



Bayesian approach for the calibration of models: application to an urban stormwater pollution model

A. Kanso, Marie-Christine Gromaire, E. Gaume, Bruno Tassin, Ghassan Chebbo

► To cite this version:

A. Kanso, Marie-Christine Gromaire, E. Gaume, Bruno Tassin, Ghassan Chebbo. Bayesian approach for the calibration of models: application to an urban stormwater pollution model. *Water Science and Technology*, 2003, 47 (4), pp.77–84. <hal-00711800>

HAL Id: hal-00711800

<https://enpc.hal.science/hal-00711800v1>

Submitted on 29 Mar 2015

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.



HAL Authorization

Bayesian approach for the calibration of models: application to an urban stormwater pollution model

A. Kalso*, M.-C. Gromaire*, E. Gaume*, B. Tassin* and G. Chebbo*,**

* CERREVE (Centre d'Enseignement et de Recherche sur l'Eau, la Ville et l'Environnement), Ecole Nationale des Ponts et Chaussées, 6–8 avenue Blaise-Pascal, Cité Descartes, 77455 Marne-la-Vallée, Cedex 2, France

** School of Engineering, Lebanese University, Route de l'aéroport, Beirut, Lebanon

(E-mail: kalso@cereve.enpc.fr; mcm@cereve.enpc.fr; gaume@cereve.enpc.fr; tassin@cereve.enpc.fr; chebbo@cereve.enpc.fr)

Abstract In environmental modelling, estimating the confidence level in conceptual model parameters is necessary but difficult. Having a realistic estimation of the uncertainties related to the parameters is necessary i) to assess the possible origin of the calibration difficulties (correlation between model parameters for instance), and ii) to evaluate the prediction confidence limits of the calibrated model. In this paper, an application of the Metropolis algorithm, a general Monte Carlo Markov chain sampling method, for the calibration of a four-parameter lumped urban stormwater quality model is presented. Unlike traditional optimisation approaches, the Metropolis algorithm identifies not only a "best parameter set", but a probability distribution of parameters according to measured data. The studied model includes classical formulations for the pollutant accumulation during dry weather period and their washoff during a rainfall event. Results indicate mathematical shortcomings in the pollutant accumulation formulation used.

Keywords Bayesian inference; conceptual model; Monte Carlo Markov chain; parameter uncertainty; urban storm water

Introduction

It is well accepted that the quality of urban stormwater is a significant source of pollution for the receiving systems and has become a major concern for urban water management.

This pollution results mostly from the erosion by the runoff of particulate pollutants accumulated on the watersheds during the dry weather period, which is mixed with the sediments eroded in the sewers.

Managers need tools to evaluate and control storm waters according to quantitative and qualitative criteria. For this purpose, models have been built to simulate the water cycle in the city for both the quantitative and qualitative aspects. Hydrologic models are now operational and widely used by managers. However, urban storm water quality models that have been proposed since the seventies, are often not reliable and remain rarely used for operational applications.

These models are conceptual because the processes involved cannot be described mechanistically. This is due to the complexity of the phenomena leading to this pollution, the heterogeneity of the system characteristics, and the great variation in space and time scale.

Since the parameters of these models either have no physical significance or can not be measured experimentally, their values have to be indirectly determined by fitting simulation results to measured data. This procedure, called model calibration, is an essential stage of model application. It necessitates measured data, criteria of fitness between simulated and measured data and an optimisation algorithm to estimate the optimal parameter values of the models.

Three main problems in developing these models can be underlined:

- First of all, in situ measurement of urban stormwater pollution is difficult and expen-

sive. Field data remains quite rare and of poor quality (uncertainty in the range of 30%). The data available rarely allow a satisfactory calibration and validation of the model (Ahyerre *et al.*, 1998).

- Secondly, there is a lack of knowledge concerning the processes involved, and the modellers tend to make the models more complex. Therefore, it is difficult to put these models into operation (Ashley *et al.*, 1999).
- Furthermore the classical criteria functions often have many local minima, and even if the optimisation algorithms find the parameter set that minimises the criteria function, which is not an easy task since the models are nonlinear, it is difficult to have sufficient confidence about the obtained results (Sorooshian and Gupta, 1983; Kuczera, 1997).

For instance, a study by Gaume *et al.* (1998) on an urban watershed in Quebec City has shown that similarly good fits between recorded data and simulated results can be obtained with very different parameter sets. In other words, large uncertainties in the values of model parameters can remain after calibration.

To improve urban storm water quality models, it seems necessary to acquire more field data, and to develop efficient calibration-validation approaches that not only attempt to identify a “best parameter set” but also help to assess, and if possible to reduce, uncertainties in the parameter values (Beck, 1991).

In this paper, we propose a methodology which allows first of all the identification of model parameters and the assessment of the parameter uncertainties, and secondly calculates the confidence limits of the model to evaluate its predictive capacity which is one of the important objectives of modelling. We will introduce the principle of the Metropolis algorithm from the family of “Monte Carlo Markov Chain Algorithm MCMC”, a robust and efficient algorithm for statistical inference of model parameters. After a description of the studied model and the measured data set, the paper presents the calibration problem using the Metropolis algorithm and the preliminary results and interpretations (Kanso, 2000).

The methodology: Bayesian paradigm

In the last decade, great attention has been given to the Bayesian approach for model calibration in particular in the case of complex hydrologic models (Beven and Binley, 1992; Kuczera and Parent, 1998; Campbell and Fox, 1999). The methodology proposed here promotes a useful calibration procedure during the modelling process that helps to design conceptual models in general and stormwater quality models in particular.

Concept: Bayesian inference

Let Y_t be the set of observed responses at time step t , $t = 1, \dots, n$. The model can be cast as a nonlinear regression model:

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

where $f(\cdot)$ is the output response simulated by the model; X_t is the input data; θ is the vector of model parameters to be estimated from these data and the residuals ε_t are an error term introduced by model uncertainty as well as by the measurement errors. Let γ be a vector of parameters characterising the statistical properties of ε_t . γ is considered, as well as θ , as a set of parameters to be estimated during calibration.

The Bayesian statistical paradigm recognises that there are two sources of information about model parameters: prior information based on historical data and expert knowledge, and data collected by experimentation. The vector of parameters θ is considered as a random variable distributed according to a probability distribution that reflects the uncertain-

ties in the parameters. From a prior distribution $P(\theta)$ the state of knowledge about the parameters is updated using the information in the data $D = \{Y_t, X_t, t=1, \dots, n\}$ to yield a posterior distribution $P(\theta | D)$ according to the Bayes rule,

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

where $P(D)$ 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.

Note that Bayes' theorem does not allow one to derive posterior distribution without prior knowledge. In many applications, in the absence of such prior information, a uniform prior distribution of the parameters is assumed. However, this assumption is not always satisfying to reduce as much as possible the uncertainties, especially when the used calibration data is not sufficient for the identification of the model parameters (Omlin and Reichert, 1999).

The posterior distribution $P(\theta | D)$ contains all the available information about the parameters θ . Bayesian statistical inference therefore reduces to summarising a posterior distribution of θ .

Estimation of the posterior distribution: MCMC method and Metropolis algorithm

In practice, it is difficult if not impossible to summarise the posterior distribution by direct analytical calculation. An approach to approximate the posterior distribution by a multinormal distribution may fail, especially when dealing with complex conceptual model using limited data. The surface of $P(\theta | D)$ mapped in the parameter space can deviate markedly from the surface characterised by the multinormal distribution (Duan *et al.*, 1992).

In recent years, much research has been devoted to the Markov chain sampling method. It represents a general method for sampling from the posterior distribution $P(\theta | D)$. Markov chain sampling generates enough samples from a random walk, which adapts to the true posterior distribution of parameters (Robert and Casella, 1999; Tanner, 1996).

One of the most commonly used MCMC algorithms, the Metropolis algorithm, has received considerable attention in the last decade in the Bayesian statistics literature. This algorithm was chosen because of its simplicity of implementation, efficiency and generality. The basic idea of this chain is to construct a recursive stochastic algorithm in the space of all possible parameter values, which generates from a random walk a sequence of parameter sets $(\theta^i, i=1, \dots, n)$ that converges to a stationary distribution of the Markov chain.

To construct this random walk we define a transition probability function describing the move $\theta^i \rightarrow \theta^{i+1}$ such that the parameter values obtained by the chain converge in distribution to the posterior (Tanner, 1996). There is typically an initial unstable transient phase before reaching the limit distribution. The parameter sets obtained during this transient phase are discarded and the remainders constitute a dependent sample from the posterior distribution.

Model assessment

To evaluate the distribution of model responses, modellers have abandoned traditional statistical inference (like the first order approximation) in favour of more general Monte Carlo simulation techniques (Spear and Hornberger, 1980; Beven and Binley, 1992). The propagation of the posterior parameters uncertainties with Monte Carlo procedures through the model to obtain the range of possible responses gives an indication of both the real predictive power of the calibrated model and its capacity to reproduce the system processes.

Unlike traditional statistic theory, this method can readily cope with nonlinearity of the

model. However, it relies on the ability to sample from the posterior distribution of the model parameter.

To explore the advantages of the Metropolis algorithm for model evaluation a case study which considers a simplified urban storm water pollution model illustrates how to use this technique for model calibration and prediction capacity tests.

Case study

The stormwater pollution model

The model used in this study is a very classical one and describes both the particulate pollutant's erosion during the storm event and their accumulation on the watershed during the preceding dry weather period.

The accumulation of pollutants is assumed to follow an asymptotic behaviour that depends on two parameters (Eq. (3)): the accumulation rate D_{accu} supposed to be linear and independent of the mass accumulated, and the erosion rate D_{ero} , induced by various phenomena like wind effect or street sweeping, proportional to the accumulated mass (Alley and Smith, 1981).

$$\frac{dMa(t)}{dt} = D_{\text{accu}} \cdot S_{\text{imp}} - D_{\text{ero}} \cdot Ma(t) \quad (3)$$

where $Ma(t)$ (kg) is the available mass of pollutants at time t and S_{imp} (ha) is the impervious area.

During the storm event, runoff is supposed to erode the mass of pollutants accumulated on the watershed. Eq. (4) represents the evolution with time of the available pollutant mass. It is supposed that the eroded mass is proportional to the available mass and to the discharge. The pollutant concentration is calculated at each time step depending on the mass of pollutants eroded during this time step. The erosion model depends on two parameters: the erosion coefficient W_{ero} and a coefficient w (Huber *et al.*, 1981):

$$C(t) = \frac{1}{q(t)} \cdot \frac{dMa(t)}{dt} \quad \text{and} \quad \frac{dMa(t)}{dt} = -W_{\text{ero}} \cdot q(t)^w \cdot Ma(t) \quad (4)$$

where $C(t)$ (mg/L) is the pollutant concentration produced by erosion and $q(t)$ (m³/s) is the discharge at the outlet of the watershed at time t .

This model is a lumped conceptual one; Eqs (3) and (4) take into account both surface and in-sewer accumulation and erosion processes.

The data

The used rain event database covers a continuous period of 16 months (1996–1997) with 151 rain events. These data were acquired on the experimental urban watershed “Le Marais” in the centre of Paris (Gromaire, 1998). The 42 ha watershed (91% imperviousness) is drained by a combined sewer system. Suspended solid pollutographs were measured for 40 rain events with various characteristics (Table 1) at the outlet of the combined sewer, and for 13 rain events at a street gully collecting discharges from a 186 m² street surface. The model has been applied to both “Le Marais” watershed (WS1) and to the street watershed (WS2).

As this paper is focused on the calibration of the water quality model, we did not use an hydraulic model to estimate the discharge $q(t)$ from the rain intensity, but we used directly the discharge measured at the catchment outlets. For WS1, the hourly fluctuations of the sanitary flow quantity and quality have been taken into account (Gromaire, 1998).

The data used for the calibration phase for WS1 are the measured pollutographs at the watershed outlet for 26 rain events, corresponding to 117 measurements of suspended

Table 1 Characteristics of storm events (Gromaire, 1998)

| | Total rainfall (mm) | Mean intensity (mm/h) | I_{\max} over 5 min I (mm/h) | Duration (hh:min) | Duration of dry weather (day) |
|---------|------------------------|--------------------------|-------------------------------------|----------------------|----------------------------------|
| Minimum | 2.7 | 1.1 | 2.6 | 00:30 | 0.1 |
| Maximum | 21.6 | 24.0 | 80.8 | 06:56 | 50.5 |
| Mean | 8.6 | 3.7 | 14.8 | 02:00 | 3.0 |

solids concentration. The remaining events are used for validation. For WS2, the measured pollutographs available for 8 rain events are used for the calibration phase. We have to notice that this quantity of calibration data is far superior to the quantity usually used in operational model applications. Indeed, owing to the cost and the difficulty of in-sewer measurements, stormwater quality models are often calibrated for fewer than five events.

Simulation procedure

Simulations were performed for two different initial mass conditions: (i) continuous simulation where the residual mass after each rain event is used to estimate the accumulated mass for the following event, and (ii) null residual mass simulation where there does not remain any more “erodable” stock after each rain event.

ε_t are assumed to be independent and normally distributed $N(0, \sigma^2)$. In this case, the likelihood function can be written in the multiplicative form:

$$P(D|\theta) = \prod_{t=1}^n \frac{1}{(2 \cdot \pi \cdot \sigma^2)^{1/2}} \cdot e^{-\frac{(Y_t - f(X_t, \theta))^2}{2 \cdot \sigma^2}} \quad (5)$$

A uniform distribution is assumed to encode the prior knowledge about the parameters. 10,000 simulations were performed with the Metropolis algorithm, and the first 1,000 samples generated were discarded in order to obtain finally the posterior distribution for each parameter. Thus, the Metropolis algorithm is computationally very intensive.

Results

Several simulations were performed with different initial parameter sets. The Metropolis algorithm converged successfully to the same posterior probability distribution of the parameters whatever the initial parameter set used.

Figure 1 presents the posterior distribution obtained for parameter D_{accu} with the Metropolis algorithm for WS1 and WS2 and for the two different initial mass conditions. For WS1, both initial mass conditions lead to a unimodal distribution of D_{accu} , with a clear identified maximum. Similar distributions are obtained for the other parameters.

However, for WS2, with the null residual mass condition, the Metropolis algorithm indicates clearly the presence of multiple optima in the likelihood function, which would not have been so straightforward with classical optimisation techniques. Let us also notice that the parameter distributions obtained with the Metropolis algorithm are not Gaussian, which is the hypothesis of the classical parameter uncertainty estimation methods based on first order approximations.

The probability distributions of the parameters for the two initial mass conditions differ significantly and in particular those of the dry weather parameters. Table 2 presents the obtained optimal parameter values for WS1 and for the two initial mass conditions. One can notice that even if the continuous simulation results are better than the ones of the null residual mass simulation, the corresponding variance of the errors $\sigma_{\text{opt}} = 126 \text{ mg/L}$

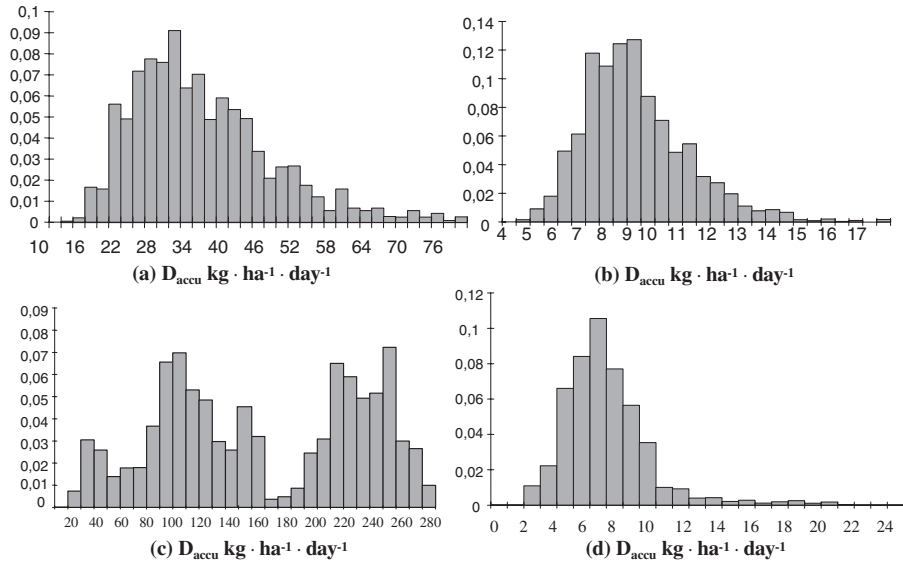


Figure 1 Probability distribution of the accumulation parameter for respectively null residual mass simulation and continuous simulation conditions at (a, b) WS1, (c, d) WS2

Table 2 Optimal parameter values for WS1

| | D_{accu} (kg/ha/day) | D_{ero} (day ⁻¹) | W_{ero} | w | σ (mg/L) |
|-------------------------------|-------------------------------|---------------------------------------|------------------|-----|-----------------|
| Continuous simulation | 7.15 | 0.098 | 0.049 | 1.3 | 126 |
| Null residual mass simulation | 25.6 | 0.38 | 0.073 | 1.2 | 145 |

obtained for the optimal parameter set is quite large compared to the variance of the data ($\sigma_{\text{data}} = 150$ mg/L). Obviously, none of the proposed models seems to be able to reproduce accurately the measured pollutographs, and the Metropolis results indicate clearly that it is not due to calibration problems.

The analysis of the posterior distributions of the parameters indicates, furthermore, a strong correlation between the values of D_{accu} and D_{ero} as shown in Figure 2(a) (correlation = 0.88 in the WS1 case). Unlike D_{ero} and D_{accu} , the value $D_{\text{accu}}/D_{\text{ero}}$, which is the maximum mass that can be accumulated during a dry period, is relatively stable (variation coefficient = 0.125 in the WS1 case). The pollutants accumulation model would probably be more easily calibrated if mathematically reformulated: replacing the two parameters D_{ero} and D_{accu} with a maximum accumulated mass and a pollutant accumulation rate.

As illustrated in Figure 2(b), the parameter uncertainties remaining after calibration can be propagated with a Monte Carlo procedure through the model to obtain the range of the possible model responses. The Monte Carlo procedure consists in repeated simulations with parameter values randomly drawn from their posterior distributions. In the present case the range of the possible responses is very large, that is not surprising regarding the variance of the calibrated residuals ($\sigma_{\text{opt}} = 126$ mg/L). In other words, the predictive power of the calibrated model is very low.

Conclusion

This study shows the power of the MCMC approach (Metropolis algorithm) for calibration. It estimates the true posterior probability distribution of parameters that may differ signifi-

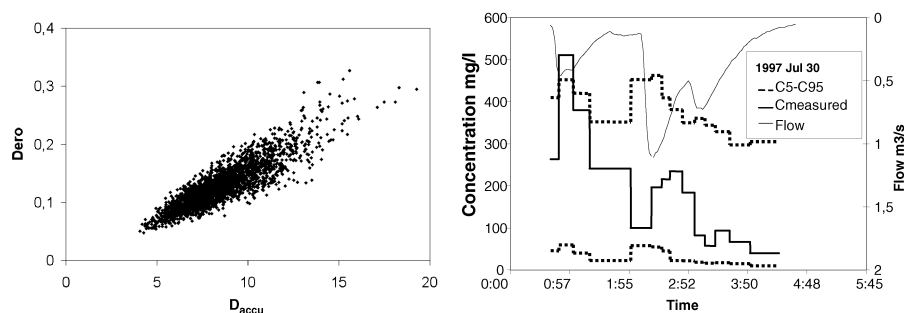


Figure 2 (a) Correlation between dry weather parameters D_{accu} and D_{ero} . (b) The 5/95% confidence interval and measured concentration for a rain event for continuous simulation condition at WS1

cantly from multinormal distributions used in classical parameter uncertainty estimation methods.

An application of the Metropolis algorithm for the calibration of a simplified lumped urban storm water quality model has shown that it produces reliable inferences of the parameters. The proposed model seems unable to reproduce accurately the measured pollutographs, and the Metropolis results indicate clearly that it is not due to calibration problems. Moreover, it has been shown that the predictive power of the studied model is very low.

Furthermore, the analysis of the posterior distributions of the parameters reveals a strong correlation between dry weather parameters, that may cast doubts on the mathematical properties of the pollutants accumulation model. This is an implicit advantage of the Metropolis algorithm, which helps in the improvement of the mathematical concept of model equations.

This behaviour of the calibrated model may be due to the fact that the hypothesis considering the watershed as one entity oversimplifies complex processes or also to the specificity of the site itself.

This result is not optimistic concerning the possibility of using this model to simulate and predict pollutant loads. However, in the literature it was shown that this model could give better results (Gaume *et al.*, 1998). In order to explain this behaviour, further work will compare our results with other studies and especially their respective databases.

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

Acknowledgements

The authors gratefully acknowledge the financial support of the “Réseau Génie Civil et Urbain” (RGC&U), the “Syndicat Interdépartemental pour l’Assainissement de l’Agglomération Parisienne” (SIAAP) and the “National Council for Scientific Research – Lebanon” (CNRSL). Help from Claire Oms and Firas Kanj is also appreciated.

References

- Ahyerre, M., Chebbo, G., Tassin, B. and Gaume, E. (1998). Storm water quality modeling, an ambitious objective?. *Wat. Sci. Tech.*, **37**(1), 205–213.
- Alley, W. and Smith, P. (1981). Estimation of accumulation parameters for urban runoff quality modelling. *Wat. Res. Res.*, **17**(6), 1657–1664.
- Ashley, R.M., Hvitved-Jacobsen, T. and Bertrand-Krajewski, J.-L. (1999). Quo vadis sewer process modelling? *Wat. Sci. Tech.*, **39**(9), 9–22.
- Beck, M.B. (1991). Principles of Modelling. *Wat. Sci. Tech.*, **24**(6), 1–8.
- Beven, K.J. and Binley, A.M. (1992). The future of distributed models: model calibration and predictive uncertainty. *Hydrol. Processes*, **6**, 279–298.

- Campbell, E. and Fox, P. (1999). A Bayesian approach to parameter estimation and pooling in nonlinear flood event models. *Wat. Res. Res.*, **35**(1), 211–220.
- Duan, Q., Sorooshian, S. and Gupta, V. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Wat. Res. Res.*, **28**(4), 1015–1031.
- Gaume, E., Villeneuve, J.-P. and Desbordes, M. (1998). Uncertainty assessment and analysis of the calibrated parameter values of an urban storm water quality model. *J. Hydrol.*, **210**, 38–50.
- Gromaire, M.-C. (1998). *La pollution des eaux pluviales urbaines en réseau d'assainissement unitaire, caractéristiques et origines*. Ph.D. thesis, ENPC, Marne-la-Vallée, France.
- Huber, W., Heaney, J., Nix, S., Dickinson, R. and Polmann, D. (1981). *SWMM3 user's Manual*. Technical report, EPA, USA.
- Kanso, A. (2000). *Calage d'un modèle de pollution du ruissellement urbain par Metropolis*. Master's thesis, ENPC, Marne-la-Vallée, France.
- Kuczera, G. (1997). Efficient subspace probabilistic parameter optimization for catchment models. *Wat. Res. Res.*, **33**(1), 177–185.
- Kuczera, G. and Parent, E. (1998). Monte Carlo assessment of parameter inference in catchments models: The Metropolis algorithm. *J. Hydrol.*, **211**, 69–85.
- Omlin, M. and Reichert, P. (1999). A comparison of techniques for the estimation of model prediction uncertainty. *J. Ecol. Model.*, **115**, 45–59.
- Robert, C.P. and Casella, G. (1999). *Monte Carlo Statistical Methods*. Springer Verlag.
- Sorooshian, S. and Gupta, V. (1983). Automatic calibrations of conceptual rainfall-runoff models: The question of parameter observability and uniqueness. *Wat. Res. Res.*, **19**(1), 260–268.
- Spear, R.C. and Hornberger, G.M. (1980). Eutrophication in Peel Inlet, II, identification of critical uncertainties via generalized sensitivity analysis. *Wat. Res.*, **14**, 43–49.
- Tanner, M. (1996). *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*. Springer Verlag.