A metamodel for stormwater detention basins design

T. Leviandier and S. Payraudeau

Centre d'Ecologie végétale et d'hydrologie, Unité mixte Université Louis Pasteur- ENGEES 1 quai Koch, 67 Strasbourg, France (E-mail: *thierry.leviandier@engees.u-strasbg.fr*; *sylvain.payraudeau@engees.u-strasbg.fr*)

Abstract The aim of storm basins is to protect urban areas against some predefined risk of exceeding a given value of downstream runoff, or a risk of overflow for a bounded storage capacity. This risk results from the combination of a natural hazard and hydraulic properties. The correct way to address this issue is to use a stochastic rainfall model, but it may require unavailable data and be cumbersome to use in the framework of an optimisation procedure. We give the end user a way to by-pass this step, by means of a metamodel. The problem is to calculate the parameters of the probability density function (pdf) of outputs as a function of the pdf of inputs and of the parameters of the dynamic deterministic system between inputs and outputs. We propose to apply a metamodel, which is a new way of designing approximate but generic derived distribution, based on conditional probabilities. For application to dimensioning of basins, the determination of the parameter(s) corresponding to an acceptable risk simply consists of solving an algebraic equation representing the metamodel. The methodology needs usual rainfall statistics and a specific parameter inferred from analysis of storms, or supposed to have a regional value.

Keywords Derived distribution; metamodel; stormwater detention

Water Science & Technology **Vol 56 No 12 pp 37–44** © IWA Publishing 2007

Introduction

The design of stormwater detention with a given risk of overflow of storage capacity is clearly related to the statistical properties of rainfall. As the storage is intrinsically cumulative, a large number of storms must be used to encompass the variability of hyetographs. As shown by Guo (2001), the design storm approach based on a unique storm carries large uncertainties. If on-site data are available, simulation of storm detention basins with different geometrical or hydraulic parameters on observed hyetographs enables an acceptable or optimal solution to be found (Phillips, 1995; Herrmann and Schmida, 1999; Vaes and Berlamont, 2001). However, generally, data are not available and only statistics of rainfall can be inferred from nearby and regional meteorological information. We want to provide a method usable in such a context, but consider that we can use models and data for designing the method that the end user will not have to use. In this case the derived probability distribution approach (Chen and Adams, 2005) is more convenient.

Methods

Two variants of hydraulic behaviour

It is generally assumed that the basin is emptying while storing (otherwise, the problem would be simple, but the efficiency dubious!), either with a constant discharge (with a pump) or in a passive way by gravity. In the latter case, the discharge is a function of the current level of storage, and the basin performs mitigation of a flood. In this paper we assume a linear relationship (only an approximation from a physical point of view). We do not simulate what happens when the storage or hydraulic capacities are saturated, but allow unrealistic simulations run and count such events.

doi: 10.2166/wst.2007.763

Two ways of introducing rainfall information

The variation of the level of the basin can be simulated on real events observed in the past or on a set of events drawn from a stochastic model. When the process is simple, there is no reason to choose a unique reference event, and using a large set is preferable. The second method can be subdivided into Monte-Carlo simulation or numerical discretisation of the probability space, less common but used in the present study.

The volumes design method

This method (Ministères, 1977; quoted in Chocat, 1997) is used to design dry stormwater detention basins. This method is based on two main hypotheses. First, the basin discharge is supposed to be constant and is usually expressed in mm per hour or litres per second per hectare. Secondly, the transfer of rainfall to the detention basin is supposed to be instantaneous without lag time. These hypotheses constrain the application of this method to small urban catchments less than some dozen hectares, without other upstream detention basins. This method requires the availability of approximately ten years of rainfall data with short time steps and can be applied with three steps.

First, each rainfall event has to be differentiated in the rainfall series on the basis of dry period thresholds with the next event. Secondly, for each event the cumulated hyetograph is performed and compared to the cumulated basin discharge to calculate the maximum level in the detention basin during this event. Applied to all the rainfall events of the series, the annual maximum levels in the detention basin are extracted and fitted to a well known probability density function (pdf). The choice of the pdf depends of the goodness of fit between observed data and the statistical model used. The Gumbel law, frequently fitted to extreme values, usually allows a good fit with the maximum annual levels of detention basin. Exponential distributions may be preferred for data sampled as peaks over thresholds, which may be convenient to analyse several storms in the same year. Thirdly, the statistical model, i.e. the Gumbel law for example, could be used to calculate the maximum level in the detention basin required to avoid failure of the basin according to a return period, 10-years for example. This maximum level associated to a return period allows the design of the detention basin by taking into account the upstream area collected.

Insight into the stochastic process with conditional probabilities

Usually, a stochastic model is built with independent random variables and exact relationships between dependent and independent variables. On the other hand, multivariate probability distribution functions, more convenient for explicit calculation, have conditional statistical relationships but these relationships do not easily represent the dynamics and causality between variables. We simultaneously use one model of each type in order to combine their advantages. The multivariate model (we will call it the conditional model) is defined by three properties:

1. the independent variable (R Rainfall) is exponentially distributed;
2. the conditional variable (output of the system) under conditions where rainfall is greater than some value, is exponentially distributed;
3. the median value of the conditional variable is linear with respect to the reduced independent variate.

With notation Q or q always taken for the output variable, which is a discharge in the case of linear outflow, but represents the maximum storage reached during the storm in the case of constant discharge, the equations are:

$$\text{for the conditional distribution}: q_R = q_0 + \lambda u_{qR} + \mu u_R \quad (1)$$

$$\text{with } \mathrm{Prob}(Q > q_R | R > r) = \mathrm{e}^{-u}{}_{qR}$$

$$\mathrm{Prob}(R > r) = \mathrm{e}^{-u}{}_R$$

and parameters q_0, λ, μ and by analytical integration, for the cumulative distribution function (cdf) of q:

$$(\mu - \lambda)e^{-u^*} = z\mu\, \mu e^{-u_\mu} - z\lambda\, \lambda \mathrm{e}^{-u_\mu} \quad (2)$$

$$\text{with} \quad \lambda u_\lambda = \mu u_\mu = q - q_0 \quad (3)$$

$$z\mu = 2^{-(\lambda/\mu)} \qquad z\lambda = 2^{-1}$$

$$\text{and} \quad \mathrm{Prob}(Q > q) = \mathrm{e}^{-u^*} \quad (4)$$

we use the notation with $*$ for this model, so that (2) is true by definition, whereas (4) is only an approximation if Q belongs to another model. Variables named u are taken for variables representing frequency, whether or not they are strictly speaking reduced variates, which is true when the random variable follows an exponential distribution.

The stochastic model is defined as the CECP (contrast enhancing clustering process, Leviandier *et al.*, 2000) model as an input of a simple deterministic model representing the hydraulic behaviour of the basin. The hypotheses of CECP are:

- the total amount of rainfall is exponentially distributed;
- real hyetographs are successively separated in two unequal intervals of constant intensity, such that the rainfall depth of each interval is kept equal to its true value during this span of time, and such that intensities and duration of intervals are linked by a "splitting equation";
- the stochastic process is given by the probability distribution function of the time of separation, at each order (order of successive separation), which is a power law.

The stochastic process is characterised by the exponent of this power law g^i at each order of separation:

$$\frac{t}{d} = F^{g^i}$$

where t is the time of splitting an interval of total duration d, and F is the cdf of t.

In opposition to an earlier paper, in which the exponent was the same at each order (some scale invariance), the power law is downgraded to a uniform distribution at order i greater than one, so that the stochastic process is defined by the unique parameter g_1 with a difference of 1, characteristic of the station, and by the two parameters of the pdf of the total rainfall for the maximum duration taken into account.

There is no theoretical equivalence between these models; in other words, the conditional model does not delivers the exact "derived distribution" derived from the combination of the stochastic and deterministic models, so that we have to check empirically if one can be considered as an approximation of the other, that is, if the conditional probabilities generated by the stochastic model satisfy the properties of the multivariate model. To be useful, the approximation must hold for a range of parameters. In other words, the models, despite the differences in their concepts, should have the same parameters. The approximation must also be validated for large return periods, and the models should have the same asymptotic behaviour. It must be noticed that the CECP model, though it provides no theoretical analytical expression of asymptotic behaviour, is very convenient to calculate conditional distribution, due to the separation of the shape of the hyetograph (given by the times of separation) from the total rainfall.

Approximation of the pdf of output variables as functions of the parameters of the system, through conditional probabilities

The conditional asymptotic probabilities of variables of interest were calculated numerically, running the stochastic model for different values of the parameters of the system. We call parameters of the system the parameters of the stochastic model of rainfall and the parameters of the detention basin. Linear regressions were calculated between parameters of conditional asymptotic probabilities, that is q_0, λ, μ of equation (1) and parameters of the system.

If the models were identical, the only thing to do would be to apply equation (2).

Due to imperfect linearity of conditional probabilities, especially in low return periods. It is necessary to fit a correction formula between the two models:

we use the notation $\mathrm{Prob}(Q > q) = \mathrm{e}^{-u}$for the stochastic model (5)

A possible simple correction is a linear relationship between reduced variates

$$u = au^* + \beta \tag{6}$$

Another thing is to consider that equation (2) is exactly satisfied with $\mathrm{Prob}(Q > q)$ on the left-hand side, but that equation (3) contributing to the right hand side, defines $\tilde{q}$, and to seek a transformation of $\tilde{q}$ into q. A relationship between logarithms was tried:

$$\mathrm{Ln}(q) = \gamma \mathrm{Ln}(\tilde{q}) + \delta \tag{7}$$

Linear regressions were also calculated between α, β and the parameters of the system.

The set of linear regressions of q_0, λ, μ, α, β on the parameters of the stochastic and "hydraulic" model, together with equation 2, is a metamodel.

Design of geometrical or hydraulic parameters by an invert method

For an application, the end user is supposed to know the characteristics of rainfall. The problem is to chose a pair (storage capacity, emptying facility), with often a constraint on one of these items, with a given frequency of overflow. He has not always the resources to run a model and may prefer a formula (Loganathan *et al.*, 1985) or a graphical method (Ministères, 1977; McEnroe, 1992).

For instance, in the case of linear outflow, he wants the outflow not to exceed a value q, with a probability F. Let us write $f(\lambda, \mu, q_0, q)$ on the right-hand side of equation (2):

- If transformation (6) has been used, we have to eliminate $u*$ between $(\lambda - \mu)\ \mathrm{e}^{-u*} = f(\lambda, \mu, q_0, q)$ and $-\mathrm{Ln}(F) = u = \alpha\ u* + \beta$ which consists of finding the root of $\mathrm{Ln}(F) + \beta - \alpha\ (\mathrm{Ln}(f) - \mathrm{Ln}(\lambda - \mu)) = 0$ for the only variable which is the basin discharge parameter b appearing in the linear regressions above-mentioned, the numerical values being given in next section.
- If transformation (7) has been used, we have to eliminate $\tilde{q}$ and find the root of $(\lambda - \mu)\ F - f(\lambda, \mu, q_0, \tilde{q}(q, \gamma\ \delta)) = 0$.

For a given station, the results can be presented in a two-dimension graph, with a curve for each frequency in axis(b,q) or a curve for each b in axis (frequency, q).

Results

Relevance of linear conditional dependence

Figure 1 shows that the conditional distribution of Q (represented here for 10 years).

Return period rainfall, and different durations) is asymptotic exponential and that the median conditional value of Q is linear with the rainfall reduced variate. As a consequence, the parameters λ and μ can be estimated as slopes of the approximate straight lines drawn on Figures 1 and 2.

Validation of the approximations by the conditional model

The model must be validated on a range of statistical parameters encountered on observed data. Several datasets were used with extraction of rainfall during 2 h exceeding a threshold of 10 mm in Northern France or 20 mm in Mediterranean regions). This threshold gives from two to three events per year. Under these conditions the scale parameter of the exponential distribution is never far from the half of the position parameter, and the exponent g_1 varies from 0.3 near Montpellier to 0.6 in Nancy (south and east of France).

The second step is to check the assumption that equation (6) or (7) is able to transform the pdf calculated by the theoretical conditional model to the pdf calculated by numerical

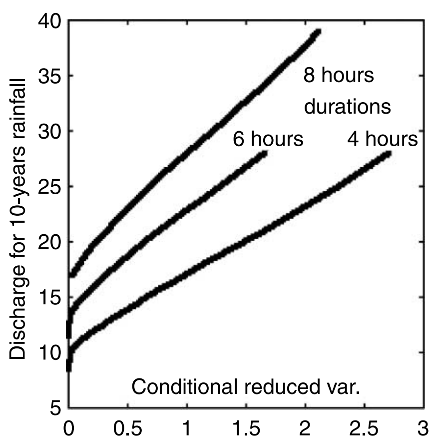

Figure 1 Conditional discharge with exponential reduced variate

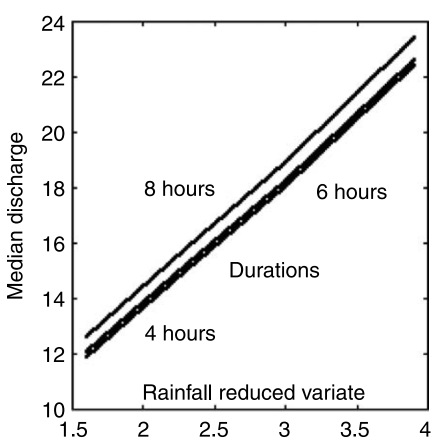

Figure 2 Median discharge in function of the rainfall frequency

integration of the CECP model. Though deep reasons are unknown, the linear correlation between the two values is very strong, with a coefficient R^2 always greater than 0.99 and often greater than 0.999.

The next step is to find relationships between parameters of the system and parameters of the conditional model. Linear regressions were calculated for an arbitrary value $p_0 = 10$ mm in the distribution of rainfall $p = p_0 + u_R g$. The coefficients R^2 of these regressions are greater than 0.99 for *q10* and λ but only about 0.9 for μ, α and β.

The full equations of linear regression (not given here) are used as initial sets of parameters for a non-linear optimisation of the full model, i.e. a minimisation of the difference between the matching reduced variates.

The most efficient correction was found to be on *u* in the first case (constant outflow) and on Ln(*q*) in the second (linear outflow).

Eventually, we find, with *b* in mm, being the depth of water, that can be filled out at constant rate during the storm.

$$\begin{pmatrix} q_{10} \\ \lambda \\ \mu \\ \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 10.03 & -0.362 & -1.823 & 1.519 \\ -5.819 & -0.104 & 2.341 & 1.716 \\ 3.38 & 0.048 & -0.238 & 0.535 \\ 0.883 & -0.001 & 0.404 & -0.005 \\ -1.1 & -0.002 & -0.715 & 0.117 \end{pmatrix} \begin{pmatrix} 1 \\ b \\ \gamma \\ g \end{pmatrix} \tag{8}$$

and for the linear outflow model, *c* representing the coefficient in the linear equation:

$$\begin{pmatrix} q_{10} \\ \lambda \\ \mu \\ \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 10.71 & -0.015 & -0.535 & 1.999 \\ -1.155 & 0.380 & 2.585 & 1.519 \\ -6.316 & -0.028 & -0.479 & 3.093 \\ 2.049 & 0.051 & 0.512 & -0.065 \\ -1.003 & -0.058 & -0.685 & 0.144 \end{pmatrix} \begin{pmatrix} 1 \\ c \\ \gamma \\ g \end{pmatrix} \tag{9}$$

The final validation is given by the graph of the two calculations of *u*, after transformation (Figures 3 and 4.).

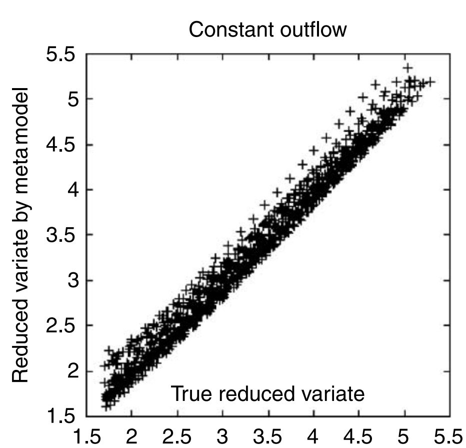

Figure 3 Comparison of model and metamodel for constant outflow

T. Leviandier and S. Payraudeau

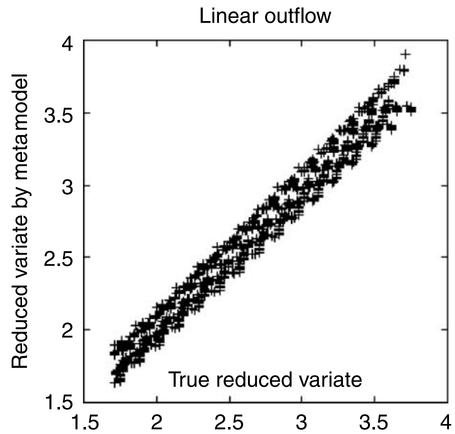

Figure 4 Comparison of model and metamodel for linear outflow

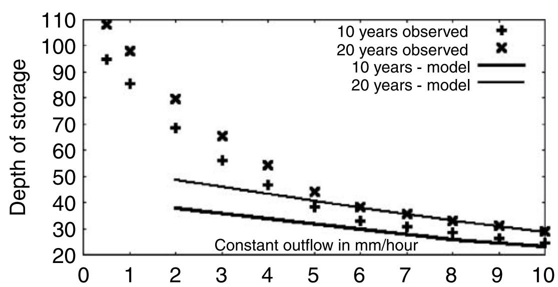

Figure 5 Calculation of critical volume by simulation of all events (points), and by the conditional model (lines) for 10 and 20 years return period as a function of basins discharge in mm/hour

The potential upper limit of efficiency is given by the efficiency of equation (6) or (7) for one set of parameters, which is very high. The errors are due to the assimilation of functional relationships to linear relationships. For an application to a real case, with $p_0 \neq 10$, parameters g and b or c must be multiplied by $10/p_0$ before application of equations 8 or 9 and the output q must be multiplied by $p_0/10$.

Validation with observed hyetographs

An example is given for a Mediterranean station (Mauguio Airport, near Montpellier). With the choice of 2 h and a threshold of 20 mm in the modelling approach, we miss only one long event 1 h which is the worst in its year, yet only for small basin discharges. However, the metamodel method underestimates maximum volumes for low basin discharges and low return periods. For this reason, if results in this range of values are expected, it is necessary to take the envelop of the derived method's results for different durations (using the same equations, in a dimensionless form). After this improvement, the metamodel's results are still lesser then those of the direct method for small basin discharge, but satisfactory for values of practical interest (Figure 5).

Conclusions and prospects

A methodology was presented to design stormwater detention basins with a given probability of overflow. It provides a very simple tool for the end user, the relationship between the free parameter of the device and the accepted risk being embedded in an algebraic formula.

Three parameters are used for the description of rainfall stochastic properties. Two of them are common since they are the first two moments of the distribution of rainfall observed in a constant duration. The third one characterises intensity-duration-frequency relationship, but in a formulation which is specific of the method. It should be considered as regional when there are no data to analyse hyetographs.

The accuracy of the method could be improved by refining the relationships of the parameters of conditional distributions with parameters of the system. However, to keep the method generic and not too much dependent on some details of the system is preferable. Relaxing the constraint of analytical solution is an advantage of the metamodel versus exact derived distribution. Though the conditional probabilities were calculated with a particular stochastic model, we believe they have some intrinsic features. The underestimation for low outflow could be due to the fact that the model was used in separating the events in only three sub-events, or more probably to the assumption that the storage is empty at the beginning of events, what can be corrected. The simple hydraulic behaviour (in power 0 or 1) assumed in this paper brings some simplification, but ongoing researches allow a possible generalisation to upstream hydrological rainfall-runoff transformation with non-linearities.

As the method takes into account the time variation of rainfall intensity, generalisation to sediment and pollutant modelling is also possible, though criteria for sizing storm water quality management facilities may be different (Heitz *et al.,* 2000).

References

Chen, J.Y. and Adams, B.J. (2005). Analysis of storage facilities for urban stormwater quantity control. *Adv. Water Res.*, **28**(4), 377–392.

Chocat, B. (coord.) Encyclopédie de l'hydrologie urbaine et de l'assainissement. (*Encyclopedia of Urban Hydraulics and waste waters management*) 1997. Technique et documentation. Paris. 1121p.

Guo, Y.P. (2001). Hydrological design of urban flood control detention ponds. *J. Hydrol. Engng.*, **6**(6), 472–479.

Heitz, L.F., Khosrowpanah, S. and Nemson, J. (2000). Sizing of surface water runoff detention ponds for water quality improvement. *J. Am. Water Res. Assoc.*, **36**(3), 541–548.

Herrmann, T. and Schmida, U. (1999). Rainwater utilisation in Germany: efficiency, dimensioning, hydraulic and environmenal aspects. *Urban Water*, **1**, 307–316.

Leviandier, T., Lavabre, J. and Arnaud, P. (2000). Rainfall contrast enhancing clustering processes and flood analysis. *J. Hydrol.*, **240**(1/2), 62–79.

Loganathan, G.V., Delleur, J.W. and Segarra, R.I. (1985). Planning Detention Storage for Stormwater management. *J. Water Res. Plan. Manage., ASCE*, **111**(4), 382–398.

McEnroe, B.M. (1992). Preliminary Sizing of Detention Reservoirs to reduce Peak Discharges. *J. Hydraul. Engng, ASCE*, **118**(11), 1540–1549.

Ministères. (1977). Instruction technique relative aux réseaux d'assainissement des agglomérations (*Technical instructions for sewer systems in towns*): IT 77 284 INT Imprimerie Nationale. Paris. 62p + annexes

Phillips, D.I. (1995). A generic method of design of on-site stormwater detention storages. *Water Sci. Technol.*, **32**(1), 93–97.

Vaes, G. and Berlamont, J. (2001). The effect of rainwater storage tanks on design storms. *Urban Water*, **3**, 303–307.