



# Managing uncertainties in urban runoff quality models: A benchmarking methodology

Assem Kanso, Bruno Tassin, Ghassan Chebbo

## ► To cite this version:

Assem Kanso, Bruno Tassin, Ghassan Chebbo. Managing uncertainties in urban runoff quality models: A benchmarking methodology. Novatech 2004, Jun 2004, Lyon, France. pp.399-'06. hal-00727149

**HAL Id: hal-00727149**

**<https://enpc.hal.science/hal-00727149>**

Submitted on 2 Sep 2012

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

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

## Managing uncertainties in urban runoff quality models: A benchmarking methodology

Un banc d'essai pour l'évaluation des incertitudes dans les modèles de pollution des eaux de ruissellement

Kanso A.\*<sup>1</sup>, Tassin B.\*, Chebbo G.\*. \*\*

(\*) CERVE - ENPC

6, 8 avenue Blaise Pascal, Champs sur Marne  
77455 Marne la Vallée, cedex 02, France

(\*\*) Faculty of Engineering – Lebanese University

Airport street, Beirut – Lebanon

(<sup>1</sup>) kanso@cerve.enpc.fr

### RESUME

Nous présentons une méthodologie d'évaluation des incertitudes dans les modèles de pollution des eaux de ruissellement. Cette méthodologie, basée sur la théorie bayésienne, utilise une technique de simulation par chaîne de Markov (l'algorithme Metropolis), pour estimer la distribution de probabilité a posteriori des paramètres du modèle, ce qui permet, d'une part, une évaluation quantitative des incertitudes liées à l'estimation des paramètres ainsi que leurs interactions, et d'autre part, fournit des informations sur la sensibilité des réponses du modèle aux paramètres. L'utilité et l'efficacité de cette méthodologie sont illustrées dans le contexte de 4 configurations de modèles d'accumulation / érosion testés sur 4 chaussées. Les résultats de calage démontrent que l'algorithme Metropolis produit des inférences fiables des paramètres qui aident à l'amélioration du concept mathématique des modèles.

### ABSTRACT

In this paper we present a benchmarking methodology, which aims at comparing urban runoff quality models, based on the Bayesian theory. After choosing the different configurations of models to be tested, this methodology uses the Metropolis algorithm, a general MCMC sampling method, to estimate the posterior distributions of the models' parameters. The analysis of these posterior distributions allows a quantitative assessment of the parameters' uncertainties and their interaction structure, and provides information about the sensitivity of the probability distribution of the model output to parameters. The effectiveness and efficiency of this methodology are illustrated in the context of 4 configurations of pollutants' accumulation/erosion models, tested on 4 street subcatchments. Calibration results demonstrate that the Metropolis algorithm produces reliable inferences of parameters thus, helping on the improvement of the mathematical concept of model equations.

### KEYWORDS

Bayesian inference, Conceptual model, Parameter uncertainty, Sensitivity analysis, Urban runoff pollution

## 1. INTRODUCTION

It looks impossible to mechanistically describe the processes involved in pollutants' runoff in urban drainage systems, due to the complexity of the processes related to the generation and transport of pollutants during rainfall, the heterogeneity of the system's characteristics and their related space and time scales. Therefore, the only possible modelling approach to be used in this case is the conceptual one, which often contains parameters that do not have a direct physical interpretation and therefore, cannot be measured in the field. Instead, these parameters must be estimated using a calibration procedure whereby the model parameters are adjusted until the system output and the model output show an acceptable level of conformity.

Automatic calibration methods quantitatively express the distance between simulation results and measured data in terms of a criteria measure and use an optimization algorithm to optimize (minimize or maximize) this measure.

However, the conceptual nature of stormwater quality models and the uncertainty level of the *in situ* measurement data [Ahyerre et al., 1998; Ashley et al., 1999], rarely allow a satisfactory calibration and validation of the model, thus the estimated parameters from these models are generally error-prone.

Furthermore, one of the great limitations of classical optimization algorithms for calibration is that they do not allow neither an estimation of the significance of the obtained optimal parameter set, nor a realistic quantification of the predictive uncertainty. Hence, the existing urban stormwater quality models are rarely used for practical application. In order to improve these models and their usefulness, we propose to develop a more robust methodology for calibration and validation of models.

In the last decade, great attention has been given to the Bayesian approach for model calibration particularly in the case of complex hydrological models [Kuczera et al., 1998, Campbell et al., 1999], but rarely in environmental modelling. Based on this approach, a "Monte Carlo Markov Chain method MCMC", was proposed by [Kanso et al., 2003] and applied to urban pollutants' stormwater modelling. Primary results have shown the robustness and effectiveness of this method. Unlike traditional calibration techniques, this method not only attempts to identify a "best parameter set", but also helps to assess, and where possible to reduce, uncertainties in the parameter values.

This paper describes a benchmarking methodology based on this method to test the existing urban stormwater quality models on different scales and for various parts of the urban catchment system (roof surface, paved surface, street surface, sewers and the entire catchment). The availability of data resulting from a 2-year survey conducted on the "Marais" catchment in the centre of Paris [Gromaire, 1998] facilitates the implementation of this methodology, and leads to an estimation of the system's sensitivity to its different components, a better understanding of the processes involved, and a reduction of uncertainties in these models.

As a first step, this benchmarking methodology as described below, will be tested in this paper in the context of pollutants' accumulation/erosion models at street surface areas, and will be extended later to other pollutants' stormwater submodels.

This paper is organized as follows: Section 2 presents the methodology proposed by introducing the Bayesian inference, the Metropolis algorithm and their use to analyse the models' uncertainties and their sensitivities to the parameters. Section 3 illustrates the usefulness and the applicability of the benchmarking methodology in the case of 4 configurations of pollutants' accumulation/erosion models, tested on 4 street subcatchments. Finally; section 4 summarizes the methodology and discusses the results.

## 2. THE BENCHMARKING METHODOLOGY

As shown in Figure 1, the benchmarking methodology consists of choosing different configurations of models to be analysed on a given system using the available *in situ* measurement data. Firstly, the uncertainties in the models' parameters are inferred using a MCMC sampling method based on the Bayesian approach and secondly the inference results are analysed and compared for the different models in order to better understand the processes involved, estimate the system's sensitivity to its different components and reduce uncertainties in these models.

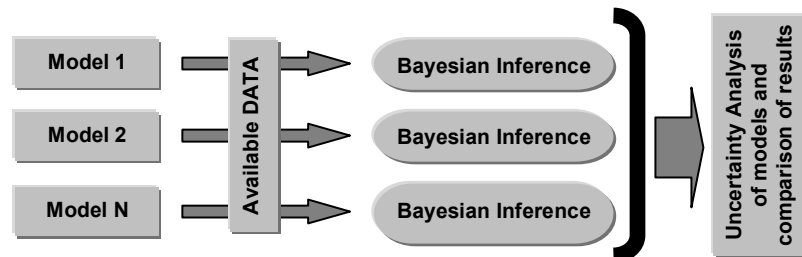


Figure 1. Illustration of the benchmarking methodology proposed

### 2.1. Bayesian inference

#### 2.1.1. Concept

We are interested in mathematical models that predict outputs from inputs. The models are indexed by parameters, which may (or may not) be physically interpretable. The model  $f(\cdot)$  can be cast as a nonlinear regression model:

$$Y_t = f(X_t, \theta) + \varepsilon_t, \quad t = 1, \dots, n \quad (1)$$

Where  $(Y_1, \dots, Y_n)$  is a vector of model predictions,  $(X_1, \dots, X_n)$  is a vector of input data,  $\theta = (\theta_1, \dots, \theta_p)$  is the vector of  $p$  unknown parameters and  $(\varepsilon_1, \dots, \varepsilon_n)$  is a vector of statistically independent errors with a mean zero and a variance  $\sigma^2$ .

Equation (1) contains various sources of uncertainties: (i) measurement uncertainties representing randomness in samples; (ii) uncertainties caused by lack of knowledge of all the causes and effects in physical systems, or by lack of sufficient data. This category of uncertainty can be subdivided in two types of uncertainties: those related to parameters and those related to the model itself. We are interested in this paper in the model's parameters uncertainties.

Assuming that the mathematical structure of the model is predetermined and fixed, the aim of model calibration procedure is to reduce the uncertainty in the parameter values by assimilating measurement data.

Bayesian approach, expresses uncertainties in terms of probability. Uncertainty in measurements can be formalized in a familiar way by assuming that the residuals are drawn from a suitable probability distribution. However, parameter uncertainty is quantified by introducing at first a prior probability distribution  $P(\theta)$  which represents the knowledge about  $\theta$  before collecting any new data, and secondly, by updating this prior probability on  $\theta$  to account for the new data collected ( $D$ ). This update is performed using Bayes' theorem, which can be expressed as:

$$P(\theta|D) = P(D|\theta) \cdot P(\theta) / P(D) \quad (2)$$

Where  $P(\theta|D)$  is the posterior distribution of  $\theta$ ,  $P(D)$  is a proportionality constant required so that  $\int P(\theta|D) d\theta = 1$ , and  $P(D|\theta)$  is the conditional probability for the

measured data given the parameters.  $P(D|\theta)$  is similar to the likelihood function of the model. In this case, the likelihood function can be written in the multiplicative form:

$$P(D|\theta) = \prod_{i=1}^n \frac{1}{(2 \cdot \pi \cdot \sigma^2)^{1/2}} \cdot e^{-\frac{(Y_i - f(X_i, \theta))^2}{2 \cdot \sigma^2}} \quad (3)$$

Note that Bayes' theorem does not allow one to derive posterior distribution without prior knowledge. However, in order to avoid favouring any initial value, the use of a uniform prior distribution over the range of parameters may seem reasonable [Beven et al., 1992]. The posterior distribution  $P(\theta/D)$  contains all the available information about the parameters  $\theta$ . In this case, the Bayesian statistical inference becomes an estimate of a posterior distribution of  $\theta$ .

### 2.1.2. The Metropolis algorithm for assessing parameter uncertainty

In practice, it is difficult, if not impossible, to estimate the posterior distribution  $P(\theta/D)$  by direct analytical calculation. In addition, classical approximations of  $P(\theta/D)$  by a multinormal distribution can be quite poor [Duan et al., 1992; Kuczera et al., 1998].

The Metropolis algorithm, derived from the MCMC family of techniques, was chosen because of the simplicity of its implementation, efficiency and generality. It is based on the idea that we can create a random walk in the space of parameters in a way to generate enough samples, which adapt to the true posterior distribution of parameters  $P(\theta/D)$  [Robert et al., 1999; Tanner, 1996]. There is typically an initial unstable transient phase before reaching the limit distribution. After removing this initial burn-in, the remainder is used as a dependent sample from the posterior distribution.

### 2.1.3. Monte Carlo method for assessing predictive uncertainty

The propagation of the parameters' uncertainty through the model with a Monte Carlo procedure in order to obtain its confidence intervals, gives an indication of both the predictive power of the model and its capacity to reproduce the system's processes.

## 2.2. Results analysis

The posterior distribution allows a quantitative estimation of both the parameters' uncertainties and the interaction between parameters. Moreover, a simplified sensitivity analysis approach given by a visual analysis of the scatter plots of the likelihood measure vs. each parameter can provide information about the sensitivity of the probability distribution of the model output to parameters: identification of input factors driving the model to have a good conformity with the data would be possible.

## 3. APPLICATION OF THE METHODOLOGY

### 3.1. A runoff modelling benchmark

Runoff models describe both the particulate pollutant's erosion during the storm event and their accumulation on the watershed during the preceding dry weather period. Table 1 represents the 4 configurations used to describe these processes. We have chosen 4 types of accumulation models and 2 types of erosion models (Table 2).

| Configuration            |        | Accumulation | Erosion     | Parameters nb |
|--------------------------|--------|--------------|-------------|---------------|
| Continuous Simulations   | SIM_01 | ACCU_TYPE_01 | ERO_TYPE_01 | 4             |
|                          | SIM_02 | ACCU_TYPE_02 | ERO_TYPE_01 | 4             |
| Single event simulations | SIM_03 | ACCU_TYPE_03 | ERO_TYPE_01 | 3             |
|                          | SIM_04 | ACCU_TYPE_04 | ERO_TYPE_02 | 2             |

Table 1 Simulations' configurations tested in the benchmark

ACCU\_TYPE\_01 [Alley et al., 1981] is commonly used in all available software. ACCU\_TYPE\_02 represents a mathematical reformulation of ACCU\_TYPE\_01 and has been proposed here in regards to the results obtained in this benchmark. Primary results encouraged us to propose 2 other models: i) ACCU\_TYPE\_03, for which the accumulation process is supposed to be instantaneous i.e. there is always a sufficient available mass  $M_{accu}$  regardless of the length of dry weather period; and ii) ACCU\_TYPE\_04 for which it is supposed that there is an infinite stock on the surface. ERO\_TYPE\_01, commonly used in literature [Huber et al., 1981] supposes that the eroded mass depends on the available mass and the rainfall rate. ERO\_TYPE\_02 supposes that erosion depends only on the rainfall rate.

| <b>Accumulation Process</b> |                          | <b>Mathematical formulation</b>                                    |
|-----------------------------|--------------------------|--|
| ACCU_TYPE_01                | Asymptotic               | $\frac{dMa(t)}{dt} = Daccu \cdot Simp - Dero \cdot Ma(t)$ (4)      |
| ACCU_TYPE_02                | Asymptotic               | $\frac{dMa(t)}{dt} = Kaccu \cdot (M_{lim} \cdot Simp - Ma(t))$ (5) |
| ACCU_TYPE_03                | Instantaneously          | $Ma(t) = M_{accu} \cdot Simp$ (6)                                  |
| ACCU_TYPE_04                | Infinite stock           | No accumulation  |
| <b>Erosion Process</b>      |                          | <b>Mathematical formulation</b>                                    |
| ERO_TYPE_01                 | Depends on runoff and Ma | $\frac{dMa(t)}{dt} = -Wero \cdot I(t)^w \cdot Ma(t)$ (7)           |
| ERO_TYPE_02                 | Only runoff              | $Me(t) = Kero \cdot I(t)^w \cdot Simp$ (8)                         |

Table 2 Models used to describe the dry and wet weather processes in the benchmark

Where  $Ma(t)$  (kg) is the available pollutants' mass,  $Me(t)$  (kg/h) is the pollutants' mass eroded by runoff,  $S_{imp}$  (ha) is the impervious area, and  $I(t)$  (mm/hr) is the rainfall intensity.  $Daccu$  (kg/ha/day),  $Dero$  (day<sup>-1</sup>),  $Mlim$  (kg/ha),  $Kaccu$  (day<sup>-1</sup>),  $M_{accu}$  (kg/ha),  $Wero$ ,  $Kero$  and  $W$  are the calibration parameters.  $Mlim$  represents the maximum accumulated mass and is equivalent to  $Daccu / Dero$ .

### 3.2. The Data

Table 3 presents the characteristics of the 4 street subcatchments used for the application. These data, for which 51 suspended solids pollutographs are available, were acquired for the urban catchment "le Marais" over the period 1996-1997.

| <b>Street name</b>             | <b>Duval</b> | <b>M.B.M.</b> | <b>St. Antoine</b> | <b>Turenne</b> |
|--------------------------------|--------------|---------------|--------------------|----------------|
| Surface (m <sup>2</sup> )      | 160          | 195           | 1017               | 1700           |
| Commercial /circulation level  | low/low      | no/low        | high/high          | low/med        |
| Pollutographs: all/calibration | 11/8         | 17/12         | 11/8               | 12/9           |

Table 3 Characteristics of the 4 street subcatchments (same slope and same imperviousness)

### 3.3. Results and discussion

A uniform distribution is assumed to encode the prior knowledge about the parameters. For each simulation, 12000 iterations were performed with the Metropolis algorithm, and the first 2000 samples generated were discarded.

Results showed that the Metropolis algorithm converged successfully to the same posterior probability distribution of the parameters regardless of the initial parameter set used. Figure 2 shows the obtained posterior distribution of parameters for the configuration SIM\_01 estimated at the Duval Street.

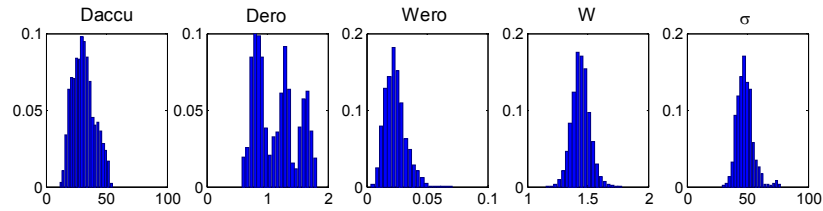


Figure 2 Posterior distribution for the SIM\_01's 4 parameters estimated at the Duval street

Contrary to the wet weather parameters *Wero* and *W*, results for SIM\_01 show a small reduction of uncertainty for the dry weather parameters, *Daccu* and *Dero*, comparing to the uniform prior distribution. By comparing the variation of the likelihood measure vs. the different parameters, one may conclude that the dry weather parameters *Daccu* and *Dero* used in SIM\_01 have no significant effect on the model's response (Figure 3.a.). Moreover, Figure 4.a. shows a strong linear correlation between these 2 parameters. We can conclude that no one of these two parameters drives the model to be more "behavioural", and that their interaction may have the main effect on the model's behaviour. The model used in SIM\_01 would probably be more easily calibrated if mathematically reformulated.

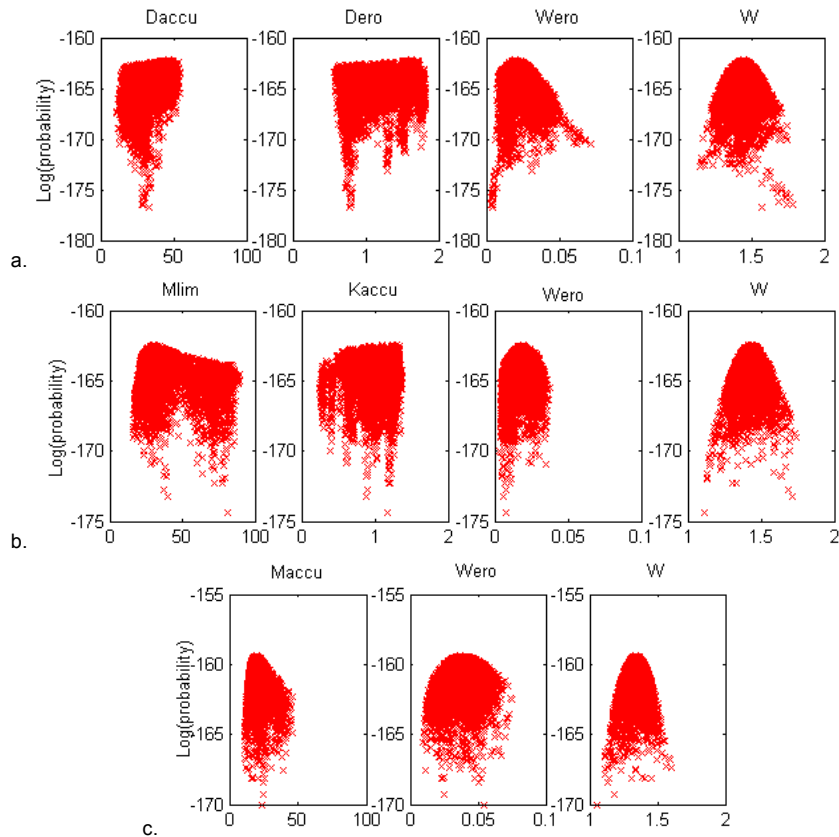


Figure 3 Scatter plots of the likelihood measure vs. the parameters for the configurations a. SIM\_01, b. SIM\_02 and c. SIM\_03 obtained at the Duval street

However, results for the reformulated model SIM\_02 show a better estimation of the *Mlim* parameter and scatter plots indicate that a clear optimum is detected for this parameter (Figure 3.b.). But great uncertainty is still obtained for the accumulation factor *Kaccu*, with a clear trend for a highest value ( $>1 \text{ day}^{-1}$ ) which is much more important than the values used in literature ( $=0.08$ ) [Bujon et al., 1990]. Such result suggests (confirming the results shown by [Gromaire, 1998]) that the dry weather period have no significant effect on the accumulation process. This led us to test two hypotheses: an instantaneous accumulation process and an infinite stock.

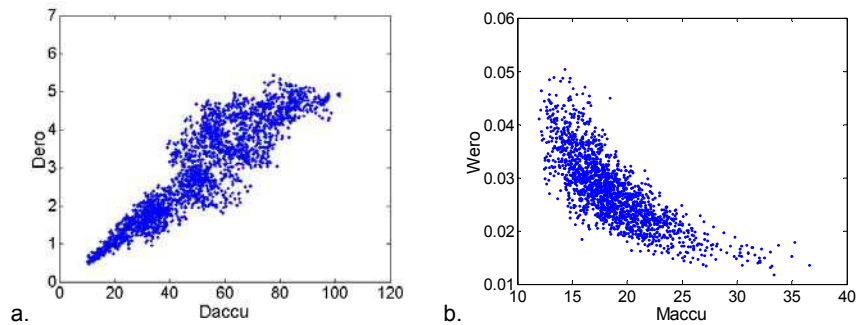


Figure 4 Types of correlations found between the model's parameters for the Turenne street: a. Dry weather parameters Daccu and Dero (SIM\_01); b. Mass of pollutants available and the erosion parameter Wero (SIM\_03).

Results obtained for SIM\_03 indicate a good estimation of a unimodal distribution of parameters. Figure 3.c. also shows that the likelihood measure is highly sensitive to the accumulated mass. Nevertheless, results indicate a clear correlation between the mass available *Maccu* and the erosion parameter *Wero* (Figure 4.b.). Such correlation is not surprising regarding the mathematical structure of the erosion model (Equation (7) represents a multiplicative form of *Maccu* and *Wero*). However, despite this correlation, results show a slight diminution of the variance of errors  $\sigma^2$  and consequently amelioration in the predictive power of the model (Table 4).

| Configuration           | Measurement | SIM_01      | SIM_02      | SIM_03     | SIM_04      |
|-------------------------|-------------|-------------|-------------|------------|-------------|
| $\sigma \text{ (mg/l)}$ | 62          | $47 \pm 10$ | $47 \pm 10$ | $42 \pm 7$ | $50 \pm 10$ |

Table 4 Comparison of the standard deviation of the measured concentrations with the standard deviation of residuals between model and observation for the 4 configurations at the Duval Street. It is to be noted that results for SIM\_04 showed deterioration in the model behaviour (Table 4). In contrast to the other three configurations, this model cannot explain the washing-effect (not showed here) of the street after a high rainfall rate that has been shown by [Gromaire, 1998].

The propagation of the parameters' uncertainty remaining after calibration of the model using a Monte Carlo procedure shows that the model confidence intervals are large, or, in other words, that the predictive power of the calibrated model is low.

#### 4. CONCLUSION

This paper has presented a benchmarking methodology, based on the Bayesian theory, for assessing parameter uncertainties in urban runoff quality models. This methodology consists firstly of choosing a number of models' configurations to describe the system studied and uses the Metropolis algorithm, a general MCMC sampling method, to infer the true posterior probability distribution of the models' parameters conditioning to data. An analysis of the obtained posterior distributions for



the different configurations allows a quantitative assessment of the parameters' uncertainties and their interaction structure, and helps to identify the main parameters that drive the model to have a good conformity with the data. The usefulness and the applicability of this methodology are illustrated in the case of 4 configurations of pollutants' accumulation/erosion models, tested on 4 street subcatchments.

Though Metropolis algorithm is computationally very intensive and needs a considerable number of iterations, results demonstrate clearly its effectiveness and efficiency to produce reliable inferences of parameters.

Results show that the mathematical concept of the accumulation model, using two parameters *Daccu* and *Dero*, contains a strong interaction between its parameters, and implies much more uncertainty in their calibration.

Furthermore, despite that a reformulation of this model using two parameters (*Mlim* and *Kaccu*) allows a better identification of the parameter *Mlim*, it seems difficult to reduce uncertainty about the accumulation parameter *Kaccu*. The high estimated value of this parameter and the good behaviour obtained of a one parameter accumulation model (*Maccu*) suggest that accumulation may happen instantaneously regardless of the length of the dry weather period. This hypothesis casts doubts on the utility of using an asymptotic behaviour to describe the accumulation process. Such a conclusion needs to be validated on other sites to test its generality.

Finally, this method delivers much information which would have been unreachable with classical calibration methods and which are very useful for modelling attempts.

## ACKNOWLEDGMENTS

Authors gratefully acknowledge the financial support of the "Réseau Génie Civil et Urbain" RGC&U and the "Syndicat Interdépartemental pour l'Assainissement de l'Agglomération Parisienne" SIAAP.

## BIBLIOGRAPHIE

- Ahyerre M., Chebbo G., Tassin B., & Gaume E. (1998). Storm water quality modeling, an ambitious objective?. *Wat. Sci. Tech.*, vol. 37, No. 1, pp. 205 - 213.
- Alley W. & Smith P. (1981). Estimation of accumulation parameters for urban runoff quality modelling. *Wat. Res. Res.*, vol. 17, No. 6, pp. 1657 – 1664.
- Ashley. R.M., Hvitved-Jacobsen T., Bertrand-Krajewski J.L (1999). Quo vadis sewer process modelling ? *Wat. Sci. Tech.*, vol. 29, No. 9, pp.9-22.
- Beven K.J. & Binley A.M. (1992). The future of distributed models: model calibration and predictive uncertainty. *Hydrol. Processes*, 6, pp. 279 – 298.
- Bujon G. and HERREMANS L. (1990). FLUPOL, modèle de prévision des débits et des flux polluants en réseaux d'assainissement par temps de pluie, calage et validation. *La Houille Blanche*, No 2, pp. 123-139.
- Campbell E. & Fox P. (1999). A bayesian approach to parameter estimation and pooling in nonlinear flood event models. *Wat. Res. Res.*, vol. 35, No. 1, pp. 211 - 220.
- Duan Q., Sorooshian S., & Gupta V. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Wat. Res. Res.*, vol. 28, No. 4, pp. 1015 - 1031.
- Gromaire M.C., (1998). *La pollution des eaux pluviales urbaines en réseau d'assainissement unitaire, caractéristiques et origines*. PhD thesis, ENPC – Marne-la-Vallée.
- Huber W., Heaney J., Nix S., Dickinson R. & Polmann D. (1981). *SWMM3 user's Manual*. Technical report, EPA.
- Kanso A., Gromaire M.C., Gaume E., Tassin B. and Chebbo G. (2003). Bayesian approach for the calibration of models: Application to an urban stormwater pollution model. *Wat. Sci. & Tech.*, vol. 47, no. 4, pp. 77– 84.
- Kuczera G. & Parent E. (1998). Monte Carlo assessment of parameter inference in catchments models: The Metropolis algorithm. *Journal of Hydrology*, 211, pp. 69 – 85.
- Robert C.P. & Casella G. (1999). *Monte Carlo statistical methods*. Springer Press.
- Tanner M. (1996). *Tools for statistical inference: Methods for the exploration of posterior distributions and likelihood functions*. Springer Press.