy), name branding and free co-developments by third parties. This last interest sometimes leads to adopting the concept of “open source in closed community”, based upon co-development of codes by a limited group of participating organisations and the right for one or more of these partners to distribute the code commercially.

It is not the place here to draw definite conclusions on what is best practice. With a common sense mind for advantages and disadvantages, the best solution can be found for each case. However, regarding point (3) the difference between open and closed source will be reduced in the near future. Water management agencies, in general, have a strong interest in openness of software, so that various components can be linked without the need to contract the original developers of such software components. This service is now being developed through the European OpenMI concept, which facilitates coupling both of open source and proprietary software components.

## **OPEN MI**

Dealing with water authorities, different consultants and own staff are implementing continuously bits and pieces of software and often there is neither a contractual obligation nor the means to embed these new service tools into a consistent hydroinformatics infrastructure for the organisation. For this reason, the concept of OpenMI (Open Modelling Interface) was developed through the European HarmonIT project, initiated by the Ministry of Transport, Public Works and Water Management in The Netherlands as an important stakeholder. OpenMI has been designed as a generic set of communication rules for linking all kinds of software components available to the water sector. For example, OpenMI facilitates the transfer of data in a generic way and at any moment in time, from one model, *e.g.* a rainfall-runoff model, to another model, such as a hydrodynamic model, vice versa. More in general,



OpenMI facilitates the flexible linking of a wide variety of simulation models, generic modelling systems, databases, GIS, decision support systems, web based services etc. (Gijsbers, 2004).

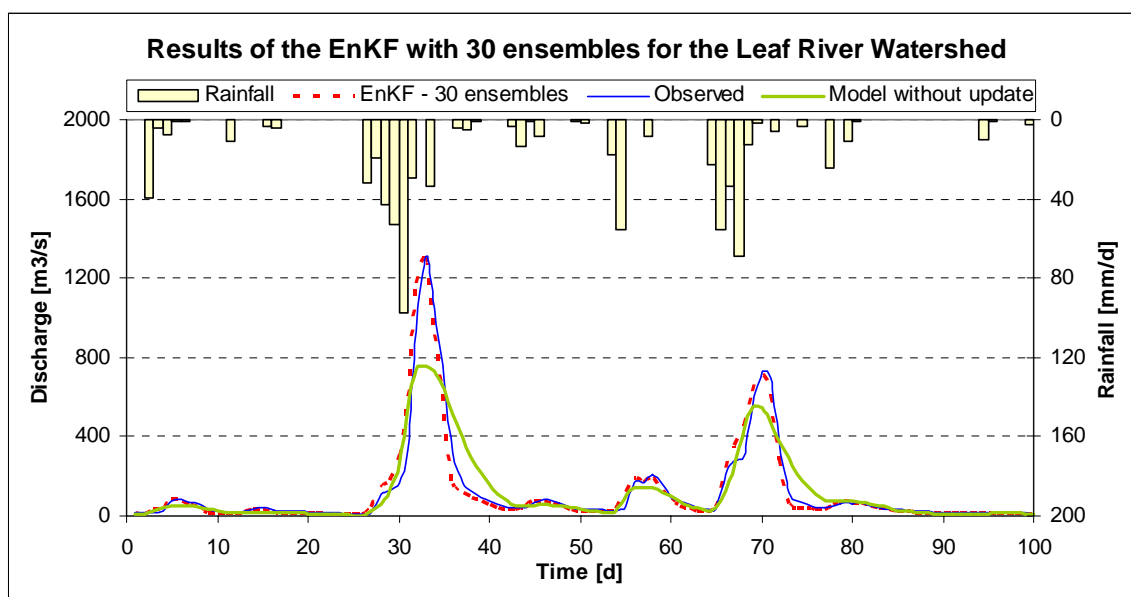
OpenMI is primarily a set of rules on how to exchange data in a pre-defined way. Ideally, within an organisation, all software tools should be linkable through a common platform. With the current trend of developing integrated models of hydraulic, hydrologic and environmental systems, the development of such platforms becomes even more desirable. Currently, the definition stage of the OpenMI standard has been completed (<http://www.harmonit.org>) and it is expected that this initiative will lead to a European standard on data exchange. In addition, the HarmonIT project is finalizing a utility library, which is available as public domain software on <http://www.OpenMI.org>. This library will facilitate application builders with the responsibility of connecting various software codes via OpenMI standards. The various organisations in the HarmonIT project are also composing additional tool sets, such as event loggers and data visualizers, which will remain their own proprietary software.

In order to improve the integration of software systems, various groups have to contribute. Fortune (2004) identifies the following OpenMI user groups: non-specialist end user, specialist end user, model integrator, model builder, model coder, application builder and tool coder. In particular, it is important that the demand for better communication between hydroinformatics components is enforced in contracts. As an example, recently the Bundesanstalt für Wasserbau (BAW) in Germany demanded OpenMI compliance for the delivery of the generic Delft3D modelling system of WL | Delft Hydraulics. There is a clear trend now that the principal participants in the HarmonIT project, such as DHI, Wallingford Software and WL | Delft Hydraulics are opening up their standard software packages with OpenMI communication links. In principle, all existing modelling systems can be made



OpenMI compliant. Fast connections can only be made if the source code is available. The adaptation is relatively easy if the code has been programmed in an object oriented way. Executables can be made OpenMI compliant by encapsulating them within a wrapper, a communication layer which transfers data that are accessible through the standard input- and output routines of the code. This, however, may lead to slower communication links.

The advantages of OpenMI were recently explored by Solomon (2005) for the application of ensemble Kalman filtering to reduce uncertainties in the outputs of a rainfall-runoff model. For the ensemble Kalman filtering, the EnKF code developed by was used. For the rainfall-runoff model, use was made of the five parameter, five state variable modelling system HYMOD on an existing model of the 1944 km<sup>2</sup> catchment of the Leaf River Watershed, Mississippi, USA (Vrugt *et al.*, 2003). A comparison was made of connecting both software components via batch files and via OpenMI calls, respectively. Objectives of this exercise were the comparison of results obtained with both methods and the analysis of execution speed differences. It was found that both methods of coupling gave nearly the same results, whereas the OpenMI coupling proved to be 40 % faster than a batch file connection, for a simulation with updates of state variables based upon 30 ensembles (see Fig. 2).



**Fig. 2** Observed, modelled and updated discharges for the OpenMI connection between EnKF and HYMOD



No effort was made to test the speed of a direct implementation of the ensemble Kalman filtering code inside HYMOD. This would go against the trend of designing modelling systems in a modular way by developing different functionalities in separate executables. However, other experiments with OpenMI have revealed that the speed difference between simulation with a fully integrated code and with separate codes in a modular design with exchange of data based upon OpenMI data exchange is not very significant.

## **FLOOD FORECASTING PLATFORM**

Quoting Fortune (2004): *Perhaps the most advanced approach to flexible model integration is taken for flood forecasting applications.* This is, indeed, the area where the need of coupling of components of a hydroinformatics systems is most pronounced. As shown in the sequel, flood forecasting requires many operations between a wide variety of components to be orchestrated in a short time. This leads to the explicit need to create a generic flood forecasting platform, where existing components can be connected and new components be added in a flexible manner. With this objective, the Delft-FEWS (Flood Early Warning System) was developed and implemented in various places in the world.

The Delft-FEWS system takes care of executing the following tasks:

- a) import of external sources of data, such as meteorological forecasts, including those based upon numerical weather models, radar data, rainfall, discharge and water level time series from telemetric systems and data from external databases;
- b) validation and interpolation of incoming data, using extensive data validation options with gap filling and hierarchy rules to allow alternative data sources to be used as a fallback for ensuring continuity in the forecasting process;



- c) data transformation in order to prepare the required inputs for reporting and for the forecasting models, such as weighting of precipitation from distributed point sources, from radar and from numerical weather models, as input to rainfall-runoff modelling;
- d) execution of the hydrologic and hydraulic forecasting models. These models may be provided by various suppliers and cover a wide range of methods, from simple regression analysis, lumped hydrological models, spatially distributed hydrological models, artificial neural networks, hydrological routing models to 1D and integrated 1D-2D hydrodynamic models;
- e) updating the state of the models through a feedback mechanism to minimize the gap between observed and forecasted data. Delft-FEWS provides some of the possible data assimilation models, such as the ARMA error correction method and ensemble Kalman filtering. Delft-FEWS also facilitates the implementation of other updating techniques;
- f) visualisation of results on maps, which can be imported from various sources, such as GIS, aerial photo's etc., including geographic navigation on these maps;
- g) dissemination of forecasts through maps and HTML formatted reports, allowing easy communication to relevant authorities and public through intranet and internet.

An example of the typical requirements of a forecasting agency is described by Werner *et al.* (2004). In 2002, the Environment Agency in the UK commissioned to WL | Delft Hydraulics and Tessella Scientific the development of the National Flood Forecasting System (NFFS) as flood forecasting platform for the complete area of England and Wales. A requirement given was the openness of the system to allow the continued use of various calibrated models which were already operational in flood forecasting systems for a number of river catchments in the area. So far, these comprised the rainfall-runoff models based upon PDM, MRCM, TCM and NAM, the hydrologic routing models based upon DODO and KW and the hydrodynamic modelling systems ISIS and Mike11. Currently, Delft-FEWS also has links to the rainfall-



runoff software HBV, Sacramento, PRMS and VFlo and the hydrodynamic modelling system SOBEK.

NFFS replaces, among others, the earlier FFS2 system developed for the UK Midlands region (Dobson and Davies, 1990). It comprises the MCRM lumped conceptual rainfall-runoff model and the DODO two layer Muskingum routing model, both equipped with updating techniques. These models had to be retained, as they have been extensively calibrated over the past years. The existing telemetry system was equipped with 124 meteorological gauges, 147 hydrological gauges and 272 forecasting points that may or may not coincide with gauge locations. The whole system represents a substantial asset value, of which many components are of great value in the newly installed Delft-FEWS. Once the new system is operational, component models can be replaced by better options if and when these will be acquired.

An immediate advantage of the new system is that forecast lead times increase through the link to more advanced weather forecasting, which is part of the overall platform. A longer term advantage is that gradually the hydrologic and hydraulic models can be replaced by state of the art products, without being bound to one single manufacturer. Similarly, with the overall platform in place, the Environment Agency can gradually increase the number of catchments where forecasting is provided.

## **EXAMPLE OF FLOOD MANAGEMENT DECISION SUPPORT SYSTEM**

Decision support systems aim at facilitating the societal, political and managerial decision making processes with sound engineering knowledge. An interesting example is the Planning Kit, developed for supporting the decision making process of improved flood management along the Rhine branches in The Netherlands (de Vriend and Dijkman, 2003). Due to climatic and land use changes and the increased awareness of their effects, triggered by the 1993 and



















