Jérémie Sage, Céline Bonhomme, Emmanuel Berthier, Marie-Christine Gromaire

# Modelling stormwater management strategies – Effect of uncertainties in pollutant wash-off dynamics

Jérémie Sage[1], Céline Bonhomme[1], Emmanuel Berthier[2] and Marie-Christine Gromaire[1]

[1]Université Paris Est, LEESU, UMR MA 102 - Agro ParisTech, Ecole des Ponts ParisTech, Cité Descartes - 6-8 av. Blaise Pascal, 77455, Champs-sur-Marne Cedex 2, France (Email: *jeremie.sage@leesu.enpc.fr, celine.bonhomme@leesu.enpc.fr, marie-christine.gromaire@leesu.enpc.fr*)
[2]CEREMA (Centre d'étude et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement), DT Ile-de-France, 12 rue Teisserenc de Bort, 78180, Trappes, France (Email: *emmanuel.berthier@cerema.fr)*

**Abstract**
This study investigates the effect uncertainties associated with pollutant wash-off dynamics in the context stormwater management practices modelling. A formal Bayesian approach is adopted for the calibration and the uncertainty analysis of a commonly used wash-off model, under (1) the unverified assumption of homoscedastic, independent and normally distributed residuals and (2) using a more correct heteroscedastic and autoregressive error model. The results obtained for each of these approaches are compared and the uncertainty associated with water quality modelling is later propagated through a conceptual Best Management Practices (BMP) model, for various stormwater management scenarios, so as to assess the effect of this uncertainty for BMP modelling and carify the benefits of a robust description of error structure. This study indicates that violation of the statistical asumptions about the residuals may result in unreliable estimation of model parameters and total predictive uncertainty. The effect of the uncertainty associated with the intra-event variability of concentations in runoff is however found to have only a limited effect on the outputs of the BMP model, regadless of the error model adopted for calibration.

**Keywords**
Best Management Practices; Calibration; Propagation; Source-control; Suspended solids;

## INTRODUCTION

In recent years, on-site runoff pollution control in small vegetated systems (referred to as Sustainable Urban Drainage Systems, Best Management Practices or Green Infrastructure) has been shown to be a relevant option to limit the adverse effects of stormwater discharge to receiving waters (Ahiablame et al., 2012). Recent literature results suggest that the performance of such source-control systems is mostly related to the volume reduction induced by infiltration and evaporation (Bressy et al., 2014) and hydrological modelling therefore offers opportunities for the development of efficient stormwater control strategies. Nonetheless, because pollutant concentrations in runoff exhibit large temporal variations during rain events (Kanso et al., 2006; Schriewer et al., 2008; Shaw et al., 2010; Vezzaro and Mikkelsen, 2012), volume reduction generally differs from pollutant load reduction and it is yet unclear whether accounting for this temporal variability is needed to assess the performance of Best Management Practices (BMP).

Various water quality models have been introduced in the past to simulate pollutant concentrations in runoff. Correct replication of the wash-off process on urban surfaces (which dictates the variability of concentrations) however remains a challenge and several findings indicate that the performance of these models has generally been overestimated and that simulated concentrations are thus subject to very large uncertainties (Dotto et al., 2010; Freni et al., 2009; Kanso et al., 2006; Sage et al., 2015; Vezzaro and Mikkelsen, 2012). In the context of stormwater management practices modelling, assessing the effect of these uncertainties on the performance of source control devices requires propagation of errors through BMP models. Unfortunately, it is yet difficult to obtain reasonable estimation of uncertainty from conventional formal Bayesian methods which have often relied on unverified statistical assumptions (Dotto et al., 2013; Evin et al., 2013), although extensions of these techniques have recently shown promising results for more rigorous bias description (Del Giudice et al., 2013; Schoups and Vrugt, 2010; Yang et al., 2007).

The purpose of this study is therefore to (1) to adequately evaluate the uncertainty in the intra-event variability of concentrations simulated from the widely used exponential wash-off model and (2) to assess the effect of this uncertainty for BMP modelling. The water quality model is calibrated from continuous turbidity and flow-rates measurements from an urban street over an 11 month period, using an autoregressive AR(1) error model to account for the autocorrelation and the non-normality of the residuals. So as to illustrate the benefits of this approach, calibration results are first compared to those obtained under the standard hypothesis of independent, homoscedastic and normally distributed residuals (standard error model). After a short discussion on calibration results, TSS concentrations simulated for 1-year rainfall period and corresponding runoff rates are used as an input to a conceptual BMP model. Parameter and predictive uncertainty associated pollutant wash-off dynamics are thus propagated through the BMP model to evaluate the effect of uncertainty under various BMP design for both the improved and standard error models.

## MATERIAL AND METHODS

### Site and data description

The experimental site consists of a small road catchment (~800m²) carrying moderate traffic loads (~8000 vehicles per day) located in "Sucy-en-Brie" municipality, a residential district nearby Paris, France. Flow-rates and turbidity measurement were recorded over an 11 month period at a 1-min time-step from a tipping bucket flow-meter system and a multi-parameter-probe installed in a storm drain. Turbidity time-series were converted to total suspended solids concentrations (TSS) using a linear TSS-turbidity relationship adjusted from event mean runoff samples collected for 7 rain events (see Sage et al., 2015 for further details on the experimental setting and the dataset).

A total of 175 rain events (considering a 30 minutes minimum inter-event time for their identification) were fully monitored from January 2013 to November 2013 and are thus used in this study for water quality model calibration and uncertainty analysis.

### Water quality modelling

The temporal variability of pollutant concentrations in runoff (from an event to another and during a storm) has traditionally been assumed to result from dry weather accumulation of pollutant on urban surfaces, followed by their removal during rain events and many conceptual water quality model thus seek to replicate these two processes from relatively simple equations (several examples can be found in Freni et al., 2009). Recent findings however suggest that the reliability of such models should be questioned and concerns have more specifically been raised about the validity of accumulation functions which relate the amount of pollutant available at the beginning of a rain event to antecedent dry period duration (Kanso et al., 2006; Sage et al., 2015; Shaw et al., 2010).

The inability of conventional water quality models to simulate pollutant accumulation for the dataset used in this study was discussed in Sage et al. (2015). In this paper, accumulation is hence deliberately left aside so as to focus on the intra-event variability of TSS concentrations and the uncertainty associated with a widely used exponential wash-off model, which may be written as:

$$C_t = \frac{M(t) \times \left[1 - \exp\left(-C_1 \times q_t^{C_2} \times \Delta t\right)\right]}{q_t \times \Delta t} \tag{1}$$

Where: $C_t$ = TSS concentration in runoff from t to t+Δt, $q_t$ = flow rate recorded at the outlet of the catchment from t to t+Δt, M(t) = sediment load available on road surface at t [$g.m^{-2}$], Δt = computation time-step (= 5 min), $C_1$, $C_2$ = wash-off model parameters. When runoff occurs ($q_t > 0$), the sediment storage is updated at each time-step from the suspended solid load washed-off at the previous time step ($M(t+1) = M(t) - C_t \times q_t \times \Delta t$). The model hence simply requires M(t) to be specified at the beginning of each rain event for $t = t_{0,i}$ (corresponding procedure is discussed in the next section).

**Model calibration and uncertainty analysis**

*Bayesian inference and MCMC sampling.* Formal Bayesian techniques have often been successfully applied in hydrological modelling and clearly offer opportunities for a robust assessment of parameter and predictive uncertainty (Bates and Campbell, 2001; Del Giudice et al., 2013; Li et al., 2011; Schoups and Vrugt, 2010; Yang et al., 2007). In this study, a Monte-Carlo Markov Chain (MCMC) sampling method based on the Metropolis-Hasting (1970) algorithm (M-H) is adopted for calibration and uncertainty analysis of the exponential wash-off model.

Under the formal Bayesian approach, model's outcome for a set of parameter θ is expressed as a probability density function of model parameters P(θ|D) that can be derived from prior knowledge about model parameters P(θ) updated by observations D. Assuming non-informative (e.g uniform) prior P(θ), the posterior probability density function of model parameters P(θ|D) can be shown to be proportional to the likelihood function L(D|θ) which measures the probability of simulation errors and reflects the structure of the residuals between observation and model outputs. Once the likelihood function specified, posterior parameter distribution P(θ|D) can be estimated numerically from the M-H algorithm which generates a random walk through the space of parameters that converges to the posterior probability function P(θ|D) (Chib and Greenberg, 1995).

*Error model formulation.* The specification of a likelihood function directly relates to the selection of a statistical error model to describe the residuals $e_t$ between model outputs and observations (Schoups and Vrugt, 2010). In many applications, residuals have been assumed to be to be independent, homoscedastic and normally distributed $e \sim N(0, \sigma_e)$. Under such hypothesis P(θ|D) can be computed from:

$$P(\theta | D) \propto L(\theta | D) = \prod_{i=1}^{N} \left[ \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{e_t^2}{2\sigma_e^2}\right) \right] \tag{2}$$

Unfortunately, such assumptions are generally unrealistic in both natural and urban hydrology, especially when dealing with high frequency flow rates or water quality measurements (Del Giudice et al., 2013; Sage et al., 2015), and recent results suggest that strong violation of these statistical hypotheses may lead to erroneous estimation of parameter and prediction uncertainty (Dotto et al., 2013; Evin et al., 2013; Schoups and Vrugt, 2010; Thyer et al., 2009). In this study, an non-normal autoregressive AR(1) error model and log-sinh variance stabilization technique (Del Giudice et al., 2013; Wang et al., 2012) are thus introduced for a more realistic bias description:

$$e_t = g\left(y_{sim,t}\right) - g\left(y_{obs,t}\right) = \rho \times e_{t-1} + \varepsilon_t \tag{3}$$

$$g\left(y_{sim}\right) = \frac{1}{b}\log\left[\sinh\left(a + by\right)\right] \tag{4}$$

Where: $y_{obs,t}$ = observations, $y_{sim,t}$ = model outputs, $e_t$ = residuals in the transformed space, $\rho$ = autocorrelation coefficient, $\varepsilon_t$ = stochastic innovations, a and b = log-sinh transformation parameters. Assuming that innovations $\varepsilon_t$ follow a Student-t distribution with standard deviation $\sigma$ and $\nu$ degrees of freedom, the likelihood function becomes (Yang et al., 2007):

$$L(\theta|D) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{\sqrt{1-\rho^2}}{\sigma\sqrt{\pi(\nu-2)}} \left[1 + \frac{e_{t_1}{}^2\left(1-\rho^2\right)}{(\nu-2)\sigma^2}\right]^{-\frac{\nu+1}{2}} \left|\frac{dg}{dy}\right|_{y_{t_0}} \prod_{i=2}^{n} \left[\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sigma\sqrt{\pi(\nu-2)}} \left(1 + \frac{\varepsilon_{t_i}{}^2}{(\nu-2)\sigma^2}\right)^{-\frac{\nu+1}{2}} \left|\frac{dg}{dy}\right|_{y_{t_i}}\right] \tag{5}$$

Where: N = number of observations, $\varepsilon_{t,i}$ = innovations at the $i^{th}$ time-step, |dg/dy| = derivative of the log-sinh function, $\Gamma$ = Gamma function.

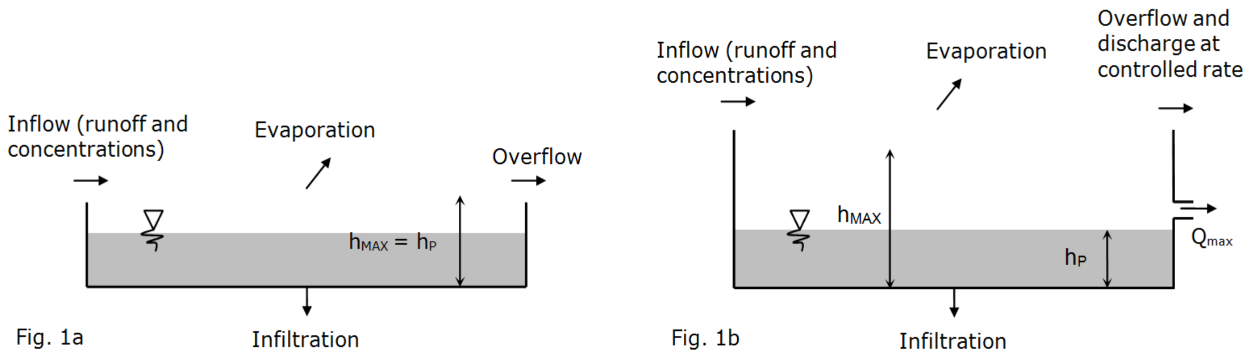
*Uncertainty analysis implementation.* So as to evaluate the benefits of a statistically correct bias description, wash-off model calibration is conducted for both the “standard” and “improved” (e.g. autoregressive, heteroscedastic and non-normal) error models with corresponding likelihood functions (Eq. 2 and 5). M-H algorithm is run for 100.000 iterations from a previously identified maximum likelihood estimate and jump probability is automatically adjusted to approximately achieve a 23% acceptance rate (see Roberts et al., 1997). Because pollutant accumulation over the surface of the road is not represented in this study, initial pollutant load $M(t_{0,i})$ is adjusted for each rain event from a simple least square optimization at each iteration of the M-H algorithm, and the calibration procedure thus solely investigates the uncertainty associated with the intra-event variability of pollutant concentrations and wash-off model parameters. Fitted $M(t_{0,i})$ values are here assumed to range from 0 to 20 g/m², in accordance with previous literature results (Deletic and Orr, 2005; Vaze and Chiew, 2002; Zhao et al., 2011).

Posterior probability distribution for wash-off model parameters $C_1$ and $C_2$ as well autocorrelation coefficient and log-sinh transformation parameters are estimated jointly within the Bayesian framework for uniform prior distribution $P(\theta)$. For the second likelihood function, optimal value for the degrees of freedom $\nu$ of the Student t-distribution is determined after performing calibration for different values of $\nu$, as suggested by Yang et al. (2007). Parameter uncertainty and stochastic errors are finally propagated by running the model for 5.000 set of parameters sampled from the estimated posterior distribution $P(\theta|D)$ to generate confidence intervals for both the “standard” and “improved” error models. For each sample, a random error term (Gaussian white noise or auto-correlated bias with Student-t innovations) is hence added to model outputs compute total predictive uncertainty. (A detailed methodology for the calculation of confidence intervals can be found in Li et al., 2011)

**BMP modelling and propagation of uncertainty**

*Stormwater management modelling.* A conceptual BMP model is adopted to assess the effect of uncertainties associated with pollutant wash-off dynamics for the modelling and the evaluation of on-site stormwater management practices. This model was initially developed to simulate both volume reduction (e.g. capture and abatement of some fraction of runoff) and flow-rate control strategies (e.g. storage and release of captured volumes). The facility consists in a simple storage unit, providing volume control through infiltration or evapotranspiration, and from which discharge may either occur as overflow (volume reduction only) or release at controlled rate through a flow limiting device (cf. Fig. 1). Because a strictly hydrological modelling approach is here adopted, specific treatment processes (e.g. settling, adsorption…) are not accounted for and pollution control hence simply results from the volume reduction associated with infiltration and evaporation.

**Figure 1**. Conceptual BMP model. 1a: volume reduction strategies, 1b: flow-rate control strategies




The BMP can be described by its size b (expressed as a ratio to drainage area), a maximum water elevation $h_{MAX}$ (mm), a permanent pool depth $h_P$ (mm) (in the case of volume reduction strategies $h_P = h_{MAX}$) and a maximum outflow rate $Q_{MAX}$ (l.s$^{-1}$.ha$^{-1}$) (flow rate control strategies only). Outflow rates are here calculated from a simple orifice function assuming that $Q_{MAX}$ is reached at the maximum water elevation $h_{MAX}$. A Green-Ampt model coupled with a conceptual soil moisture redistribution scheme introduced by Milly (1986) is implemented to simulate infiltration and evapotranspiration fluxes are calculated from meteorological records (Penman-Monteith reference evapotranspiration). Because the infiltration-redistribution model requires soil hydrodynamic parameters to be specified, a soil description based on the USDA classification (Rawls et al., 1982) is here adopted (soil type may hence be seen as an additional model parameter). Concentrations in the storage unit are finally computed considering the BMP as a perfect reactor where runoff inflow instantaneously mixes with stored water.

*Propagation of uncertainty*. So as to clarify the benefits of a solid bias description (satisfying statistical assumptions about the residuals), predictive and parameter uncertainty is propagated through the BMP model for both the “standard” and “improved” error models and BMP model response is evaluated from the annual pollutant load reduction efficiency η simulated for 2 designs scenarios (cf. table 1). The approach adopted for the propagation of uncertainties is similar the one describe in previous section (for the production of confidence intervals), although a lower number of run (500) is here performed due to the computational cost associated with BMP modelling.

**Table 1**. Configuration of the BMP model for the propagation of uncertainties ([1]corresponding saturated hydraulic conductivity is 6.8 mm/h)

| Design scenario | b (%) | $h_P$ (mm) | $h_{MAX}$ (mm) | Qmax (l/s/ha) | Soil Type |
|---|---|---|---|---|---|
| Volume reduction | 1 to 10 | 100 | 100 | 0 | Silt Loam[1] |
| Flow rate control | 5 | 0 | 400 | 5 | Silt Loam[1] |

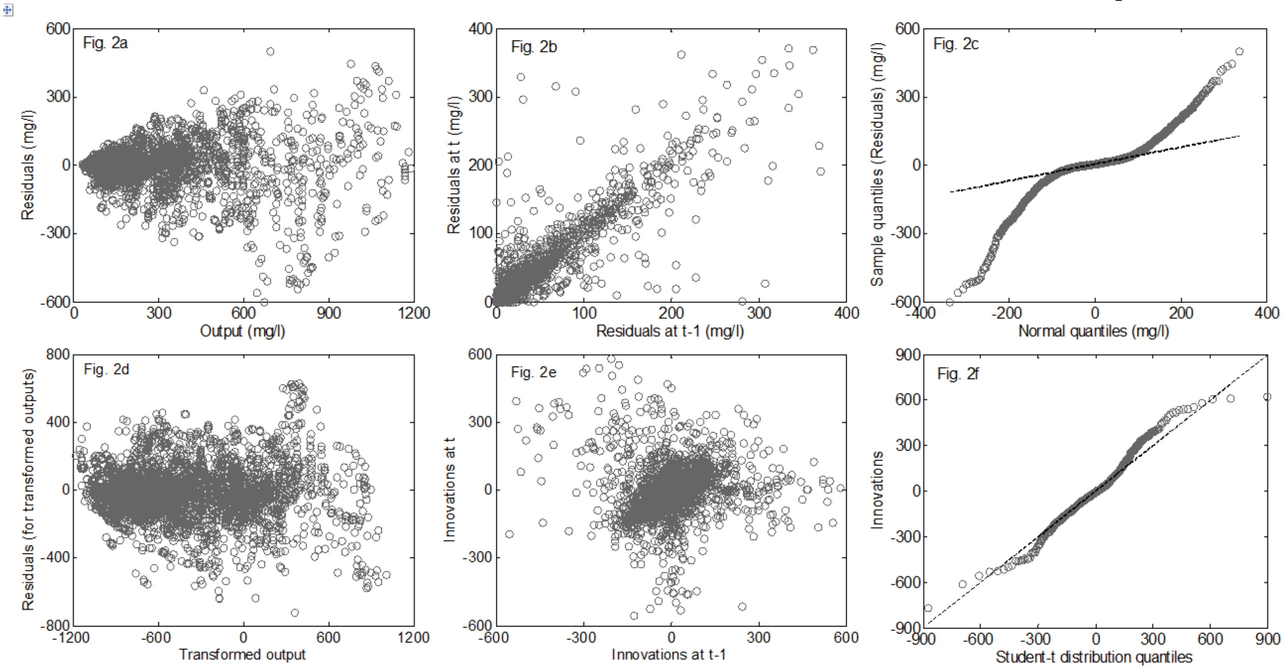
A simple linear reservoir model is implemented to generate runoff volumes from an urban street for a 1-year rainfall record (Paris region), considering 1mm initial losses, a $10^{-8}$ $m.s^{-1}$ infiltration through road surface (in addition of evaporation) and a 1-min lag time to simulate flow routing (proposed parameterization is based on the observations of Ramier et al, 2011). TSS concentrations are thus computed from simulated runoff volumes and used as inputs of the BMP model. Because the wash-off equation requires initial sediment load $M(t_{0,i})$ to be specified at the beginning of each rain event, hypotheses regarding pollutant accumulation are necessary to generate concentration time series. The widely used Alley and Smith (1981) accumulation model is therefore adopted, assuming that accumulation occurs whenever rainfall stops at very fast rate (95% of a 5g/m² equilibrium load reached within a day), which has been found to be acceptable in numerous studies (Kanso et al., 2006; Sage et al., 2015; Shaw et al., 2010).

## RESULTS AND DISCUSSION

### Water quality modelling

*Error model consistency*. As shown in figure 2, the standard assumptions of independent, homoscedastic and normally distributed residuals clearly do not hold in this study. It can be noted that residuals here exhibit a very strong first order autocorrelation (R=0.81) (Fig. 2b) which may result from model structural errors (Beven, 2005) and is probably exacerbated by the relatively high frequency of the measurements (5 min time step) (Del Giudice et al., 2013). In comparison, the improved (autoregressive with Student-t innovations and log-sinh transformation) error model appears to be much more consistent and diagnostic plots (Fig. 2d to 2e) indicate that corresponding statistical assumptions are not strongly violated (best agreement to observed residuals is obtained for $\nu = 3$). More specifically, figure 2f shows that the heavy tailed Student-t distribution better describes large and relatively infrequent errors than the Gaussian distribution does.

**Figure 2**. Diagnostic plots for the standard and improved error model (2a to 2c and 2d to 2f respectively). 2a and 2d: residuals as a function of model outputs, 2b and 2e: residuals and innovations autocorrelation, 2c and 2f: observed vs. theoretical residual or innovation quantiles.

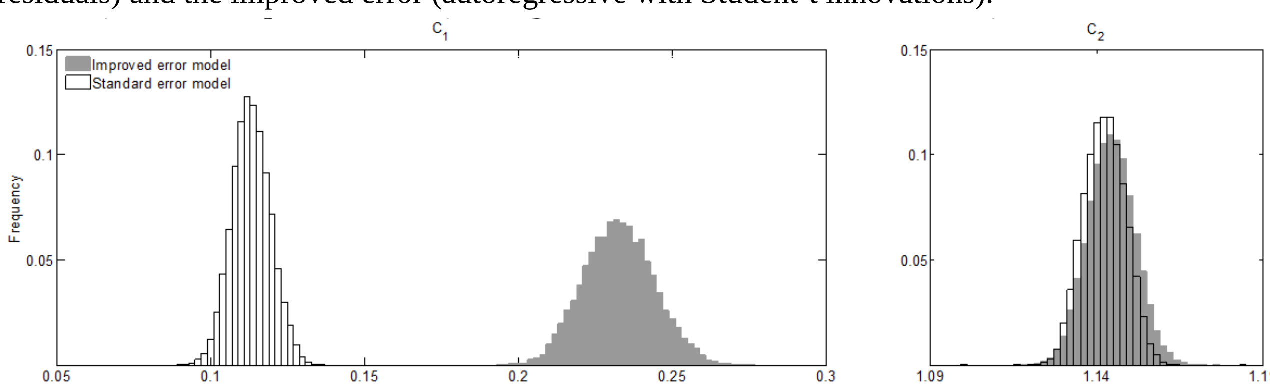


*Model performance.* The accuracy of the wash-off model is first evaluated using the Nash-Sutcliffe efficiency coefficient (E) computed for non-zero values of TSS concentrations. The overall performance the model remains very similar for both calibration approaches, although a slightly lower value of E (0.84 vs. 0.86) is obtained for the improved error model as a result of variance stabilization (log-sinh transformation) which requires the model to fit a wider portion of the pollutographs (Dotto et al., 2013). This difference nonetheless remains very moderate as the log-sinh transformation tends to preserve the least square nature of the likelihood better than other variance stabilization method do (Del Giudice et al., 2013).

As shown in Sage et al. (2015), turbidity time-series recorded at studied site however exhibit a large seasonal variation which limit the applicability of the Nash-Sutcliffe criterion since the average of observation becomes a poor predictor of reference time series (Schaefli and Gupta, 2007). Besides, because initial pollutant load $M(t_{0,i})$ is adjusted for each rain event, the model should be expected to naturally replicate the inter-event variability of TSS concentrations. Application a simple constant concentration model adjusted for each rain event indeed results in an only slightly lower efficiency (E=0.79), which suggests that model performance is in fact relatively poor (cf. Fig 2a and 2c) despite high E values. Nevertheless, detailed inspection of simulation results indicate that the wash-off model still remains a better predictor than event mean concentrations for 78% of the events and that the mean absolute percentage error for TSS predictions does not exceed 20% for half of them.

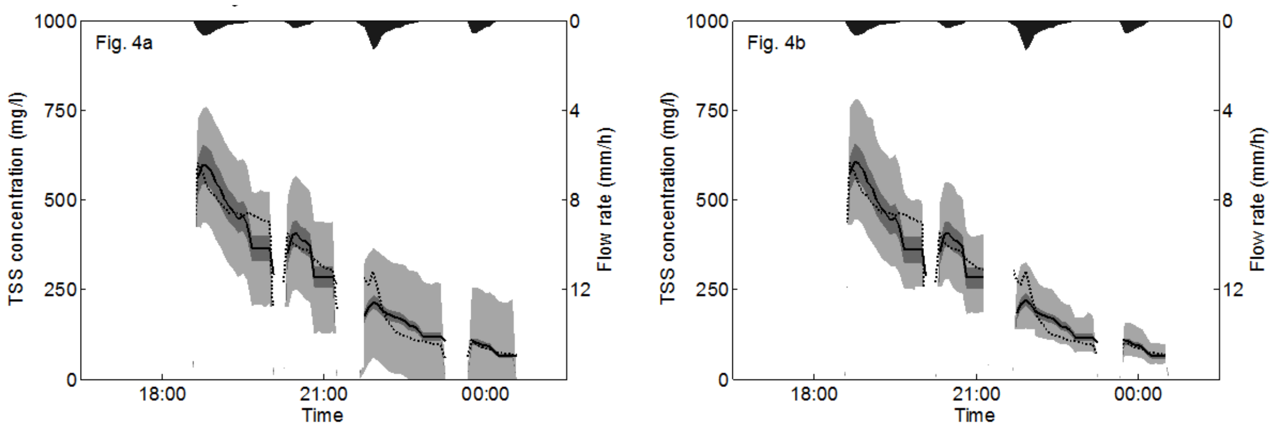
*Uncertainty analysis results.* Posterior parameter distributions estimated through the M-H algorithm for each error model are presented in figure 3. In the case of wash-off coefficient $C_1$, calibration results clearly demonstrate that a change of the likelihood function can dramatically affect the posterior distribution of model parameters, which is consistent with previous findings (Bates and Campbell, 2001; Dotto et al., 2013; Schoups and Vrugt, 2010; Yang et al., 2007). While the value of $C_1$ is mostly driven by high concentrations for the standard error model (as a result of the least-square nature of Eq. 3), the improved approach requires the model to fit a larger portion of the measurements and assigns greater probability to large prediction errors (heavy tailed distribution of the innovations) (Schoups and Vrugt, 2010), resulting in higher uncertainty and significantly different values for $C_1$. Contrariwise, posterior distribution for the second wash-off parameter $C_2$ does not significantly differ from a calibration approach to another and $C_2$ values thus probably remains equally acceptable regardless of the magnitude of output concentrations.

**Figure 3**. Calibration results for the standard (independent, homoscedastic and normally distributed residuals) and the improved error (autoregressive with Student-t innovations).



Further comparison between the standard and improved calibration approaches can be done by comparing confidence intervals generated for each error model. As shown in figure 4, the effect of parameter uncertainty does not significantly differ from an approach to another, despite the differences in both optimal parameter values and the dispersion of posterior distributions. Conversely, the standard error model clearly produces unrealistic confidence intervals for the total predictive uncertainty as it does not account for the output dependence of the residuals (cf. Fig 4a). Besides, the simple addition of a random Gaussian noise to simulated concentrations does not only result in unreliable coverage of uncertainty but also fail to capture the temporal variability of the stochastic error (Dotto et al., 2011), and its yet unclear whether such approach is acceptable if one seeks to propagate uncertainties though another model (cf. next section).

**Figure 4**. Simulation results for the 10/04/2013 event. 4a: standard error model, 4b: improved error model. The black dashed represents measured concentrations, the solid black line is simulated concentrations, the light shaded area is 5-95% total uncertainty, the dark shaded area is 5-95% parameter uncertainty and black area is flow rate over street surface

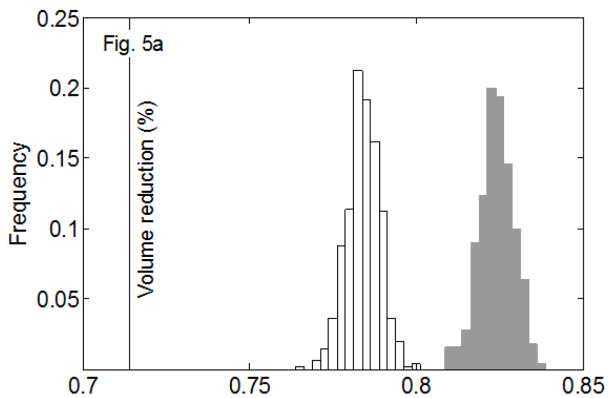



**Propagation of the uncertainty in the BMP model**

Simulation results for the propagation of the uncertainty associated with pollutant wash-off dynamics through the BMP model (cf. table 1) are shown in figure 5. For both design scenarios, total pollutant load reduction η simulated for the standard and improved error model significantly differs as a result of water quality model parameterization. Because estimates of wash-off coefficient $C_1$ based on the improved calibration approach are almost twice as large as for the standard one (cf. Fig. 3), the water quality model indeed simulates a much faster sediment removal and therefore produces a more pronounced decrease of TSS concentrations at early stages of runoff. As a consequence, first millimeters of runoff, easily captured in the BMP, represent a larger fraction of the total washed-off sediment load, resulting in higher values for η. Besides, it may be noted that η remains systematically higher than the volume reduction efficiency (cf. Fig. 5). Previous results thus suggest that accounting for the temporal variability of pollutant concentrations is probably necessary to assess the performance of stormwater management strategies, although simulation results may depend strongly on water quality model parameterization.

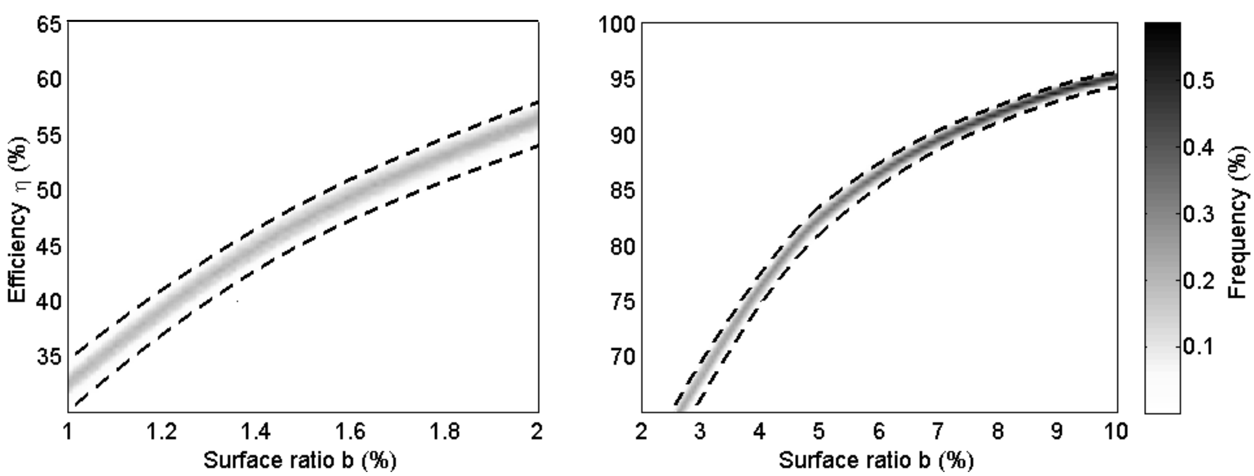
Surprisingly, while the magnitude of η clearly varies from a posterior distribution to another (depending on the value of $C_1$), the uncertainty in model outputs, represented by the dispersion of η, remains very similar for the two error models and thus presumably do not depend on error structure. This uncertainty in BMP model outputs however does not solely originate from calibration uncertainty: in the case of the volume reduction scenario, simulated efficiencies for the first and last percentile of $C_1$ (for the posterior distribution computed with the improved error model) for instance exhibit only a 1.4% percent difference, which remains relatively small as compared to the dispersion η shown in figure 5a. This result therefore indicate that the propagation of a stochastic error term to account for the uncertainty in TSS concentrations does influence BMP model outputs, although the structure of this error (statistical properties and temporal variability) apparently has no effect on simulated efficiencies. Nonetheless, the similarity in the dispersion of η for the two approaches might be related to the hypothesis of a nearly invariant initial pollutant load $M(t_{0,i})$ which reduces the variability of simulated concentrations and thus limits the incidence of the homoscedasticity assumption for the standard error model.

**Figure 5**. Results for the propagation of the uncertainty in the BMP model: distribution of load reductions. 5a: volume reduction scenario, 5b: flow-rate control scenario.




Comparison between the volume reduction and flow rate control scenario does not reveal major differences in the uncertainty associated with BMP model outputs. However, because previous approach only partially explores the relation between BMP design and the dispersion of η, propagation of the uncertainty is additionally performed for different values of the surface ratio b in the case the volume control strategy.

**Figure 6**. Results for the propagation of the uncertainty in the case of the volume reduction scenario with the improved error model: uncertainty in simulated efficiency as a function of BMP area. The dashed black line gives the 1-99% confidence interval for simulated efficiency. (Results interpolated from the distributions calculated for 10 values of b)



As indicated in figure 6, the dispersion of η clearly tends to decrease as the surface ratio b increases and the difference between the first and last percentile in simulated efficiencies ranges from 4.6% to 1.3% for b = 1% and 10% respectively. However, the magnitude of η is as well strongly influenced by b: the reduction in model output uncertainties for large values of b is therefore very expectable since an important fraction of the total runoff volume is captured in BMP, resulting in a very limited effect of the variability in TSS concentration. Besides, it is finally important to acknowledge that, regardless of the value b, the dispersion of η remains very moderate, and should probably be regarded as negligible given the numerous assumptions associated with BMP modelling.

## CONCLUSION AND PERSPECTIVES

Calibration and uncertainty analysis of a commonly used wash-off model was conducted using a formal Bayesian approach, considering two different error models, either (1) based on the unverified assumption of homoscedastic, independent and normally distributed residuals (standard error model) or (2) assuming heteroscedastic and autoregressive errors (improved error model). For both approaches, the uncertainty associated with water quality modelling was propagated through a conceptual BMP model, whose response was evaluated from the total pollutant load reduction efficiency simulated over a 1-year period. The results of this study can be summarized as follow:

- In the case of pollutant wash-off modelling, good agreement with the statistical assumptions about the residuals could be achieved with the heteroscedastic and autoregressive error model (for Student-t innovations). Parameter distribution estimated for the improved calibration approach significantly differed from the one obtained with the standard and unverified hypotheses. Besides, the standard error model was found produce unreliable predictive confidence intervals due the heteroscedasticity of the residuals. Further research is however believably needed to identify the most important statistical hypotheses to be verified for a robust assessment of parameter uncertainty.

- The magnitude of pollutant removal efficiency simulated by the BMP model after propagation of the uncertainties associated with wash-off dynamics significantly differed from an error model to another as a result of the differences in parameter posterior distributions. This finding therefore indicates that BMP model outputs are in fact quite sensitive to the intra-event variability of inflow concentrations and suggests that erroneous representation of the pollutant wash-off dynamics may bias the assessment of the performance of Best Management Practices (BMP).

- For both the standard and improved approaches, pollutant removal efficiencies simulated by the BMP model were found to exhibit a very similar dispersion. While the use of a statistically correct error model is clearly needed for calibration, it yet unclear whether it is justified for the propagation of uncertainty through another model. Besides, the uncertainty in model outputs apparently remained very moderate regardless of BMP design, which casts doubt on the necessity of accounting for the uncertainty associated with the intra-event variability of concentrations in runoff.

## ACKNOWLEDGEMENT

This research was carried out under the OPUR research program. The authors gratefully acknowledge OPUR partners and the French Ministry of Ecology, Sustainable Development and Energy for financial support.

## REFERENCES

Ahiablame, L.M., Engel, B.A., Chaubey, I., 2012. Effectiveness of Low Impact Development Practices: Literature Review and Suggestions for Future Research. Water, Air, Soil Pollut. 223, 4253–4273. doi:10.1007/s11270-012-1189-2

Alley, W.M., Smith, P.E., 1981. Estimation of accumulation parameters for urban runoff quality modeling. Water Resour. Res. doi:10.1029/WR017i006p01657

Bates, B.C., Campbell, E.P., 2001. A Markov chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling. Water Resour. Res. 37, 937–947. doi:10.1029/2000WR900363

Beven, K., 2005. On the concept of model structural error. Water Sci. Technol. 52, 167–175.

Bressy, A., Gromaire, M.C., Lorgeoux, C., Saad, M., Leroy, F., Chebbo, G., 2014. Efficiency of source control systems for reducing runoff pollutant loads: Feedback on experimental catchments within Paris conurbation. Water Res. 57, 234–246. doi:10.1016/j.watres.2014.03.040

Chib, S., Greenberg, E., 1995. Understanding the Metropolis-Hastings algorithm. J. Am. Stat. Assoc. 49, 327–335. doi:10.2307/2684568

Del Giudice, D., Honti, M., Scheidegger, A., Albert, C., Reichert, P., Rieckermann, J., 2013. Improving uncertainty estimation in urban hydrological modeling by statistically describing bias. Hydrol. Earth Syst. Sci. 17, 4209–4225. doi:10.5194/hess-17-4209-2013

Deletic, A., Orr, D., 2005. Pollution Buildup on Road Surfaces. J. Environ. Eng. 131, 49–59. doi:10.1061/(ASCE)0733-9372(2005)131:1(49)

Dotto, C.B.S., Deletic, A., McCarthy, D.T., 2013. Uncertainty analysis in urban drainage modelling: should we break our back for normally distributed residuals? Water Sci. Technol. 68, 1271–9. doi:10.2166/wst.2013.360.

Dotto, C.B.S., Kleidorfer, M., Deletic, A., Fletcher, T.D., McCarthy, D.T., Rauch, W., 2010. Stormwater quality models: performance and sensitivity analysis. Water Sci. Technol. 62, 837–843. doi:10.2166/wst.2010.325

Dotto, C.B.S., Kleidorfer, M., Deletic, A., Rauch, W., McCarthy, D.T., Fletcher, T.D., 2011. Performance and sensitivity analysis of stormwater models using a Bayesian approach and long-term high resolution data. Environ. Model. Softw. 26, 1225–1239. doi:10.1016/j.envsoft.2011.03.013

Evin, G., Kavetski, D., Thyer, M., Kuczera, G., 2013. Pitfalls and improvements in the joint inference of heteroscedasticity and autocorrelation in hydrological model calibration. Water Resour. Res. 49, 4518–4524.

Freni, G., Mannina, G., Viviani, G., 2009. Urban runoff modelling uncertainty: Comparison among Bayesian and pseudo-Bayesian methods. Environ. Model. Softw. 24, 1100–1111. doi:10.1016/j.envsoft.2009.03.003

Hasting, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57, 97–109. doi:10.1093/biomet/57.1.97

Kanso, A., Chebbo, G., Tassin, B., 2006. Application of MCMC-GSA model calibration method to urban runoff quality modeling. Reliab. Eng. Syst. Saf. 91, 1398–1405. doi:10.1016/j.ress.2005.11.051

Li, L., Xu, C.-Y., Xia, J., Engeland, K., Reggiani, P., 2011. Uncertainty estimates by Bayesian method with likelihood of AR (1) plus Normal model and AR (1) plus Multi-Normal model in different time-scales hydrological models. J. Hydrol. 406, 54–65.

Milly, P.C.D., 1986. An event-based simulation model of moisture and energy fluxes at a bare soil surface. Water Resour. Res. doi:10.1029/WR022i012p01680

Ramier, D., Berthier, E., Andrieu, H., 2011. The hydrological behaviour of urban streets: long-term observations and modelling of runoff losses and rainfall–runoff transformation. Hydrol. Process. 25, 2161–2178. doi:10.1002/hyp.7968

Rawls, W.J., Brakensiek, D.L., Saxton, K.E., 1982. Estimation of Soil Water Properties. Trans. ASAE 25, 1316–1320 & 1328.

Sage, J., Bonhomme, C., Al Ali, S., Gromaire, M.-C., 2015. Performance assessment of a commonly used "accumulation and wash-off" model from long-term continuous road runoff turbidity measurements. Water Res. 78, 47–59. doi:http://dx.doi.org/10.1016/j.watres.2015.03.030

Schaefli, B., Gupta, H. V., 2007. Do Nash values have value? Hydrol. Process. doi:10.1002/hyp.6825

Schoups, G., Vrugt, J.A., 2010. A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors. Water Resour. Res. 46. doi:10.1029/2009WR008933

Schriewer, A., Horn, H., Helmreich, B., 2008. Time focused measurements of roof runoff quality. Corros. Sci. 50, 384–391. doi:10.1016/j.corsci.2007.08.011

Shaw, S.B., Stedinger, J.R., Walter, M.T., 2010. Evaluating Urban Pollutant Buildup/Wash-Off Models Using a Madison, Wisconsin Catchment. J. Environ. Eng. doi:10.1061/(ASCE)EE.1943-7870.0000142

Thyer, M., Renard, B., Kavetski, D., Kuczera, G., Franks, S.W., Srikanthan, S., 2009. Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis. Water Resour. Res. 45. doi:10.1029/2008WR006825

Vaze, J., Chiew, F.H.S., 2002. Experimental study of pollutant accumulation on an urban road surface. Urban Water 4, 379–389. doi:10.1016/S1462-0758(02)00027-4

Vezzaro, L., Mikkelsen, P.S., 2012. Application of global sensitivity analysis and uncertainty quantification in dynamic modelling of micropollutants in stormwater runoff. Environ. Model. Softw. 27–28, 40–51. doi:http://dx.doi.org/10.1016/j.envsoft.2011.09.012

Wang, Q.J., Shrestha, D.L., Robertson, D.E., Pokhrel, P., 2012. A log-sinh transformation for data normalization and variance stabilization. Water Resour. Res. 48. doi:10.1029/2011WR010973

Yang, J., Reichert, P., Abbaspour, K.C., 2007. Bayesian uncertainty analysis in distributed hydrologic modeling: A case study in the Thur River basin (Switzerland). Water Resour. Res. doi:10.1029/2006WR005497

Zhao, H., Li, X., Wang, X., 2011. Heavy metal contents of road-deposited sediment along the urban-rural gradient around beijing and its potential contribution to runoff pollution. Environ. Sci. Technol. 45, 7120–7127. doi:10.1021/es2003233