Combined Hydraulic and Black-Box Models for Flood Forecasting in Urban Drainage Systems

Michael Bruen, M.ASCE¹; and Jianqing Yang²

Abstract: Rapid urbanization and its implications for both water quality issues and floods have increased the need for modeling of urban drainage systems. Many operational models are based on deterministic solutions of hydraulic equations. Improving such models by adding a “black-box” component to deal with any systematic structure in the residuals is proposed. In this study, a conventional deterministic stormwater drainage network model is first developed for a rapidly developing catchment using the HYDROWORKS (now called Infoworks) package, from Wallingford Software in the United Kingdom. However, despite the generally satisfactory results, the HYDROWORKS model tended to underestimate the flow volume. In this paper, a black-box or “systems” model is fitted to the hydraulic urban drainage model in order to improve its overall efficiency. A study was conducted of suitable black-box models, which included the nonlinear artificial neural network model (ANN), and the linear time series models of Box and Jenkins in 1976. They were added to either the output (in simulation mode) or, in updating mode, to the residuals (i.e., difference between modeled and measured output) of the deterministic hydraulic model. The updating procedure provided a considerable improvement in the overall model efficiency for different lead-time forecasting. In simulation mode, however, only the nonlinear ANN model gave better performance in calibration, and a slight improvement in validation.

DOI: 10.1061/(ASCE)1084-0699(2006)11:6(589)

CE Database subject headings: Hydraulic models; Stormwater management; Urban areas; Drainage; Simulation; Floods; Forecasting.

Introduction

Rapid urban growth has been taking place world wide, and its hydrological impacts are well documented (cf. Hall 1984; Shaw 1994). The increase in impervious areas due to urbanization reduces infiltration, increases runoff, and reduces catchment response times, all leading to increased flood peaks and volumes. Water tables in aquifers and base flow in channels are reduced. In addition, urbanization in a catchment can often cause a degradation of water quality (Buren et al. 1997; Bertrand-Krajewski et al. 2000).

Urbanization and its implications for both flood risk and damage and water quality issues have increased the need for modeling of urban drainage systems. Computer models have been extremely useful in devising and assessing the various engineering and operational strategies for managing both water quantity and quality. For instance, Zoppou (2001) discusses the important features of 12 such models, representing a wide range of capabilities and spatial and temporal resolutions. They are all deter-

ministic, most are physically based, and many are event-based models.

Such models can be used (1) to simulate past events (simulation mode) where all of the input and output data for the event have already been recorded and are available to the modeler or (2) in real-time operation in which the individual measured data values become available to the modeler at successive time intervals. In such a case, the model may operate in “updating” mode in which it tries to model its own residuals, i.e., the difference between the model prediction and measured output at each step.

However, there are unavoidable uncertainties associated with mathematical modeling. Willems and Berlamont (1999) demonstrated a number of uncertainties involved in the modeling of a sewer system, including those relating to (1) data inputs, (2) model simplifications of the physical reality, and (3) uncertainties in model parameter values. The first of these, relating to data inputs, generally has a stochastic or random nature principally associated with measurement or sampling errors, but the second generally has a deterministic nature and the third has a mixture of deterministic and stochastic components. Thus, the difference (residual) between modeled and measured output (e.g., flow or water quality indices) contains a mixture of both systematic and stochastic components.

There are a number of possible ways to reduce the residuals and improve the performance of such models. The obvious approach is to address directly the deficiencies in the deterministic model and improve the accuracy of its physical representation of the system being modeled (i.e., improve the system geometric description and/or the physical equations describing the operation of the system). However, increasing the complexity of the model does not always produce better performance because of data limitations and increasing difficulty in estimating model parameters.

¹Director, Centre for Water Resources Research, Civil Engineering Dept., Univ. College Dublin, Earlsfort Terrace, Dublin 2, Ireland (corresponding author). E-mail: michael.bruen@ucd.ie

²Graduate Student, Centre for Water Resources Research, Civil Engineering Dept., Univ. College Dublin, Earlsfort Terrace, Dublin 2, Ireland. E-mail: jianqing.yang@ucd.ie

Note. Discussion open until April 1, 2007. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on December 17, 2004; approved on April 4, 2006. This paper is part of the *Journal of Hydrologic Engineering*, Vol. 11, No. 6, November 1, 2006. ©ASCE, ISSN 1084-0699/2006/6-589-596/\$25.00.

Fig. 1. Black-box simulation improvement on hydraulic model output (x_{sim} =flow calculated by the HYDROWORKS hydraulic model; y_{est} =improved estimate of flow from the combined model; and e =residual, i.e., the difference between the measured flow and that estimate by the combined model)

An alternative approach, developed in this paper, is to fit a “black-box” or “systems” model to either the output of the deterministic hydraulic model or to its residuals. The hypothesis is that a simple black-box model is able to fit the deterministic component of the deficiencies in the hydraulic model. If this approach works, it is likely to be faster and much less expensive than the first approach.

This paper describes the methods and their results based on a case study to test the hypothesis, and is organized as follows: the first section is this introduction. The next section is a brief description of the methodology and models used. The following section describes the case study. Results and discussions are presented in the next section, and the last section lists the conclusions from the study.

Methodology

The method developed in this study is to fit a black-box or systems model to either the output of the deterministic hydraulic model or to its residuals to improve the overall flow forecast performance. The simplest case is when the model is used in simulation mode, i.e., all the input and output data have already been measured and the model is run with the measured input, but has not given any information about the measured output. In this case, the output of the hydraulic model is the input to the black-box model and its output is an improved estimate of the discharge. It is fitted by minimizing the sum of squares of the differences between the output from the black-box model and the measured discharge, Fig. 1. In updating mode (i.e., in real-time forecasting), measurements of both the input and output are available up to a given point in time (typically the time at which

Fig. 2. Black-box model moving average (MA) updating: Base case (x_{sim} =flow calculated by the HYDROWORKS hydraulic model; y_{est} =improved estimate of flow from the combined model; and e =residual, i.e., the difference between the measured flow and that estimate by the combined model)

Fig. 3. Black-box model (AR, ARMA, ANN) updating: Base case. (AR=autoregressive model; ARMA=autoregressive moving average model; and ANN=artificial neural network model)

the model is run) and both are given to the model. As measured values of the discharge are available in real-time, the black-box model can be fitted to the known recent past residuals of the hydraulic model output, Figs. 2 and 3. In this case, the black-box model output is a correction for the hydraulic model output.

An analogous approach has proved successful in the deterministic modeling of natural catchments for real-time flow forecasting (Shamseldin and O'Connor 2001; Brath et al. 2002) but has not yet been applied to hydraulic urban runoff models. In its basic form, it can be used for both flow simulation (off-line), e.g., for risk assessment, flow forecasting, and real-time flood warning/management. However, a further refinement of the method, of particular use in real-time flood forecasting, is to add a forecast updating procedure. This approach was also shown to be effective in rural catchment modeling, using a catchment simulation model coupled with a procedure for forecast updating (WMO 1990, 1992). The overall forecast performance thus can be expected to improve (Serban and Askew 1991), cf. Fig. 4.

In this study, the improvement produced by adding a black-box model to a hydraulic urban drainage model is investigated for a rapidly urbanizing catchment in Dublin, Ireland. First, a conventional stormwater drainage network model is developed to quantify the flow contributions to the existing watercourses in order to assess the risk of flooding. In this study HYDROWORKS version 6.0 (HRWallingford 1997) is used, but any urban drainage hydraulic model could be used. The catchment is a combination of urban and rural areas. This also causes some difficulties in the model development and uncertainties in the model performance. To address these problems, some simple black-box models were added in series with the hydraulic model. The black-box models tested were (1) an artificial neural network (ANN), (2) an autoregressive (AR), (3) a moving-average (MA), and (4) an autoregressive moving average (ARMA) (Box and

Fig. 4. The principal effect of updating (adapted from Serban and Askew 1991). Legend: QM=measured hydrograph; QF=forecasted hydrograph; QS=simulated hydrograph; J =time of preparation of forecast; e =simulation error; and ϵ =forecasting error.

Jenkins 1976). These were tested separately in simulation mode and in adaptive updating mode (Wood and O'Connell 1985).

Hydraulic Model

HYDROWORKS is a widely used urban drainage modeling system developed at the Hydraulics Research Institute, Wallingford, U.K. It is a physically based hydraulic model capable of modeling storm water or sewer systems or combined systems, for both hydraulic and water quality purposes. In the HYDROWORKS package, rainfall data (hyetographs) can either be defined as input (measured data), or be generated as synthetic design rainstorms using a number of techniques (e.g., depth-duration-frequency relationships) depending on the circumstances. The rainfall-runoff model employs a modified rational method, which is essentially the same as the rational method with the inclusion of a routing coefficient (Zoppou 2001). The runoff coefficient of the rainfall-runoff model takes into account the proportion of impervious area, soil type, evapotranspiration, and antecedent conditions. The hydraulic calculations are done with an implicit finite difference method. HYDROWORKS has been applied in several countries for different purposes (cf. Masse et al. 2001; McIntyre 2003) and a detailed description of the model is given in HRWallingford (1997).

Black-Box Models

There are various forms of forecast improvement procedures used in practice, depending on which variables are modified (Serban and Askew 1991). For instance, the output may be modified (Brath et al. 2002), the model parameters or its internal states may be modified (Sholl and Wolfe 1985), or the inputs may be modified as in the Hydrologiska Byråns Vattenbalansavdelning (named after the Hydrological Bureau Waterbalance Section of the Swedish Meteorological and Hydrological Institute) (known internationally as HBV) model (Lindström et al. 1997). In this study the first method is used and the model output is modified by passing it through a black-box model. The black-box models chosen for the study include both linear and nonlinear models. They are used to update the simulation model outputs to improve performance both in direct simulation (Fig. 1) and in real-time forecasting (Figs. 2 and 3).

Linear Time Series Models

The classical time series models of Box-Jenkins (1976), including AR, MA, and their combination ARMA, are linear in that predicted values are linear functions of current and/or past observations. Because of their relative simplicity, these are highly popular in many fields, including hydrology and economics (Zhang 2003).

Equations of the time series models in both simulation and updating operations are as follows:

For simulation

$$\text{AR model: } \hat{y}_t = \alpha_1 \hat{y}_{t-1} + \alpha_2 \hat{y}_{t-2} + \cdots + \alpha_p \hat{y}_{t-p} + \beta_0 x_t \quad (1)$$

$$\text{MA model: } \hat{y}_t = \beta_0 x_t + \beta_1 x_{t-1} + \cdots + \beta_q x_{t-q} \quad (2)$$

$$\text{ARMA model: } \hat{y}_t = \alpha_1 \hat{y}_{t-1} + \alpha_2 \hat{y}_{t-2} + \cdots + \alpha_p \hat{y}_{t-p} + \beta_0 x_t + \beta_1 x_{t-1} + \cdots + \beta_q x_{t-q} \quad (3)$$

For updating

$$\text{AR model: } \hat{e}_t = \alpha_1 \hat{e}_{t-1} + \alpha_2 \hat{e}_{t-2} + \cdots + \alpha_p \hat{e}_{t-p} + a_t \quad (4)$$

$$\text{MA model: } \hat{e}_t = \beta_0 x_t + \beta_1 x_{t-1} + \cdots + \beta_q x_{t-q} \quad (5)$$

$$\text{ARMA model: } \hat{e}_t = \alpha_1 \hat{e}_{t-1} + \alpha_2 \hat{e}_{t-2} + \cdots + \alpha_p \hat{e}_{t-p} + \beta_0 x_t + \beta_1 x_{t-1} + \cdots + \beta_q x_{t-q} \quad (6)$$

where x_t =deterministic hydraulic model output at time t ; y_t =measured discharge at time t ; $\hat{y}_t$ =final estimated discharge (flow estimates after the black-box model improvement) at time t ; $\hat{e}_t$ =residual of the hydraulic model output at time t ; α_j ($j=1, 2, \dots, p$) and β_i ($i=0, 1, 2, \dots, q$)=finite set of weight parameters; and a_t =independent, random term. The integers p and q =orders of the autoregressive and moving-average components, respectively, and are determined by fitting a range of models with different values of p and q in model calibration.

The AR model, Eqs. (1) and (4), is a special class of the family of time series models defined by Box and Jenkins (1976). Structurally, the AR model is possibly the simplest model of the Box-Jenkins family. The model output is a linear combination of a finite number of previous model outputs plus an uncorrelated random term. It is included in this study, partly because its process has a natural systems interpretation, and partly because it is flexible and easy to apply. The MA model, Eqs. (2) and (5) can be thought of as a linear filter with a transfer function when the input is white noise. The ARMA, Eqs. (3) and (6), is a flexible mixture of an AR process and a finite MA component. These models represent the basic building blocks of the Box-Jenkins time series modeling approach (Box et al. 1994). To provide a stationary process, it is necessary for the processes in the models to remain in equilibrium about a constant mean level. This is achieved by transforming the data series by subtracting its mean value.

Artificial Neural Networks

In addition to the classical linear models, an artificial neural network is used in this study to account for possible, more complex, nonlinear relationships. ANNs are powerful tools for solving a wide range of practical problems and they have been increasingly applied to predict and forecast water resources variables (French et al. 1992; Shamseldin 1997; Bodri and Cermak 2000). They have been used for event-based stormwater forecasting with short time steps (Yang and Bruen 2003). Basic and attractive characteristics of ANNs is their ability to learn from data and their potential to accurately describe the behavior of complex nonlinear systems. Their data-driven approach is useful for analyzing field data sets where no prior guidance is available to suggest either a linear or nonlinear relationship. Neural network techniques are useful for time series modeling and forecasting because they do not require detailed prior knowledge of the underlying physical processes.

An ANN is made up of several layers of neurons (nodes) which are interconnected by links using weights (parameters). Multilayer feedforward neural networks are the most popular and widely used network paradigm (Maier and Dandy 2000). These contain one input layer, one output layer, and one or more intermediate hidden layers (cf. Fig. 5). Signals pass through the network in the forward direction only, i.e., start from the input layer, move to the hidden layer (or layers), and then to the output layer. The input neurons transmit external signals across the links to the next layer (hidden or output) of neurons. In the hidden and output layers, each neuron sums the weighted values (signals) from the previous layer and puts the result through a transfer function to give a scalar output value (signal). The neural transfer function is also called an activation function. Some forms of transfer func-

Fig. 5. Three-layer feedforward neural network used for updating forecasting

tion popular in current use are linear, logistic (i.e., sigmoid), and hyperbolic tangent. Of these, the logistic function is the most commonly employed (Haykin 1994), and is used in this study.

Fig. 5 depicts an example of a three-layer feedforward neural network used in the study for residual updating forecasts, where the inputs to the ANN are residuals of the previous hydraulic simulation model output ($Y_{obs} - X_{sim}$) up to the time of forecast, and the outputs are the forecasted residuals at the required lead times. The forecasted residuals are then added to the hydraulic model output to give the updated forecasts. The appropriate numbers of input and hidden neurons are determined by fitting and comparing the performance of a range of different model structures during the model calibration (training) phase. The conjugate gradient (CG) optimization algorithm (Press et al. 1986) is used for model training (calibration) in which it minimizes the sum of squares of the difference between the model output and the measured values.

Case Study

Citywest catchment is a new development area, situated on the west side of Dublin, with industrial, business, and residential components. Upstream of the catchment are the foothills of the Dublin Mountains, which are located to its south. The catchment is elongated with a steep slope and a quick response to rainfall. Its area is 1.83 km², approximately 1.8 km long by 1 km wide and the average runoff lag time is estimated from recorded thunderstorms as 39.4 min, and the minimum lag time of any recorded individual storm event is approximately 20 min. At the time of this study, the average impervious area covered approximately 25.4% of the total catchment area.

The catchment is drained for the most part by a pre-existing open channel passing through the center of the catchment, into

Fig. 6. Black-box model simulation on the Hydraulic model output: (a) calibration; (b) validation

which the new pipe drainage network flows. The Citywest stormwater drainage system includes below ground pipes and above ground ponds and channels used to convey stormwater. The modeled network has 309 nodes in total including channels, which gives 310 connections with a total conduit length of 16.5 km. The system serves an area of 1.83 km² with pipe sizes ranging from 225 mm up to 1,200 mm. Rain-gauges and water-level recorders were installed in the catchment to provide measurements of rainfall and to estimate the flows in open channels downstream of the outlets of the pipes, where small weirs established a stable rating relationship. Storm events were extracted from the record for modeling purposes.

Results and Discussions

The hydraulic simulation results and black-box model updated results are discussed separately, and then compared.

Table 1. Simulation Results of Basic Hydraulic Model versus Observed Storm Events at Catchment Outlet

Period	Storm event	Peak flow (m ³ /s)			Storm volume (m ³)		
		Hydraulic model	Measured	Difference (%)	Hydraulic model	Measured	Difference (%)
Calibration	E020822	1.188	0.974	+21.9	4,401.3	4,837.0	-9.0
	E030117	1.065	0.974	+9.3	11,339.3	12,597.2	-9.9
	E030208	0.342	0.337	+1.4	9,411.6	9,462.0	-0.6
Validation	E020808	0.312	0.313	-0.3	10,757.7	10,143.6	+6.1
	E020813	0.525	0.531	-1.1	2,746.7	3,028.4	-9.3
	E030124	0.305	0.343	-11.1	6,248.8	6,628.4	-5.8

Table 2. Results of Black-Box Model Simulation Improvement on HYDROWORKS at Catchment Outlet

Model ^a	Calibration		Validation	
	R^2	MSE	R^2	MSE
Base hydraulic model (HW)	85.29	1.01E+04	83.24	2.72E+03
HW+ANN(32,26,1)	96.76	2.29E+03	85.64	2.33E+03
HW+AR(7)	90.70	6.46E+03	80.18	3.21E+03
HW+ARMA(3,12)	91.53	5.88E+03	80.20	3.21E+03
HW+MA(21)	92.11	5.53E+03	79.68	3.29E+03

^aANN(l, m, n), where, l , m , and n represent input, hidden, and output layer sizes, respectively; AR(p), MA(q), and ARMA(p, q), where p and q represent order of autoregressive and moving-average components, respectively.

Simulation with the Hydraulic Model

The hydraulic model is calibrated and validated using part of the measured storm data. Moderate to heavy storms were chosen with rainfall amounts from 5.2 to 13.0 mm per event. A split-sample technique for model calibration and validation was used. Six storm events are used in these tests, with three used for calibration and three for validation. The model was simulated with 2 min time steps.

The hydraulic model parameters, including geometric information such as sizes, slopes, and connections of the pipes; and hydraulic factors such as friction coefficients, etc. have a full physical meaning and values can be directly given to these parameters (Willems and Berlamont 1999). Therefore, in model calibration, most effort was directed at determining the effective runoff area for each node and the parameters of the runoff model. The model was evaluated by comparing the simulation results with the measured flow in terms of peak flow, event volume, and hydrograph shape.

The simulation results at the drainage outlet are summarized in Table 1. These show a generally acceptable agreement between the hydraulic simulation and the observed storms for all events. However, they also indicate a consistent trend toward underestimating flow volumes, in particular, for the recession part of the hydrographs, for example, compare the HYDROWORKS output with the measured flows in Fig. 6. This may be caused by uncertainties or inadequacies in either the modeling approach and/or the available data, as previously discussed. For instance, there may be some leakage of water from the unsaturated soil zone into the pipe or channel network or it may be due to temporal and spatial variability of rainfall across the catchment. Regardless of its origin, it clearly has a deterministic component and so a

Table 3. Lag-1 (2 min) Forecasting Performance

Updating model	Calibration		Validation	
	R^2	MSE	R^2	MSE
Base hydraulic model (HW)	85.29	1.01E+04	83.24	2.72E+03
HW+ANN(8,12,1)	99.26	5.16E+02	98.45	2.47E+02
HW+AR(4)	99.47	3.67E+02	99.53	7.53E+01
HW+ARMA(3,5)	99.61	2.73E+02	99.49	8.13E+01
HW+MA(0,27)	92.45	5.27E+03	83.37	2.66E+03

Table 4. Lag-2 (4 min) Forecasting Performance

Updating model	Calibration		Validation	
	R^2	MSE	R^2	MSE
Base hydraulic model (HW)	85.29	1.01E+04	83.24	2.72E+03
HW+ANN(8,12,1)	97.65	1.64E+03	95.57	7.07E+02
HW+AR(4)	98.21	1.25E+03	98.67	2.11E+02
HW+ARMA(3,5)	98.70	8.99E+02	98.84	1.87E+02
HW+MA(27)	91.80	5.72E+03	82.98	2.72E+03

deterministic systems model is added to the hydraulic model output with the expectation that it will improve the overall model performance.

Black-Box Models Improved Results

Black-box models added to the basic hydraulic simulation model output were tested in both simulation and real-time forecasting (updating) modes. As stated earlier, the split sample test is used for the black-box model calibration. The storm data are divided into a calibration and a validation set. In calibrating the models, the CG algorithm (Gill et al. 1981) is used for the ANN model training (calibration), whereas the simple least-squares method is applied to the linear Box-Jenkins models. The overall model performance was evaluated using the Nash and Sutcliffe (1970) R^2 efficiency and mean square error (MSE) criteria.

Simulation Mode Operation

For each black-box model, a large number of different orders were tried. Table 2 presents the results for simulation mode operation of the hydraulic plus black-box models for the best performing model in each class. The addition of any of the black-box models provided considerable improvement in simulation performance in calibration. In validation, however, only the ANN model improved on the basic hydraulic model output, in terms of R^2 efficiency and the MSE criteria. These were calculated by adding together the contribution from each time step for all events considered, i.e., as if the events occurred sequentially in time. Fig. 6 compares the simulation hydrographs for the hydraulic model with and without the ANN model in both calibration and validation periods. Note that, for the validation period, the improvement is mainly for the recession and there is more variability in the shape at the peak. Thus this technique is of little benefit here and cannot be recommended for simulation mode.

Real-Time Forecasting Mode Operation

In contrast, the real-time (updating) forecasting results are presented in Tables 3–5 for different forecast lead times. The lead

Table 5. Lag-15 (30 min) Forecasting Performance

Updating model	Calibration		Validation	
	R^2	MSE	R^2	MSE
Base hydraulic model (HW)	85.29	1.01E+04	83.24	2.72E+03
HW+ANN(8,12,1)	85.29	1.03E+04	70.03	4.79E+03
HW+AR(4)	88.58	7.98E+03	85.89	2.25E+03
HW+ARMA(3,5)	91.16	6.13E+03	89.06	1.76E+03
HW+MA(27)	87.33	8.84E+03	81.22	3.00E+03

Fig. 7. Performance of updating procedures at different lead times: (a) calibration; (b) validation

time is defined as the time from the making of the forecast to the time at which a forecasted value is expected to occur. It is clear that all the four black-box updating models performed well with short lead times. All except the MA model provided considerable improvement over the nonupdating hydraulic model in both calibration and validation periods for a wide range of lead times. Tables 3 and 4 show the model performance for lead-times of 2 and 4 min. The three updating procedures, i.e., ANN, AR,

ARMA, are a significant improvement on the raw output of the hydraulic model. Overall, the ARMA model gave the best performance for all cases in both calibration and validation, whereas the MA model had the worst performance among the updating models. However, the performance of all models declines as the lead-time increases. This is natural and expected. Table 5 shows that the AR model updating provides an improvement, over the basic hydraulic model, for up to at least 30 min (15 time steps), Fig. 7. This is significant when compared to the 39.4 min average catchment lag-time and 20 min minimum lag time of any recorded individual storm event. Thus the updating procedure could provide substantially improved information for the operational and automatic control of stormwater drainage systems. Figs. 8 and 9 illustrate the validation forecast hydrographs for storm event of January 24, 2003 (E030124) for lead times of 2 and 8 min. Only the nonlinear ANN updating procedure and the best linear model (ARMA) are shown in Figs. 8 and 9 and are compared with the raw hydraulic model forecasts. Figs. 8 and 9 show the improvement made by the updating procedures over the entire hydrograph (peak and recession). Note that only the updating models track the measured recession, in particular its later part which decreases very slowly for this case, possibly augmented by unrecorded inflows from minor rainfalls on parts of the catchment far from the raingauges after the main event. The results also indicated that, as an updating procedure, the simple linear time series model is as good as or even better than the nonlinear ANN model in this study.

Conclusions

In this study, a model combination technique was developed to improve the performance of hydraulic models for urban drainage systems using simple black-box models. This technique was applied to a stormwater drainage system in a small urbanizing catchment using the HYDROWORKS modeling software package. Four black-box models were tested as an updating procedure to improve on the output of the hydraulic model for real-time forecasting. Updated forecasts were calculated for a range of lead times and the overall model efficiencies were compared. The ap-

Fig. 8. Two minute lead-time forecasting using black-box updating procedure coupled with hydraulic simulation model

Storm event E030124 in validation

Fig. 9. Eight minute lead-time forecasting using black-box updating procedure coupled with hydraulic simulation model

proach can provide improved information for the operational and real-time control of stormwater drainage systems. It is useful for small urban catchments and/or for catchments with steep slopes and with short catchment response times. The approach could be applied to any urban drainage hydraulic model. Due to the effects of discontinuous data (i.e., event storms) as well as the small time-step utilized in the study, relatively large size black box models were used in both simulation and updating cases (cf. Tables 2–5). However, this is acceptable since automatic optimization procedures are used in the model calibration. In comparison with the two types of model operations (simulation and updating) using the black-box models, a simple linear model is sufficient for forecasting updating but a non-linear model is the best for simulation. This is most likely due to the use of previously measured flows to define the residuals which are the input in the updating procedure, while there are no directly measured flows input to the black-box models in simulation mode. The following conclusions may be drawn from the study.

The basic hydraulic stormwater drainage system model performed well in comparison with the observed data at the catchment outlet. All the updating models, except the MA model, provided a considerable improvement on the overall model efficiency for different lead times. This is due, in part, to the use of previously measured flows to define the residuals which are the input to the updating procedure, whereas there are no directly measured flows used as input to the black-box models in simulation mode. The ARMA updating procedure provided the best performance in lead-time forecasting in both calibration and validation, whereas the MA model provided the worst results in this study. Simple black-box models, used for updating the output of a physically based hydraulic model in real-time forecasting, are capable of improving flow forecasts for a wide range of lead times in urban drainage systems. Either a linear or nonlinear black-box model may be used for such a purpose. However, in simulation mode operation, the improvement on model performance was achieved in calibration by the addition of a black-box model. This is to be expected, but only the nonlinear ANN model produced a better R^2 and MSE in validation and this reflected an improved modeling of the recession, but not of the peak flow. Thus, this method is not recommended when only off-line simulation results are required.

Acknowledgments

The study is supported by Fahy Fitzpatrick Consulting Engineers, Dublin, Ireland, on behalf of Citywest developments and by the Centre for Water Resources Research at University College Dublin (UCD).

Notation

The following symbols are used in this paper:

- a_t = an independent, random term;
- $\hat{e}_t$ = residual of the hydraulic simulation model output at time t ;
- x_t = hydraulic simulation model output at time t ;
- y_t = measured discharge at time t ;
- $\hat{y}_t$ = final estimated discharge (updated flow forecasts from black-box model) at time t ;
- α_j = $j=1, 2, \dots, p$, are a finite set (order of p) of weight parameters; and
- β_i = $i=0, 1, 2, \dots, q$, are a finite set (order of q) of weight parameters.

References

- Bertrand-Krajewski, J., Barraud, S., and Chocat, B. (2000). "Need for improved methodologies and measurements for sustainable management of urban water systems." *Environ. Impact. Assess. Rev.*, 20(3), 323–331.
- Bodri, L., and Cermak, V. (2000). "Prediction of extreme precipitation using a neural network: Application to summer flood occurrence in Moravia." *Adv. Eng. Software*, 31(5), 311–321.
- Box, G. E. P., and Jenkins, G. M. (1976). *Time series analysis—Forecasting and control*, Revised Ed., Holden-Day Inc., San Francisco.
- Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994). *Time series analysis—Forecasting and control*, 3rd Ed., Prentice-Hall, Englewood Cliffs, N.J.
- Brath, A., Montanari, A., and Toth, E. (2002). "Neural networks and non-parametric methods for improving real-time flood forecasting

- through conceptual hydrological models." *Hydrology Earth Syst. Sci.*, 6(4), 627–640.
- Buren, M. A. V., Watt, W. E., and Marsalek, J. (1997). "Application of the log-normal and normal distributions to stormwater quality parameters." *Water Sci. Technol.*, 31(1), 95–104.
- French, M. N., Krajewski, W. F., and Cuykendall, R. R. (1992). "Rainfall forecasting in space and time using a neural network." *J. Hydrol.*, 137(1/4), 1–31.
- Gill, P. E., Murray, W., and Wright, M. H. (1981). *Practical optimization*, Academic, London.
- Hall, M. J. (1984). *Urban hydrology*, Elsevier Applied Science Publishers, London.
- Haykin, S. (1994). *Neural networks—A comprehensive foundation*, Macmillan College, New York.
- HRWallingford. (1997). *HYDROWORKS: Online manual*, Oxon, U.K. (<http://www.oxtowns.co.uk/wallingford/home.html>) (Oct. 2002).
- Lindström, G., Johansson, B., Persson, M., Gardelin, M., and Bergström, S. (1997). "Development and test of the distributed HBV-96 model." *J. Hydrol.*, 201(1–4), 272–288.
- Maier, H. R., and Dandy, G. C. (2000). "Neural networks for the prediction and forecasting of water resources variables: A review of modelling issues and applications." *Environ. Modell. Software*, 15(1), 101–124.
- Masse, B., Zug, M., Tabuchi, J. P., and Tisserand, B. (2001). "Long term pollution simulation in combined sewer networks." *Water Sci. Technol.*, 43(7), 83–89.
- McIntyre, K. (2003). "Aspects of storm water management." *Urban Hydrology Stormwater Management in Proc., National Hydrology Seminar*, Tullamore, Ireland, Office of Public Works, Dublin, Ireland.
- Nash, J. E., and Sutcliffe, J. V. (1970). "River flow forecasting through conceptual models. Part I: A discussion of principles." *J. Hydrol.*, 10(3), 282–290.
- Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1986). *Numerical recipes—The art of scientific computing*, Cambridge University Press, Cambridge, U.K., 301–353.
- Serban, P., and Askew, A. J. (1991). "Hydrological forecasting and updating procedures." *Proc., IAHS Conf. on Hydrology for the Water Management of Large River Basins*, Vienna, Publication No. 201, IAHS Press, Wallingford, U.K., 357–379.
- Shamseldin, A. (1997). "Application of a neural network technique to rainfall-runoff modelling." *J. Hydrol.*, 199(3–4), 272–294.
- Shamseldin, A., and O'Connor, K. M. (2001). "A non-linear neural network technique for updating of river flow forecasts." *Hydrology Earth Syst. Sci.*, 5(4), 577–597.
- Shaw, E. M. (1994). *Hydrology in practice*, 3rd Ed., Chapman and Hall, London.
- Sholl, P., and Wolfe, R. K. (1985). "The Kalman filter as an adaptive forecasting procedure for use with Box-Jenkins ARIMA models." *Comput. Ind. Eng.*, 9(3), 247–262.
- Willems, P., and Berlamont, J. (1999). "Probabilistic modelling of sewer system overflow emissions." *Water Sci. Technol.*, 39(9), 47–54.
- Wood, E. F., and O'Connell, P. E. (1985). "Real-time forecasting." *Hydrological forecasting*, M. G. Anderson and T. P. Burt, eds., Wiley, New York, 505–558.
- World Meteorological Organisation (WMO). (1990). "Hydrological models for water-resources system design and operation." *Operational Hydrology Rep. No. 34* (WMO-No. 740), Geneva, Switzerland.
- World Meteorological Organisation (WMO). (1992). "Simulated real-time intercomparison of hydrological models." *Operational Hydrology Rep. No. 38* (WMO-No. 779), Geneva, Switzerland.
- Yang, J., and Bruen, M. (2003). "Artificial neural network modelling of runoff from storms in urban areas." *Proc., 7th Int. Conf. on the Application of Artificial Intelligence to Civil and Structural Engineering*, B. H. V. Topping, ed., Civil-Comp Press, Edinburgh, U.K., 129–130.
- Zhang, G. P. (2003). "Time series forecasting using a hybrid ARIMA and neural network model." *Neurocomputing*, 50, 159–175.
- Zoppou, C. (2001). "Review of urban storm water models." *Environ. Modell. Software*, 16(3), 195–231.