Data-based mechanistic rainfall-runoff continuous-time modelling in urban context

D. KUSS * **V. LAURAIN** ** **H. GARNIER** ** **M. ZUG** *** **J. VAZQUEZ** *

* *Ecole Nationale du Génie de l'Eau et de l'Environnement de Strasbourg, UMR Hydraulique urbaine, 1 quai Koch, 67070 Strasbourg Cedex, France*
** *Centre de Recherche en Automatique de Nancy, Nancy-Université, CNRS, BP 239, 54506 Vandoeuvre-les-Nancy Cedex, France*
*** *Veolia Water, Direction Technique Ouest, 3 rue Marcel Sembat 44925 Nantes Cedex 9, France*

Abstract: This paper presents a data-based mechanistic modelling (DBM) approach to rainfall-runoff modelling based on the direct identification and estimation of continuous-time models from discrete-time series. It is argued that many mechanistic model parameters are more naturally defined in the context of continuous-time, differential equation models. As a results, there are advantages if such models are identified directly in this continuous-time form rather than being formulated and identified as discrete-time models. An illustrative example based on the analysis of rainfall-flow data from the Kerinou catchment in Brest demonstrates the relevance of the proposed modelling approach.

Keywords: data-based mechanistic, contiunous-time models, rainfall-runoff modelling, urban hydrology

1. INTRODUCTION

The integrated urban wastewater system has a major impact on the quality of the receiving waters, especially during wet weather. During this period, in the case of combined systems, part of wastewaters collected are discharged at the level of combined sewer overflows, causing a degradation of the quality of receiving waters (rivers, sea,...).

The Water Framework Directive (WFD) adopted on October 23, 2000 by the European Union defines quality objectives for all water masses. A good ecological and chemical status is required before 2015 for all surface waters, including rivers and coastal waters. For good chemical status, this means that the concentrations of pollutants in the water body must not exceed the limit values specified in the directive.

Moreover the new directive on bathing waters (2006/7/EC) is more restrictive in terms of water quality than the previous one (76/160/EEC). Bathing water quality is assessed on the basis of two bacteriological criteria. As raw wastewaters are the major source of fecal contamination of coastal waters, the discharge of wastewaters at the level of combined sewer overflows (CSOs) is very harmful for the quality of the receiving waters.

Many recent studies have shown the necessity to implement an integrated management of urban wastewater systems in order to minimize the load of discharged pollutants, (see *e.g.* Vanrolleghem et al. [2005], Butler and Schütze [2005]). The Integrated Management of Coastal Wastewater Discharges project was born in 2006 in this context. This project is managed by Veolia water© and involves some French environmental companies and research institutes. Two sites have been selected to carry out the experimental studies; both are located in France: the first is in Brittany (Saint-Malo and Brest) while the second is in the Mediterranean sea. The main objective is to minimize the impacts of the urban wastewaters discharged in coastal wasters thanks to an integrated real-time control of the urban wastewater system.

Before minimizing the load of discharged pollutants, a first study consists in assessing the volume, frequency, and flow rate of the wastewaters discharged in the receiving waters and their corresponding impacts. In order to reach this goal, measurements of flow rate and pollution are crucial all along the urban wastewater system. Those measurements permit to design urban drainage network models that will be used to test new control and management strategies. The paper presents a data-based approach to rainfall-runoff modelling in urban context.

In most urban hydrology modelling softwares, the rainfall-runoff relationship is approximated with lumped conceptual continuous-time models (Ministère de l'écologie et du développement durable (MEDD) and CERTU [2003]). Here, the *a priori* conceptual model structure is effectively a (normally simple) theory of hydrological behaviour based on the perception of the hydrologist/modeller. The models are formulated on the basis of natural laws and are often expressed in terms of continuous-time linear or nonlinear differential equations. However, the resulting model is strongly conditioned by assumptions that derive from current hydrological paragidms.

An alternative modelling approach known as data-based mechanistic modelling (DBM) is used here (see *e.g.* Young [1998]). In DBM modelling, the most parametrically efficient (parsimonious) model structure is first inferred statistically from the available time series data in an inductive manner. After this initial back-box modelling stage is complete, the model is interpreted in a physically meaningful, mechanistic manner based on the nature of the system under study. Paradoxically, even if most mathematical models of hydrological systems are formulated in continuous-time, the model to be identified are almost always presented in the alternative discrete-time form. One reason for this paradox is that hydrological data are normally sampled at regular intervals over time, forming discrete time series that are in a more appropriate form for DT modelling. Another ground is that most of the technical literature on the statistical identification and estimation (calibration) of transfer function models deals with these discrete-time models. Closer review of this literature, however, reveals apparently less well-known publications (see *e.g.* Young and Garnier [2006], Garnier and Wang [2008], Sinha and Rao [1991]) dealing with estimation methods that allow the direct identification of continuous-time models from discrete-time data.

The present paper outlines an approach to DBM of urban rainfall-runoff data based on the recent advanced direct method of continuous-time model identification and estimation (Garnier and Wang [2008]). Previous DBM modelling of rainfall-flow data has confirmed many aspects of earlier hydrological research and identified that the rainfall-flow model can be stated as an 'Hammerstein' nonlinear form, with an input nonlinearity that converts the measured rainfall into effective rainfall (*i.e.* the rainfall that is effective in causing variations in flow). This effective rainfall then passes through the linear transfer function to yield the river flow.

A DBM-based approach is developed to jointly identify both linear and nonlinear parts of the continuous-time model. This method is then used to investigate the nonlinear relationship within the context of rainfall and flow time series from the Kerinou catchment in Brest, France.

2. DATA-BASED MECHANISTIC CONTINUOUS-TIME MODELLING

Every rainfall-runoff model requires two essential components (Desbordes [1994], Beven [2003]):

- the runoff production component which determines how much of the rainfall becomes part of the hydrograph;
- the runoff routing component which determines the distribution and the transfer time of the surface runoff up to the outlet of the catchment.

A schematic representation of the components of an urban rainfall-runoff model is depicted in Figure 1. If we assume that continuous losses are proportional to the intensity of the rainfall, the rainfall-runoff relationship can be broken down in a non-linear part including only the initial losses and a linear part including the continuous losses and the runoff routing component.

Thus, the objectives of the proposed approach are:

- to use a simple model for the estimation of the initial losses compatible with real-time modelling;
- to identify and estimate the continuous losses and the runoff routing model component by using recent advanced direct continuous-time modelling approaches.

The measured data used to identify a rainfall-runoff model consists of time-series data of total rainfall intensities (mmh^{-1}) for the input(s) and time-series data of flow rates (m^3s^{-1}) at the outlet of the catchment.

According to the size of the catchment, it can be useful to identify multiple input single output (MISO) rainfall-runoff models including more than one input (*i.e.* one rain gauge) in order to take into account the spatial distribution of the rainfall. Indeed, modelling a rainfall-runoff relationship on the basis of only one input is in some cases dangerous because it is thus assumed that the rainfall fallen in the whole catchment is the same than the one measured in one discrete point. For this reason, referring to the literature, the distance between two rain gauges that have to be installed in urban areas should be comprised between 1 to 4 km (Einfalt et al. [1998], Berne et al. [2004]). In urban hydrological models, a possible method to take into account the nearest rain gauges is the method of the Thiessen polygons which delimit the area of influence of each rain gauge. This method consists in drawing on a base map polygons around each rain gauge such that any location within a particular polygon is nearer to the rain gauge within the polygon than to any other rain gauge.

2.1 Initial loss model

Initial losses refer to hydrological processes occurring at the beginning of a rainfall event, such as interception by the canopy or retention in surface depressions. Models of initial losses used in urban areas are generally function (Chocat [1981], Falk and Niemczynowicz [1978]):

- of a parameter independent of the rainfall and representative of the surface of the catchment, for example the slope,
- or of a parameter representing the initial wetness of the catchment at the beginning of the rainfall event.

Nevertheless, the use of such models requires to preliminarily divide the rainfall chronicle in independent rainfall events in order to identify the beginning of each rainfall event. This step carried out manually is time-consuming and is not suited to real-time modelling.

In urban areas, values of initial losses may not exceed 2 to 3 mm (Brulé et al. [1997]) and can sometimes be neglected because interception losses due to canopy and retention in surface depressions are less important than in natural areas.

Initial losses introduce non-linearities in the rainfall-runoff relationship. If their effects are important, for example in the case of mixed or rural catchments, the proposed approach considered here might not be the most appro-

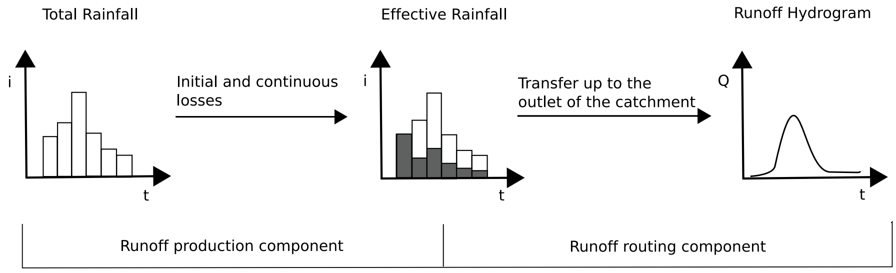

Fig. 1. Schematic representation of the components involved in urban rainfall-runoff

priate and two of the authors of this paper are currently investigating some new approaches which aim to identify directly a non-linear Hammerstein continuous-time Box-Jenkins model (Laurain et al. [2008]).

For urban areas we suggest a simple algorithm that transforms the total rainfall intensity $i_T(t)$ in an "effective" rainfall intensity $i_E(t)$ where the initial losses are deducted from the total rainfall.

As rainfall events are not manually detected, we have to find an automatic procedure that realizes this task. The automatic detection of the beginning of a rainfall event is difficult, even impossible to realize. To get around this issue, for a given time-step j we defined a sliding window of duration d, starting at time-step $j-d$ and ending at time step j. In this sliding window, if the total rainfall intensity of the current time-step $i_T(j)$ (mmh^{-1}) is positive, reference initial losses L_{in} (mm) are compared with the sum of the rainfall height $S(j)$ (mm) fallen between time-step $j-d$ and time-step j :

$$\begin{cases} \text{if } i_T(j) > 0 \text{ and } S(j) < L_{in} \Rightarrow i_E(j) = 0 \\ \\ \text{if } i_T(j) > 0,\ S(j) < L_{in} \text{ and } S(j+1) \geq L_{in} \\ \quad \Rightarrow i_E(j+1) = \dfrac{(S(j+1) - L_{in})}{\Delta t} \\ \text{else } i_E(j) = i_T(j) \end{cases} \tag{1}$$

with:

$$S(j) = \sum_{j-d}^{j} i_T(j)\Delta t \text{ , where } \Delta t \text{ is the sampling time} \tag{2}$$

This approach has a physical meaning but requires the specification of two user-parameters: L_{in} which represents the reference initial losses and d which denotes the length of the sliding window necessary for the evaporation of reference initial losses stored in soil depressions or in the canopy. An approach to automatically estimate these two user parameters will be presented in Section 2.4.

2.2 Direct continuous-time linear model estimation

Dynamic systems in the physical world are naturally described in continuous-time (CT), differential equation terms because the physical laws have been evolved mainly in this form. Paradoxically, however, the best known system identification schemes have been based on discrete-time (DT) models, without much concern for the merits of natural continuous-time model descriptions and their associated identification methods. In fact, the development of CT system identification techniques occurred in the last century, before the development of the DT techniques, but was overshadowed by the more extensive DT developments (Garnier and Wang [2008]).

Moreover, identification of continuous-time models can be done either with a direct or indirect approach. In the direct approach, a continuous-time model is directly identified from the sampled data. In the indirect approach a discrete-time model is first identified and estimated. This estimated DT model is then converted to a CT model, using some assumption about the nature of the input signal over the sampling interval.

Direct identification of CT models proved to have many advantages (Garnier and Wang [2008], Rao and Garnier [2002]):

- The model parameters are unique: unlike discrete-time models, they are not a function of the data sampling interval.
- The model is in an ordinary differential equation form that can be related directly to the formulation of physically meaningful models, such as those derived from mass, energy and momentum conservation.
- Model parameter estimation is superior over a wider range of sampling intervals, particular at fast sampling intervals (*e.g.* 5 min. for rainfall-flow measurements).

These advantages, together with the parsimony that is a natural consequence of DBM transfer function modelling, should mean that any relationships between the CT model parameters and physical measures of the catchment characteristics should be clearer and better defined statistically, as required for hydrological system applications.

The theoretical basis for the statistical identification and estimation of linear, continuous-time models from discrete-time data can be outlined by considering the case of a linear, single input, single output system:

$$q(t) = \frac{B(s)}{A(s)} i_E(t-\tau) \tag{3}$$

$i_E(t)$ denotes the input signal, *i.e.* the intensity of the total rainfall minus the initial losses (mmh^{-1}). $q(t)$ is the output signal, *i.e.* the unmeasured noise free runoff flow rate (m^3s^{-1}). τ is a time delay in time units representing the delay between the input and the output signal. Note consequently, that the continuous-time model that has to be identified also includes the continuous losses. $A(s)$ and $B(s)$ are polynomials in the derivative operator$s$ ($s^n x(t) = \dfrac{d^n x(t)}{dt^n}$):

$$A(s) = s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n \quad (4)$$
$$B(s) = b_0 s^m + b_1 s^{m-1} + \cdots + b_{m-1}s + b_m \quad (5)$$

n and m can take on any positive integer values with $m \leq n$. Of course, the model can also be written in the usual following differential equation form:

$$\frac{d^n q(t)}{dt^n} + a_1 \frac{d^{n-1} q(t)}{dt^{n-1}} + \cdots + a_n q(t) = b_0 \frac{d^m i_E(t-\tau)}{dt^m} + \cdots + b_m i_E(t-\tau) \quad (6)$$

In the case of uniform sampling, at a constant sampling interval T_s, the sampled signals will be denoted by $i_E(t_k)$ and $Q(t_k)$ and the output observation equation can be written in the following form:

$$Q(t_k) = q(t_k) + \xi(t_k) \quad k = 1, \ldots, N \quad (7)$$

where $\xi(t)$ is a discrete-time colored noise associated with a sample output measurement $Q(t_k)$. It is assumed to be independent of the input signal $i_E(t)$ and represents the aggregate effect of all the stochastic inputs to the system, including distributed unmeasured inputs, measurement errors and modelling error. The structure of the CT model is denoted by the triad $[m\ n\ \tau]$. The objective is then to identify a suitable model structure ($[m\ n\ \tau]$) for (6) and to estimate the parameters that characterize this structure, based on the observed sampled measurement of the input (rainfall intensity) and output signal (runoff flow rate) $i_E(t_k)$ and $Q(t_k)$ ($k = 1, \ldots, N$).

Various statistical methods of identification and estimation have been proposed to implement both direct and indirect approaches to continuous-time model identification outlined above (Garnier and Wang [2008]). We have chosen to use the Simplified Refined Instrumental Variable algorithm (SRIVC) (Young and Jakeman [1980]). This algorithm is available in both CAPTAIN (see www.es.lancs.ac.uk/ cres/captain/) and CONTSID (see www.cran.uhp-nancy.fr/contsid/) Toolboxes for Matlab. The SRIVC method can be interpreted in optimal statistical terms, so providing an estimate of the parametric error covariance matrix and, therefore, estimates of the confidence bounds on the parameter estimates when the measurement noise $\xi(t_k)$ is assumed to be white.

Model structure identification is based on two statistical criteria. First, the coefficient of determination, R_T^2, based on the error between the sampled output and the simulated CT model output se:

$$R_T^2 = 1 - \frac{\sigma_{\hat{\xi}}^2}{\sigma_Q^2} \quad (8)$$

where $\sigma_{\hat{\xi}}^2$ is the variance of the estimated noise $\hat{\xi}(t_k)$ and σ_Q^2 is the variance of the measured output $Q(t_k)$. Second, the YIC statistic, which is a measure of model identifiability and is based on how well the parameter estimates are defined statistically, (see *e.g.* Young and Garnier [2006], Garnier and Wang [2008]).

$$\text{YIC} = \log_e \frac{\hat{\sigma}^2}{\sigma_y^2} + \log_e\{\text{NEVN}\}; \quad \text{NEVN} = \frac{1}{n_\theta} \sum_{i=1}^{n_\theta} \frac{\hat{p}_{ii}}{\hat{\theta}_i^2} \quad (9)$$

Here, n_θ is the number of estimated parameters; $\hat{p}_{ii}$ is the ith diagonal element of the block-diagonal covariance matrix $\mathbf{P}_\theta$, and so is an estimate of the variance of the estimated uncertainty on the ith parameter estimate. $\hat{\theta}_i^2$ is the square of the ith parameter estimate in the $\boldsymbol{\theta}$ vector, so that ratio $\hat{p}_{ii}/\hat{\theta}_i^2$ is a normalized measure of the uncertainty on the ith parameter estimate.

Note that the closer R_T^2 is to 1, the better the accuracy of fit between the model and the measured output is. In the same manner, the more negative YIC is, the more parsimonious the model is. The best structure is thus the result of a compromise between accuracy and parsimony.

2.3 Equivalence between a first order model and a linear tank model

It can be seen that a first order model denoted by the triad [1 1 0] (10) is close to the linear tank model widely used for the modelling of the runoff routing component in urban rainfall-runoff models (11) (see Beven [2003] or Ministère de l'écologie et du développement durable (MEDD) and CERTU [2003]) :

$$\frac{dQ(t)}{dt} + a_1 Q(t) = b_0 i_E(t) \quad (10)$$

$$\frac{dQ(t)}{dt} + \frac{1}{K} Q(t) = \frac{1}{K} Q_E(t) \quad (11)$$

In the linear tank model, K denotes the lag-time, *i.e.* the lag-time between the center of gravity of the effective hyetograph and the center of gravity of the hydrograph. Note also that the term $i_E(t)$ in (10) is replaced by:

$$Q_E(t) = C_R \cdot A \cdot i_E(t) = A_E \cdot i_E(t), \quad (12)$$

where $Q_E(t)$ is the flow rate generated by the effective rainfall on the whole catchment, C_R is the runoff coefficient, and A_E is the effective area of the catchment, *i.e.* the area that contributes effectively to runoff production. Thus, the link between the physical parameters A_E and K and the estimated parameters of (10) can be stated in the form:

$$\begin{cases} a_1 = \dfrac{1}{K} \cdot 24 \cdot 60, \\ b_0 = \dfrac{A_E}{360 \cdot K}. \end{cases} \quad (13)$$

Equation (13) shows that the physical parameters of the linear tank model (K, C_R) can be automatically and optimally determined from a_1 and b_0 for a first order model structure.

2.4 Proposed iterative methodology

As we saw earlier, in some cases, the initial losses can be neglected. Here, the problem simply consists in identifying a DBM model from the total rainfall.

In other cases, an initial losses model is required to improve the results. Thus, an iterative algorithm was developed in order to estimate at the same time the user-parameters of the proposed initial losses model and those of the linear part of the continuous-time transfer function model:

(1) The algorithm begins with the identification of the mathematical structure of the DBM model from the total rainfall. The best mathematical structure, *e.g.* the best triad $[m\ n\ \tau]$, is chosen on the basis of both R_T^2 and YIC criteria,
(2) In the second step, we search for the best couple (L_{in}, d) that maximizes the coefficient of determination between the simulated and the measured flow rate.
(3) Finally, we check that the best structure $[m\ n\ \tau]$ chosen from total rainfall in step 1 is still the same from the total rainfall calculated in step 2. If not, we start again the algorithm from step 2.

Note that this iterative algorithm is simple but not optimal since the best couple (L_{in},d) is not estimated for all linear model structures. Nevertheless, due to the low influence of initial losses in urban areas, this algorithm gives good results and prevents us to test all the possible combinations.

3. AN ILLUSTRATIVE EXAMPLE

The developed methodology was tested on a part of the Kerinou catchment located in Brest, in France. An overview of the catchment is given in Figure 2.

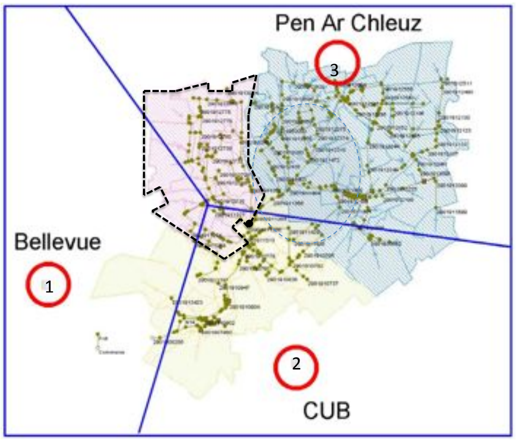

Fig. 2. Schematic representation of the urban drainage network of the Kerinou catchment. The red circles represent the location of the rain gauges and include their labels. The subcatchment is delimited by the dotted line. The bold lines delimits the Thiessen polygons.

The urban drainage network system mainly consists of combined sewers which collect both dry and wet weather wastewaters. The total area of the catchment Kerinou is about 500 ha but the area drained by the subcatchment where the measurement of flow rate was carried out is about 107 ha. A pumping station also discharges the wastewaters collected from another catchment in the main sewer. The estimated effective area of the other catchment was estimated to at least ten or so ha. The runoff coefficient of the Kerinou catchment is about 40% and was calculated on the basis of measurements of rainfall and runoff volumes. The catchment is mainly residential and characterized by steep slopes from 3 up to 10%. This indicates that lag-times may be low.

The measurements of rainfall and flow rate were carried out during 49 days with a sampling interval of respectively 1 and 5 minutes. The flow rate was assessed in the main sewer with an ultrasonic probe for the measurement of the water level and a Doppler velocimeter for the velocity. For the measurement of rainfall intensity, no rain gauge was available within the catchment. The Thiessen polygons are represented in Figure 2. We used the first half of the data for the calibration and the other part for the validation of the identified model.

Referring to the criteria R_T^2 and YIC and according to the proposed methodology, we identified some SISO models and MISO (with two rain gauges as inputs) models. Nevertheless, only first order SISO and one MISO models have real poles that can be interpreted in a physically meaningful manner. The other ones have complex or negative values that do not have any sense for an hydrological system. The estimation results obtained for the two models are summarized in Table 1.

The second column of the table indicates the rain gauge taken into account for the identification of the model. No model including 3 rain gauges could have been interpreted physically because of negative flow rate generated by one of the three processes. This shows that the influence of one of the three rain gauges (rain gauge N1 "Bellevue") is negligible. Thus, in this case, the method of the Thiessen polygons is not relevant.

Moreover, it can be observed that for both SISO and MISO models R_T^2 and YIC are very close. This suggests that taking into account 2 rain gauges only improves slightly the results. It was previously checked that the 2 signals are not linearly correlated ($r^2 = 0.23$ for the calibration dataset).

Nevertheless, a physical interpretation can be given for both models. According to (13), the effective area of the subcatchment is 67 ha. As the measured output also includes a flow rate collected from another subcatchment, the value of 67 ha is totally realistic. The values of lag-time found are also relevant.

The MISO model provides another information concerning the area of influence of each rain gauge. Referring to (13), the effective areas influenced by the rain gauges N2 and N3 are respectively 36.2 and 36.1 ha and the corresponding lag-times have values of 12 and 39 min. Those values seem also relevant. It can be observed in Figure 2 that the rainfall-runoff process occurring in the area influenced by the rain gauge N3 which is further from the outlet of the subcatchment has a larger lag-time. The measured and the simulated flow rate obtained from the MISO model are plotted in Figure 3.

Finally, the proposed methodology makes it possible to identify the two user-parameters of the model of initial losses in an automatic way.

4. CONCLUSION

This paper has outlined an approach to data-based mechanistic (DBM) modelling for urban hydrological sys-

Model type	Rain gauge	$[m\ n\ \tau]$	R_T^2	YIC	L_{in}	d	A_E	K
SISO	N2	[1 1 0]	0.85	-11.87	0.4 mm	3 h	67.4 ha	18 min
MISO	N2	[1 1 0]	0.87	-10.32	0.4 mm	3 h	36.2 ha	12 min
	N3	[1 1 0]			0.2 mm	3 h	31,1 ha	39 min

Table 1. Characteristics of the two best data-based identified continuous-time models

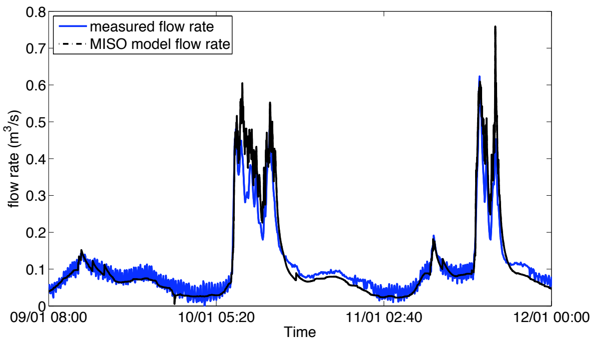

Fig. 3. Comparison of the measured and MISO model flow rate

tems based on the direct identification and estimation of continuous-time (transfer function or differential equation) models. A practical example based on the analysis of rainfall-flow data from the Kerinou catchment in Brest has been used to illustrate the relevance of the proposed modelling approach. The differences between this approach and the one carried out in most urban drainage network modelling softwares can be summarized as follows:

- No *a priori* assumption is made on the mathematical structure of the model. This is a major advantage even if the example studied showed that the best mathematical structure identified correspond to a linear tank.
- There is no need to fully describe the catchment, and break it down into sub-units because the model is directly identified at the catchment scale. Thus, the approach makes it possible to simplifiy the model.
- The model is automatically calibrated. This is also a great advantage of the developed method, because in urban drainage network modelling softwares, a lot of time is spent to choose the appropriate models and to estimate the corresponding parameters.
- The mechanistic validation of the model allows us to evaluate if the estimated parameters are relevant.
- The method directly identifies a continuous-time model (*i.e.* a linear differential equation). Thus, the estimated parameters do not depend on the sampling interval.

REFERENCES

A. Berne, G. Delrieu, J. D. Creutin, and C. Obled. Temporal and spatial resolution of rainfall measurements required for urban hydrology. *Journal of Hydrology*, 299 (3-4):166–179, 2004.

K. J. Beven. *Rainfall-Runoff Modelling: The Primer*. John Wiley & Sons, Chichester, 2003.

D. Brulé, F. Blanchet, and J. Rousselle. Study of runoff losses on impervious surfaces in an urban environment. *Revue des Sciences de l'Eau*, 10(2):147–166, 1997.

D. Butler and M. Schütze. Integrating simulation models with a view to optimal control of urban wastewater systems. *Environmental Modelling & Software*, 20(4): 415–426, April 2005.

Bernard Chocat. *Un système d'aide à la gestion, la prévision et la conception des équipements d'assainissement.* PhD thesis, INSA Lyon and University Claude Bernard of Lyon (France), 1981.

M. Desbordes. *Modélisation en hydrologie urbaine. Recherches et applications.* USTL. Laboratoire d'Hydrologie et Modélisation, ENGREF., 1994.

T. Einfalt, G. Johann, and A. Pfister. On the spatial validity of heavy point rainfall measurements. *Water Science and Technology*, 37(11):21–28, 1998.

J. Falk and K. Niemczynowicz. Characteristics of the above-ground runoff in sewered catchments. *Urban Storm Drainage*, pages 159–171, 1978.

H. Garnier and L. Wang, editors. *Identification of continuous-time models from sampled data.* Springer-Verlag, London, 2008.

V. Laurain, M. Gilson, H. Garnier, and P. Young. Refined instrumental variable methods for identification of Hammerstein continuous-time Box–Jenkins models. In *47th IEEE Conference on Decision and Control, CDC'08*, pages 210–216, Cancun (Mexico), december 2008.

Ministère de l'écologie et du développement durable (MEDD) and CERTU. La ville et son assainissement : principes, méthodes et outils pour une meilleure intégration dans le cycle de l'eau. Technical report, CERTU Editions, 2003.

G. P. Rao and H. Garnier. Numerical illustrations of the relevance of direct continuous-time model identification. In *15th Triennial IFAC World Congress on Automatic Control*, Barcelona (Spain), 2002.

N. K. Sinha and G. P. Rao, editors. *Identification of Continuous-Time Systems: Methodology and Computer Implementation.* Kluwer Academic Publishers, Norwell, USA, 1991.

P. A. Vanrolleghem, L. Benedetti, and J. Meirlaen. Modelling and real-time control of the integrated urban wastewater system. *Environmental Modelling and Software*, 20(4):427–442, 2005.

P. C. Young. Data-based mechanistic modelling of environmental, ecological, economic and engineering systems. *Environmental Modelling and Software*, 13(2):105–122, April 1998.

P. C. Young and H. Garnier. Identification and estimation of continuous-time, data-based mechanistic (dbm) models for environmental systems. *Environmental Modelling & Software*, 21(8):1055–1072, August 2006. doi: 10.1016/j.envsoft.2005.05.007.

P. C. Young and A. Jakeman. Refined instrumental variable methods of recursive time-series analysis - part III. extensions. *International Journal of Control*, 31(4): 741–764, 1980.