Model based approaches for rainfall estimation in urban catchments

Estimation de la pluie à l’aide de modèles de bassins versants urbains

Günther Leonhardt*, Siao Sun**, Wolfgang Rauch*, Jean-Luc Bertrand-Krajewski**

* University of Innsbruck, Unit of Environmental Engineering, Technikerstrasse 13, A-6020 Innsbruck, Austria
guenther.leonhardt@uibk.ac.at
wolfgang.rauch@uibk.ac.at

** Université de Lyon, INSA Lyon, LGCIE, 34 avenue des arts, F-69621 Villeurbanne Cedex, France
siao.sun@insa-lyon.fr
jean-luc.bertrand-krajewski@insa-lyon.fr

RÉSUMÉ

Dans cet article, nous comparons deux méthodes pour estimer l’intensité moyenne surfacique de la pluie sur des bassins versants urbains. La première méthode est un modèle inverse d’un modèle pluie-débit. Avec ce modèle, on estime les intensités de la pluie et leurs incertitudes à partir du débit mesuré uniquement. La deuxième approche estime les paramètres d’un modèle d’erreur de la pluie observée avec une méthode bayésienne. Cette approche nécessite les données de débit et de pluie. Les deux méthodes reposent sur le même modèle hydrologique global. Bien que les deux méthodes s’appuient sur des concepts différents, elles traitent du même problème, à savoir la quantification de l’intensité moyenne surfacique de la pluie et des erreurs associées. Leur comparaison est donc pertinente. Les résultats montrent que les deux méthodes estiment les hyérommes avec des hauteurs totales et des intensités maximales de pluie réalisistes. Les incertitudes estimées avec le modèle inverse sont relativement élevées.

ABSTRACT

We compare two different approaches to estimate mean areal rainfall intensity in urban catchments. The first method is based on a reverse model, i.e. an inverse formulation of a rainfall runoff model. Based on calibrated model parameters, rainfall intensities and their uncertainties are estimated from flow data only. The second method estimates parameters of a rainfall error model using a Bayesian approach. It requires measurements of both flow and rainfall. Both methods are based on the same lumped hydrological model. Although the two approaches are conceptually rather different, they address the same issue – the quantification of areal rainfall intensities and their related measurement errors, respectively – and are therefore considered interesting to be compared. Results show that both methods provide best estimates of hyetographs with maximum intensities and total depths in a realistic order of magnitude, whereas the uncertainty of rainfall estimated with the reverse model is rather large.

KEYWORDS

Conceptual hydrological model, Error model, Rainfall estimation, Rainfall uncertainty, Reverse model
1 INTRODUCTION – ESTIMATION OF RAINFALL ERRORS

Good information about rainfall intensity as the driving force of many relevant processes in urban drainage systems is of great importance. Rainfall is highly variable in space and time, and this variability is assumed to have a great influence on runoff quantity and quality at the outlet of a typical urban catchment. It should therefore be considered in simulation models (Schilling, 1984).

For many applications in urban drainage lumped conceptual models are used. These models are usually based on areal rainfall as model input, i.e. rainfall is assumed to be uniform in space. Rain gauges as common devices for rainfall observation perform point measurements, thus it is difficult to get information about the spatial variability and areal rainfall. Areal rainfall for a catchment can be estimated from several rain gauges by interpolation. However, as the rain gauge network is usually sparse with respect to the size of urban catchments, data from a single rain gauge is often assumed to represent areal precipitation and used as model input. Alternative methods such as rainfall radar (Einfalt et al., 2004; Marshall et al., 1947) or microwave links (Messer et al., 2006) provide information about the spatial variability, but their accuracy is still limited as they estimate rainfall indirectly.

To improve model results, methods to estimate the rainfall error, i.e. the error of the rainfall data used as model input, have been developed. They range from simple approaches as areal reduction or correction factors (Vaes et al., 2005) to determine design storm intensities to very sophisticated methods based on dynamic error models, considering variation of the rainfall errors with time.

In this paper, we compare two model based methods to estimate rainfall errors with the purpose to determine the “true” areal rainfall. Although the two methods are based on lumped hydrological models, they are conceptually rather different. However, a comparison is interesting as they address the same important question.

The first method is based on a reverse model, i.e. a model to simulate rainfall based on observed runoff. The reverse model is an inverse formulation of a lumped rainfall runoff model. Net areal rainfall intensities and their uncertainties are estimated based on runoff data and corresponding uncertainties by Monte Carlo (MC) simulation. No further model inputs are required. The second method uses a combination of an error model and a rainfall runoff model. Parameters of the error model representing rainfall measurement errors are estimated by Bayesian inference, i.e. a Markov-Chain Monte Carlo (MCMC) approach, based on measurements of both rainfall and flow. Both methods are based on the same lumped conceptual model. However, whereas the reverse model estimates the rainfall using flow data only, the error model considers also rain gauge data. The comparison is performed by estimating areal rainfall in an urban catchment drained by storm sewers in Lyon, France.

The estimated rainfall can be used as input to other models, e.g. to simulate stormwater quality or to assess the quality of the rain gauge data. However, further applications are beyond the scope of this paper.

2 MATERIALS AND METHODS

As the rainfall runoff model and the reverse model are based on the same model concepts, they have the same model parameters. Model parameters are calibrated by a state of the art procedure, based on a calibration data set comprising rainfall and flow data. The reverse model is then applied for an application data set. For the same data set, the parameters of the error model are estimated and rainfall errors are computed.

2.1 Hydrological Model and Calibration

As already mentioned, a lumped conceptual hydrological model is considered in this study. The model consists of two parts: a rainfall loss component and a routing function. The gross rainfall is transformed to net rainfall after rainfall loss deduction. The net rainfall is then fed into the routing function to generate runoff.

The rainfall loss incorporates an initial loss and a proportional loss. It is formulated as:

\[
L(t) = \begin{cases} 
  i & \text{if } \int_0^t i \, dt \leq L_0 \\
  ip_{cons} & \text{if } \int_0^t i \, dt > L_0 
\end{cases}
\]

where \( L(t) \) is the rainfall loss (mm/h), \( i \) (mm/h) is the rainfall intensity, \( L_0 \) and \( p_{cons} \) are the two model parameters representing the initial loss (mm) and the proportional loss (-). Net rainfall \( i_{net} \) can then
simply be computed as follows:
\[ i_{\text{net}}(t) = i(t) - L(t) \]  
(2)

The routing function is modelled with two cascaded linear reservoirs. The net rainfall \( i_{\text{net}} \) is converted to the inflow to the first reservoir with a lag time \( T_{\text{lag}} \) by multiplying with the impervious catchment area \( A \):
\[ Q_{\text{in}}(t) = i_{\text{net}}(t - T_{\text{lag}}) \cdot A \]  
(3)

A linear reservoir has its outflow \( Q_{\text{out}} \) varying linearly with its storage volume \( V \):
\[ Q_{\text{out}}(t) = V(t) / K \]  
(4)

where the parameter \( K \) (min) is called reservoir constant. Eq. (4) is combined with an equation of continuity:
\[ Q_{\text{in}}(t) - Q_{\text{out}}(t) = dV(t) / dt \]  
(5)

The analytical solution of Eq. (4) can be derived by integration of the differential equation over the time interval \([t - \Delta t, t]\):
\[ Q_{\text{out}}(t) = \exp\left(-\frac{\Delta t}{K}\right)Q_{\text{out}}(t - \Delta t) + \left[1 - \exp\left(-\frac{\Delta t}{K}\right)\right]Q_{\text{in}}(t) \]  
(6)

A second linear reservoir is placed in series after the first reservoir. The reservoir constants of the two cascaded reservoirs are set to the same value because they are heavily correlated when determining them in calibration according to some preliminary tests. A baseflow \( q \) to account for dry weather flow is simply added to the outflow from the second reservoir as a contribution to the runoff. In total five parameters, i.e. initial rainfall loss \( L_0 \), proportional loss \( p_{\text{cons}} \), lag time \( T_{\text{lag}} \), reservoir constant \( K \) and baseflow \( q \) are required in the conceptual hydrological model.

The model is calibrated with the DREAM algorithm developed by Vrugt et al. (2009). DREAM employs the concept of Bayesian statistics that considers model parameters as probabilistic variables. The joint posterior probability density function of model parameters is updated according to observed inputs and outputs to capture the probabilistic belief about parameters. DREAM is an adapted MCMC approach using a differential evolution with multiple chains in parallel to provide a robust exploration of the search space of calibration parameters. DREAM has been demonstrated to be effective for calibration in hydrological related problems (e.g. Schoups and Vrugt, 2010). For more details about DREAM please refer to Vrugt et al. (2009).

### 2.2 Reverse model

The reverse model \( i_{\text{net}} = f(Q, \theta_r) \) simulates rainfall from runoff \( Q \) based on the same concepts as the rainfall-runoff model described above. The rainfall \( i_{\text{net}} \) (mm/h) is the part of rainfall contributing to runoff. It corresponds to gross rainfall after subtraction of the initial loss \( L_0 \). The parameters of the reverse model \( \theta_r \) correspond to the parameters of the rainfall runoff model \( \theta \) except for the initial loss. The model can thus not estimate the beginning of gross rainfall. Applications of reverse models to estimate rainfall can be found in Hino (1986), Marceau (1997) or Leonhardt et al. (2012).

To derive the reverse model, we formulate a part of the rainfall runoff model as a set of linear equations:
\[ M \cdot \bar{q} = \bar{q}_r \]  
(7)

A similar formulation can be found in Hino (1986). The vectors \( \bar{r} \) and \( \bar{q}_r \), both of length \( n \), correspond to the inflow and outflow of the cascade of linear reservoirs mentioned above. The matrix \( M \) represents the model transfer function, i.e. the two reservoirs in series with the corresponding parameters \( K \), and is of size \( n \times n \):
The elements of each column below the main diagonal \( h = (h_1, h_2, \ldots, h_n)^T \) correspond to the model's response to unit net rainfall input, which can be computed by applying equation (6) twice. In the case of a linear reservoir the values of the elements of \( h \) are approaching zero asymptotically. \( \tilde{q}_r \) is computed from measured total flow by deduction of the baseflow \( q \), but is always greater than or equal to zero. Basically \( \tilde{q}_r \) can be estimated by solving the linear equations as follows:

\[
\tilde{q}_r = M^{-1} \cdot \tilde{q}_r.
\]  

However, only if the model describes the catchment's behaviour perfectly and \( \tilde{q}_r \) is not perturbed by any measurement noise, the solution \( \tilde{q}_r \) is physically meaningful, i.e. \( \tilde{q}_r \geq 0 \), and fulfils the mass balance. If \( \tilde{q}_r \) is noisy and the model is not perfect, constraints must be imposed in order that the estimated \( \tilde{q}_r \) is physically meaningful and fulfils the mass balance. We use the least squares solution of equation (9) as estimate for \( \tilde{q}_r \). To obtain rainfall \( \hat{i}_{\text{net}} \) contributing to runoff, \( \tilde{q}_r \) is shifted backwards by the lag time \( T_{\text{lag}} \) and divided by the contributing area \( A \cdot (1 - p_{\text{cons}}) \):

\[
\hat{i}_{\text{net}}(t) = \frac{\tilde{q}_r(t + T_{\text{lag}})}{A \cdot (1 - p_{\text{cons}})}.
\]  

The values of the reverse model parameters are taken from the direct model calibration. As already reported by other authors (e.g. Hino, 1986; Koussis \textit{et al.}, 2012), reverse models amplify errors and therefore cause strong oscillations in estimated hyetographs. This can only be partly tackled by imposing constraints as described above. A single solution is not unrealistic, but as it is very sensitive to small changes in runoff data, we consider it as a sample from a population of many possible estimates of areal rainfall. To obtain a more reliable picture of possible areal rainfall and estimate its uncertainty, we perform a MC simulation. Flow samples are generated, according to the standard error of the flow data, and the reverse model is applied for each flow sample.

### 2.3 Rainfall error model estimation

The measured rainfall is considered with accompanied uncertainty:

\[
x = \tilde{x} + e_x
\]  

where \( x \) and \( \tilde{x} \) represent the real and the measured rainfall, \( e_x \) is the rainfall error.

The Bayesian method provides a way to update posterior probability distributions of unknown parameters. Assuming the rainfall errors are unknown parameters needing to be determined as:

\[
p(e_x | \hat{y}, \hat{x}, \theta) \propto p(\hat{y} | \hat{x}, e_x, \theta) p(\hat{x} | e_x) p(e_x)
\]  

where \( \hat{y} \) is the measured output (runoff). This formula is very similar to the Bayesian total error analysis methodology used by Kavetski \textit{et al.} (2006) but without involving evaluation of the model parameters \( \theta \). In this case of rainfall error estimation, we assume that the hydrological model has been identified by pre-executed calibration that is independent of error estimation. Instead of calibrating \( \theta \) in a normal calibration problem, model errors \( e_x \) are calibrated. Assuming errors in runoff measurements are independent and normally distributed:

\[
p(\hat{y} | \hat{x}, e_x, \theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi \sigma_{\hat{y}_i}^2}} \exp\left(-\frac{(\hat{y}_i - y_i(x))^2}{2\sigma_{\hat{y}_i}^2}\right)
\]  

Similarly if rainfall measurement errors have independent normal distributions:
Theoretically the error of each rainfall measurement can be viewed as a parameter needing to be determined. However, this approach makes the dimensionality of the problem massive and the problem might be over-parameterized. Thus the dimension of the problem is reduced by grouping several neighbouring rainfall measurements, with the assumption that rainfall measurements in a group share the same rainfall error. The algorithm DREAM is employed to search for proper rainfall errors.

2.4 Case study

The catchment of Chassieu, located in the East of Lyon, France, is selected as the study catchment to test the two above-mentioned methods. Chassieu is one of the experimental sites in the OTHU project (Field Observatory for Urban Hydrology - www.othu.org). It covers 185 ha and is mainly an industrial area that is equipped with a separate storm sewer system. In the hydrological model, a rough impervious area of 80 ha is assumed and the imperviousness is adjusted in calibration by the proportional rainfall loss parameter.

Catchment runoff is measured with two-minute time steps and rainfall with one-minute time steps. The hydrological model is calibrated using five events in 2007. They cover different rainfall characteristics, with the total rainfall depth ranging from 6 to 60 mm and maximum intensities between 19 and 88 mm/h. Observed peak flows range from 1062 to 2656 L/s. Uncertainty in flow measurements is derived using the law of propagation of uncertainty (LPU), as runoff is not measured directly but computed from several other measurements including geometric parameters of the channel, water depth and mean flow velocity (refer to Bertrand-Krajewski and Bardin, 2002 for details). Relative standard uncertainties in flow measurements approximately range from 15 to 25%. Although it is a storm sewer, a certain amount of dry weather flow is constantly observed. To minimize computational effort, an event based approach was used for calibration. The calibrated model is then applied to another event also measured in 2007 to estimate the rainfall (see Figure 3).

3 RESULTS

3.1 Model calibration results

The conceptual hydrological model is calibrated using DREAM with uniform prior probability for each parameter with the range identified according to the physical meaning of the parameter or catchment characteristics. In this case, a total of 100 000 model evaluations are sufficient to identify a stable calibration result. Parameter distributions are identified according to values evolved in the last 20% of the model simulations that constitute posterior samples. Figure 1 shows the uncertainty in estimated parameters as well as their initial ranges. The evaluated baseflow approaches the limit of its searching range. We did not relax the prior range because we respect the fact that the baseflow, by observation under dry weather conditions, is normally under 5 L/s. The observed and simulated runoffs (using the highest-likelihood parameters) of one calibration event are given in Figure 2 as an example. With the highest-likelihood parameter set, the Nash-Sutcliffe (N-S) indexes (Nash and Sutcliffe, 1970) of simulations of all events are higher than 0.9. Uncertainty in runoff simulations due to parameter estimations is hardly observable and is thus not shown in the figure. The optimal parameter set of \([1.36, 0.34, 1.92, 11.3, 5.0]\) will be used in the following studies and uncertainty in model parameter estimation will be ignored.
3.2 Reverse model

A MC simulation comprising 10 000 model runs was performed with the reverse model to estimate the net areal rainfall and corresponding uncertainty. Flow samples were randomly generated based on the provided standard error of the data, assuming a normal distribution. In addition, a single hyetograph was deterministically estimated from the flow data, without consideration of uncertainties.

Table 1 lists total rainfall depths and maximum intensities of simulation results and observed data. With regard to these quantities the median of the MC samples and the deterministic model run show very similar results. In Figure 3 plots (a) and (b) show the flow data used as model input and results of the MC simulation in terms of minimum and maximum values of all samples as well as the 2.5- and 97.5-percentiles. The median of all samples, the deterministically simulated rainfall and measured rainfall intensities are also compared in plot (c) (Figure 3). The estimations of the MC simulation cover a rather wide range, from zero to extremely high intensities as 28.4 mm/h for single time steps. However, the 95 % uncertainty band is significantly smaller (12.6 mm/h maximum intensity), indicating distributions with flat tails and a high kurtosis. The median of all samples as well as the deterministically estimated rainfall from flow data both show rainfall intensities in the same order of magnitude as the measured rainfall (initial loss was deducted from observed gross rainfall for these plots). Some peaks are shifted by a positive lag of about 25 minutes.
To check the simulated rainfall, runoff is simulated from all estimated rainfall samples. The results of these simulations show good agreement with the observed flow (see plot (d) in Figure 3, the N-S index of the median is above 0.99), which is of course expected. The uncertainty of the simulated flow is smaller than that of the input data.

Table 1: Total rainfall depth and maximum intensity of observed and simulated rainfall; initial loss was deducted from observed rainfall and error model results.

<table>
<thead>
<tr>
<th></th>
<th>observed median</th>
<th>reverse model p2.5</th>
<th>reverse model p97.5</th>
<th>deterministic reverse model best</th>
<th>error model p2.5</th>
<th>error model p97.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>total depth in mm</td>
<td>30.3</td>
<td>35.1</td>
<td>14.2</td>
<td>62.7</td>
<td>36.1</td>
<td>31.6</td>
</tr>
<tr>
<td>maximum intensity in mm/h</td>
<td>5.4</td>
<td>6.51</td>
<td>3.8</td>
<td>12.6</td>
<td>6.3</td>
<td>6.7</td>
</tr>
</tbody>
</table>

Figure 3: Reverse model input and output: (a) observed flow data and corresponding expanded uncertainty U(Q); (b) results of the MC simulation with the reverse model and observed rainfall; (c) median of the estimates from the reverse model; (d) simulated flow.
reverse model, estimate of the deterministic reverse model and observed rainfall (converted to 2 min time step, initial loss deducted); (d) flow simulated from all rainfall estimates and observed flow.

3.3 Error model estimation

Uncertainty in rainfall measurements has been evaluated by experts to have a joint independent normal distribution with standard deviation being of the order of 0.1 times the measurement. Because the inverse model provides a rainfall estimation for each time step, for comparison reason, a small number - five rainfall measurements - are placed in a group to share a single error parameter to be evaluated using the Bayesian method (Eq. (12)). With a trial and error approach, it was found that a total of 100 000 model evaluations are sufficient to provide a stable evaluation of parameters. The posterior probability distributions of normalized rainfall errors (rainfall errors divided by the standard deviations) are shown in Figure 4 in the form of box plots as well as the measured rainfall. Figure 5 displays the simulated runoff with and without rainfall error consideration. The N-S indexes of runoff simulations with and without rainfall error consideration are respectively 0.94 and 0.87.

To show results which can be compared, net rainfall intensities were calculated using the error model (Figure 6). The initial loss was deducted from gross rainfall. Percentiles of the remaining rainfall were calculated based on the distributions of the rainfall error \( c_r \) and considering the grouping. The changes made by the error models are moderate because they are computed based on measurements. The relative uncertainty is smaller in comparison to the reverse model.
4 DISCUSSION

Rainfall estimated by the two methods is in the same order of magnitude as measured data in terms of both total depth and maximum intensities. Both results can therefore be considered as physically meaningful and realistic estimates of the areal rainfall. They depend of course on the model concept and real areal rainfall is unfortunately unknown.

For the investigated event, estimates of total rainfall depth as well as maximum intensity by both approaches are higher than the measured values. This can be explained by the performance of the calibrated rainfall runoff model (see Figure 5), which underestimates peak flows. As the model also underestimates the total flow volume, the total rainfall depth simulated with the reverse model is higher, because mass conservation is imposed as constraint for the solution. The reverse model transforms all flow exceeding the baseflow into rainfall. In contrast, the error model “revises” measured rainfall moderately with respect to the likelihood function in Eq (12). In addition, the revision has no effect if the observed value equals zero due to the error assumptions. Therefore, the total depth simulated with the error model is lower than that by the reverse model. Those differences are of course event specific. The lag between estimated rainfall by the reverse model and the observed rainfall towards the end of the events (last three peaks) reflects the lag between observed and simulated runoff. Possibly, the storm was moving and areal intensities therefore are different with respect to peak time and maximum peak intensity (last peak). This effect cannot be observed in the results of the error model, as corrections are constraint by the observed intensities.

Uncertainties of estimated rainfall are significantly larger for the reverse model than for the error model. It is worth mentioning that the uncertainty sources by the two approaches are essentially different: uncertainty of the reverse model originates in uncertainty of runoff measurements, while uncertainty estimated by error models evaluates uncertainty due to measurements of rainfall itself. In the reverse model, even a small change in the flow can cause a large change in single estimated rainfall intensities, but not in mean or total rainfall because the reverse model necessarily compensates such short term fluctuations. The range therefore represents uncorrelated, random uncertainties. Each hyetograph randomly sampled from these ranges results in a hydrograph within the uncertainty bands of the observed flow data. The median of all rainfall samples can be considered as the best estimate of areal rainfall. In the case of larger uncertainties of model parameters, they should of course be considered in the reverse model.

The uncertainty in rainfall estimation by the reverse model can only be reduced with improvement of runoff measurements. Regarding to error models, uncertainty evaluations in rainfall in principle should not be affected by runoff measurement uncertainty. However the estimation depends much on the
correct judgements on the distributions of uncertainty in rainfall and runoff that are used as likelihood functions in Bayesian inference if the model is assumed to be correct. Such judgements should be made based on the measuring techniques and methods. In addition, model structure uncertainty is not considered in this paper but should be incorporated if it contributes significantly to the total uncertainty. Although the estimated areal rainfall is conditional on the model and true areal rainfall remains unknown, the approaches could possibly be used for calibration of indirect rainfall observation methods as radar or microwave links.

5 CONCLUSION

In this paper, we presented and compared two methods to estimate the error of measured rainfall and areal precipitation, respectively. The methods are conceptually rather different, but they both address the issue of areal rainfall estimation, as they are based on the same lumped hydrological model. The main difference between the two approaches lies in the used information and the sources of uncertainty. The reverse model only requires flow data. A MC simulation is performed to propagate the uncertainties in flow data through the model and estimate the uncertainty of areal rainfall. The parameters of the error model are estimated based on both measured rainfall and flow using information on distributions of uncertainty in rainfall and flow. Generally, the uncertainty derived with the reverse model is greater than the estimates based on the error model.

The best estimates of both models show similar and realistic rainfall intensities and total depths and can therefore be considered as good estimates of areal rainfall. They are of course dependent on the model concept.

6 ACKNOWLEDGEMENTS

The first author was supported by a Marietta-Blau-Grant from OeAD – GmbH, funded by the Austrian Federal Ministry of Science and Research (BMWF), to stay at LGCIE, INSA Lyon as a visiting fellow.

LIST OF REFERENCES


