Weather Radar Information and Distributed Hydrological Modelling (Proceedings of symposium HS03 held during IUGG2003 at Sapporo, July 2003). IAHS Publ. no. 282, 2003.

Propagation of precipitation uncertainty through a flood forecasting model

SHREEDHAR MASKEY
IHE-Delft, PO Box 3015, 2601 DA Delft, The Netherlands
maske@ihe.nl

VINCENT GUINOT
Maison des Sciences de l'Eau, Université Montpellier 2, F-34095 Montpellier Cedex 5, France

ROLAND K. PRICE
IHE-Delft, PO Box 3015, 2601 DA, Delft, The Netherlands

Abstract Uncertainty in the precipitation forecast is one of the major sources of uncertainty in real-time flood forecasting. Precipitation uncertainty generally consists of uncertainty in: (a) the quantity, (b) the temporal distribution, and (c) the spatial distribution of precipitation. This paper presents a methodology for propagating the three forms of precipitation uncertainty through a rainfall–runoff routing type deterministic flood forecasting model. It uses fuzzy set theory combined with a genetic algorithm. The methodology is particularly useful where the probabilistic forecast of precipitation is not available. The results show that the output uncertainty due to uncertain temporal and spatial distributions of precipitation can be significantly dominant over the uncertainty due to the uncertain quantity of precipitation.

Key words flood forecasting; fuzzy set; genetic algorithm; precipitation; uncertainty

INTRODUCTION

Forecasting a flood using rainfall–runoff routing type models requires the forecasting of precipitation for the forecast period. Over the last decade significant progress has been made in the quantitative forecast of precipitation using sophisticated radar technology. The uncertainty in forecast precipitation remains, however, one of the major sources of uncertainty in real-time flood forecasting. It consists of uncertainty in: (a) the quantity, (b) the temporal distribution over the forecast period, and (c) the spatial distribution over the watershed. This paper presents a methodology for propagating the three forms of precipitation uncertainty through a rainfall–runoff routing type deterministic flood forecasting model. The extension principle of the fuzzy set theory is used to propagate the uncertainty. This principle is implemented applying the so-called α-cut method. A genetic algorithm (GA) is used to determine the maximum and minimum of the model output, which is an essential part of the extension principle by α-cuts. The uncertainty in temporal and spatial distributions of precipitation is treated using the concept of the principal precipitation pattern (PPP). This methodology is independent of the structure of the forecasting model. That is, the methodology can be used with any rainfall–runoff routing type deterministic models. The method is applied to a watershed model of Klodzko (Poland), built with HEC-1 (USACE, 1998) and HEC-HMS (USACE, 2000).

METHODOLOGY

The methodology for the propagation of precipitation uncertainty introduces the notions of a three-valued quantitative precipitation (3-VQP) and a principal precipitation pattern (PPP).

Three-valued quantitative precipitation (3-VQP)

Let P_i represent the total precipitation for a sub-basin i ($i = 1, \ldots, m$; m = number of sub-basins) accumulated during the forecast period (T). P_i is assumed to be obtained from a quantitative precipitation forecast (QPF). Since this value is not absolutely certain, assume it has an error of $\pm\, e$ percent. In the absence of a probabilistic forecast of precipitation, the notion of the 3-VQP is introduced. This is defined as the most credible, minimum and maximum forecast precipitations, where:

The most credible forecast precipitation, $P_{i,mc} = P_i$ (1)

The minimum forecast precipitation, $P_{i,\min} = P_i - eP_i/100$ (2)

The maximum forecast precipitation, $P_{i,\max} = P_i + eP_i/100$ (3)

A triangular membership function (Fig. 1) is constructed from the 3-VQP. The qualitative meaning of the triangular membership is the following. The "true" value of precipitation, P_i, is certainly included between $P_{i,\min}$ and $P_{i,\max}$ and is likely to be close to $P_{i,mc}$. The key words are "included" and "close". These words constitute the only information that one has about the problem (Rivelli & Ridolfi, 2002).

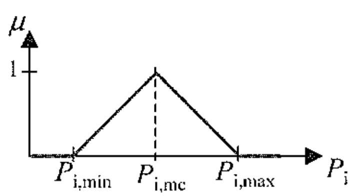

Fig. 1 Fuzzy triangular membership function based on the 3-VQP.

Principal precipitation pattern (PPP)

To take into account the uncertainty in temporal and spatial distribution of precipitation, the notion of the PPP is introduced. A forecast period is divided in to n subperiods where the length of a subperiod, $t = T/n$. Then the PPP is the pattern such that the total precipitation for a sub-basin occurs during one of the subperiods and no precipitation occurs during the remaining subperiods. For example, assuming three subperiods the 3 PPPs (Fig. 2) will be (P_i, 0, 0), (0, P_i, 0) and (0, 0, P_i). It should be noticed however that the PPP is not used directly as inputs to the model. The actual input precipitation pattern will be based on the PPP and randomly generated temporal

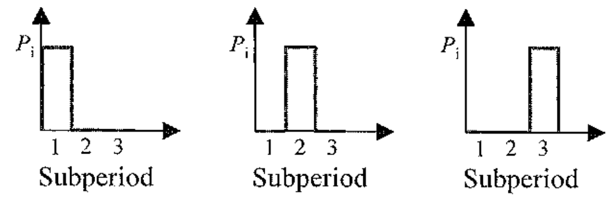

Fig. 2 Principle precipitation patterns (PPPs) considering three subperiods of forecast.

and spatial distribution factors (pattern coefficients). A matrix of temporal and spatial distribution factors, $\boldsymbol{B}$, is introduced such that:

$$\boldsymbol{B} = \begin{bmatrix} b_{1,1} & \cdots & b_{1,n} \\ \vdots & b_{i,j} & \vdots \\ b_{m,1} & \cdots & b_{m,n} \end{bmatrix} \tag{4}$$

where $0 \leq b_{i,j} \leq 1$ (for all i, j) and $b_{i,1} + ... + b_{i,n} = 1$ (for all i), and the matrix of the precipitation pattern, PT, is given by $PT = P_i\boldsymbol{B}$. Thus by generating different values of $b_{i,j}$, as many numbers of precipitation patterns as necessary can be generated. Normally, the temporal distribution is incorporated by considering a fixed number of predefined precipitation patterns. The advantage of the present approach is that there are equal chances that all possible precipitation patterns (the subperiod level) are checked. In the precipitation uncertainty processor (PUP) of Kelly & Krzysztofowicz (2000), the temporal disaggregation of precipitation was treated using the notions of wet subperiods and timing patterns.

Propagation of precipitation uncertainty

Uncertainty represented by a fuzzy membership function can be propagated through a model using the extension principle of fuzzy set theory. First introduced by Zadeh (1975) the extension principle enables us to extend the domain of a function on fuzzy sets. The extension principle is generally carried out at different level α-cuts. An α-cut of a fuzzy set A, denoted as A^α is the set of elements which have memberships greater than or equal to α. For each α-cut two values of the output are required called the lower bound (LB) and the upper bound (UB). The LB is the minimum and the UB is the maximum of the model output. In the present example, representing the model function by f and the model output discharge by Q, the extension principle gives:

$$Q_{LB}\alpha = \min[f([P_{1,LB}, P_{1,UB}]^\alpha, \quad [f([P_{1,LB}, P_{1,UB}]^\alpha, \boldsymbol{B} \quad \alpha \in [0,1] \tag{5}$$

$$Q_{UB}\alpha = \max[f([P_{1,LB}, P_{1,UB}]^\alpha, \quad [f([P_{1,LB}, P_{1,UB}]^\alpha, \boldsymbol{B} \quad \alpha \in [0,1] \tag{6}$$

where Q^α_{LB} and Q^α_{UB} are the lower and upper bounds of the model output discharge, respectively; $P^\alpha_{i,LB}$ and $P^\alpha_{i,UB}$ are the lower and upper bounds of the forecast precipitation (accumulated for the forecast period), respectively, for the given α-cut.

Thus, finding the membership function of an output using the extension principle is a problem of finding the minimum/maximum of functions (equations (5) and (6)) in which the constraints are the bounds in the input variables defined for the given α-cut. The optimization problem is therefore subject to the constraints:

$$\left.\begin{array}{ll} P_{i,LB}^{\alpha} \le P_i^{\alpha} \le P_{i,UB}^{\alpha} & \forall i \\ 0 \le b_{i,j} \le 1 & \forall (i,j) \\ \sum_{j=1}^{n} b_{i,j} = 1 & \forall i \end{array}\right\} \tag{7}$$

A genetic algorithm is used to generate a range of precipitation patterns in order to deduce the lower and upper bounds of the model output discharge.

Algorithms for minimum and maximum determination

Algorithms for solving optimization problems range from linear, nonlinear to global. The global optimization algorithms (GOAs) possess a particular advantage in solving optimization problems using off-the-shelf software, for which the details of the underlying algorithms may not be known (Maskey *et al.*, 2002). One of the most famous GOAs is the genetic algorithm (GA). A comprehensive evaluation of optimization algorithms is not the intent of this study. However, the successful and extensive use of GAs in various fields of engineering, including water-related problems, inspired the authors to choose a GA scheme as an optimizer.

Genetic algorithms are search algorithms based on the mechanics of natural selection and natural genetics (Goldberg, 1989). The version of the GA used in this study consists of the tournament selection, uniform crossover, two children per pair of parents, jump and creep mutations, elitism and niching.

APPLICATION EXAMPLE

Model overview

A flood forecasting model of the Klodzko watershed has been considered. HEC-1 (USACE, 1998) and HEC-HMS (USACE, 2000) by the Hydrologic Engineering Centre of the US Army Corps of Engineers were used to build the model. HEC-HMS (its Calibration Module) was used for model calibration and HEC-1 was used for the simulation. HEC-1 is one of the most frequently used rainfall–runoff models in the United States (Melching *et al.*, 1991). The methods used to represent different processes in the present model are: (a) The Soil Conservation Service (SCS) Curve Number (CN) method for the direct runoff volume (precipitation excess) computation; (b) Clark's unit hydrograph (UH) method for the transformation of excess precipitation to runoff; (c) the exponential recession method for baseflow computation; and (d) the Muskingum method for the flow routing.

A schematic diagram of the model is shown in Fig. 3. The basin consists of nine sub-basins (SBs) and covers the total area of 1744 km^2 with sub-basin areas ranging

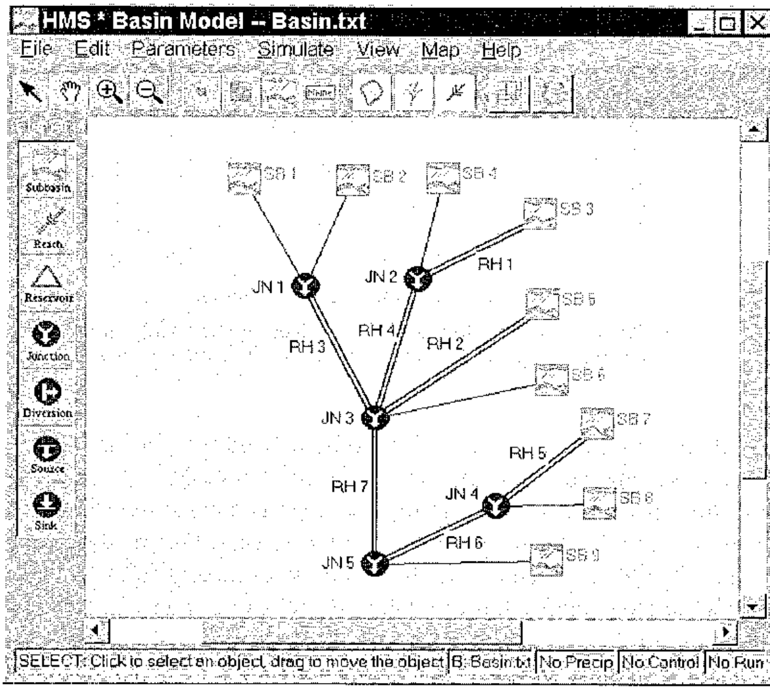

Fig. 3 Schematic diagram of the watershed (SB = sub-basin, RH = reach, and JN = junction).

from 64 km^2 to 280 km^2. It has seven reaches (RHs) and five junctions (JNs). Observed precipitation data (cumulative for every 3 h) are available for the flood of July 1997. Due to the lack of forecast precipitation data, the observed precipitation is assumed as the forecast precipitation. Since the frequency of the available precipitation is 3 h, the size of the forecast period is also taken as 3 h. The period is divided into three subperiods of 1 h each. For a practical application in real time, the 3 h forecast period may be too short. Given the abrupt rise of the flood in a short period, this is fairly appropriate for the application of the methodology. For each forecast, only the uncertainty in the forecast precipitation during the same forecast period is considered. For example, to forecast the flood for the time 6 h, no uncertainty is assumed in the precipitation up to the time 3 h, and so on. An error of 30% is assumed in the given precipitation to estimate the 3-VQP. Although the error is taken arbitrarily, discussion with experts suggested that the assumption is not too far from reality. The starting date for the simulation is 4 July 1997 at 06:00 h, while the forecast of precipitation started on July 7 1997 at 06:00 h.

RESULTS

All results presented and discussed here are based on the forecast downstream of JN 5 (Fig. 3). Firstly, the simulation is carried out to compute the flood discharge without considering the uncertainty. The simulated flood together with the basin-average

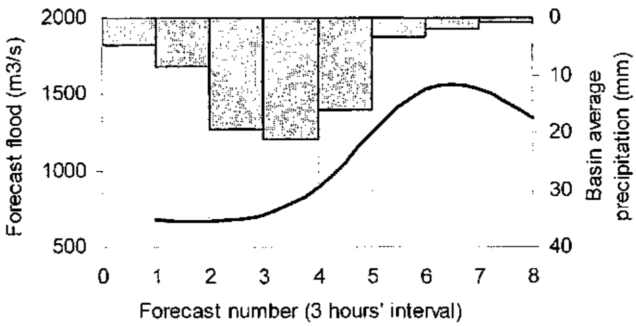

Fig. 4 Simulated flow without considering uncertainty in precipitation with the basin average precipitation. Only a part of the simulation is shown.

precipitation is presented in Fig. 4. It is to be noticed that the simulation was carried out taking the accumulated precipitation in each sub-basin, not with the basin-average precipitation. Secondly, the present methodology is applied to estimate the uncertainty in the forecast flood due to the uncertainty in the forecast precipitation. This is done in two steps: (a) for the uncertainty in the quantity of precipitation only, and (b) for the uncertainty in the quantity of precipitation plus the uncertainty in the temporal and spatial distributions. The former assumes a uniform distribution of precipitation throughout the forecast period. This means that the pattern coefficients $b_{i,1} = b_{i,2} = b_{i,3} = 1/3$. For simplicity, these two cases are referred to as precipitation quantity uncertainty (PQU) and precipitation quantity and pattern uncertainty (PQPU), respectively.

For each forecast, $Q_{UB} - Q_{LB}$ is calculated from the results at a zero level α-cut for both cases: PQU and PQPU. This quantity expresses the maximum amount of estimated uncertainty in the forecasted flood (Fig. 5). It is clearly seen that in all forecasts the output uncertainty is dominated by the uncertainty in the precipitation patterns. As can be seen in Fig. 5, the maximum uncertainty occurred for the fourth and fifth forecasts. The precipitation patterns that give the maximum flow for the fifth forecast is shown in Fig. 6. The results shown in Figs 5 and 6 are for a zero level α-cut only. To construct a complete membership function (MF) of the output, the computation needs to be carried out at different level α-cuts. Four more α-cuts at α = 0.25, 0.5, 0.75 and 1 were selected. The MFs obtained for the fifth forecast due to PQU and PQPU are shown in Fig. 7.

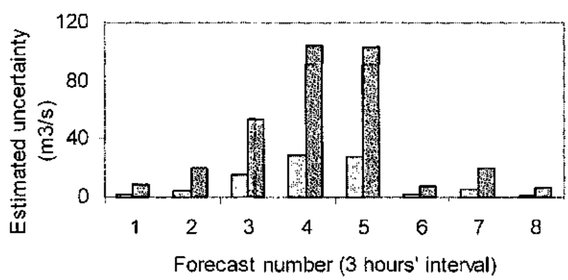

Fig. 5 Estimated uncertainty in the forecast flow due to PQU and PQPU.

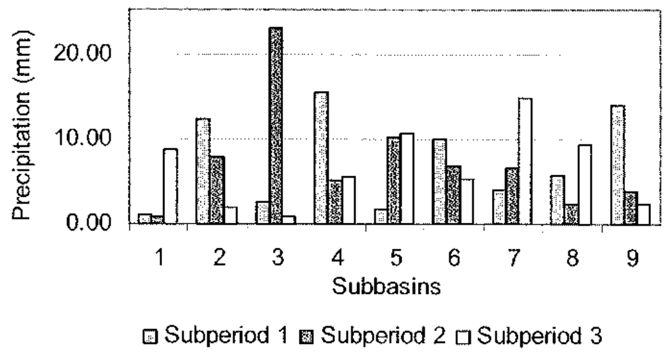

Fig. 6 Precipitation pattern resulting in the maximum flow for the fifth forecast.

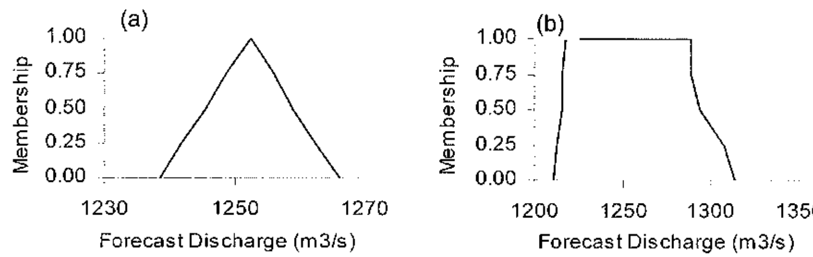

Fig. 7 Membership function of the forecast flow (fifth forecast): (a) due to PQU and (b) due to PQPU.

CONCLUSION

The presented methodology for the propagation of precipitation uncertainty can be very useful in the absence of a probabilistic precipitation forecast. This methodology is independent of the structure of the forecasting model. In other words, it can be used with any rainfall–runoff routing type deterministic model. The methodology uses two important concepts: 3-VQP and PPP. The concept of 3-VQP is used to represent uncertainty in the quantity of forecast precipitation by a fuzzy membership function from limited information about uncertainty. The concept of PPP is used to represent uncertainty in the temporal and spatial distributions of the precipitation. The results show the good potential of the fuzzy set theory combined with a GA for the propagation of uncertainty. The results also show that the output uncertainty due to the uncertainty in the temporal and spatial distributions can be significantly dominant over the uncertainty due to the uncertain in the quantity of precipitation. Therefore, it suggests that we might be making a big mistake by taking basin average precipitation uniformly distributed throughout the forecast period for forecasting, which is generally the case in many forecasting systems. The estimated uncertainty in the output (Fig. 5) may seem small compared to the magnitude of the flood (Fig. 4). This is due to the relatively small forecast period (3 h) considered. Obviously, increasing the forecast period significantly increases the uncertainty in the forecast precipitation and thereby increases the output uncertainty.

Acknowledgements This research was initiated by the EC project Operational Solutions for the Management of Inundation Risks in the Information Society (OSIRIS), contract IST-1999-11598. The authors extend their gratitude to the IMWM, Poland for providing the data used in this study.

REFERENCES

Goldberg, D. E. (1989) *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, Reading, Massachusetts, USA.

Kelly, K. S. & Krzysztofowicz, R. (2000) Precipitation uncertainty processor for probability river stage forecasting. *Water Resour. Res.* **36**(9), 2643–2653.

Maskey, S., Jonoski, A. & Solomatine, D. P. (2002) Groundwater remediation strategy using global optimization algorithms. *J. Water Resour. Plan. Manage. Div. ASCE* **128**(6), 431–440.

Melching, C. S., Yen, B. C. & Wenzel, H.G., Jr (1991) Output reliability as guide for selection of rainfall-runoff models. *J. Water Resour. Plan. Manage. Div. ASCE* **117**(3), 383–398.

Revelli, R. & Ridolfi, L. (2002) Fuzzy approach for analysis of pipe network. *J. Hydraul. Engng* **128**(1), 93–101.

USACE (1998) HEC-1 *Flood Hydrograph Package User's Manual*. Hydrologic Engineering Centre, USACE, Davis, California, USA.

USACE (2000) *Hydrologic Modelling System HEC-HMS Technical Reference Manual*. Hydrologic Engineering Centre, USACE, Davis, California, USA.

Zadeh, L. A. (1975) The concept of a linguistic variable and its application to approximate reasoning. Part I. *Inform. Sci.* **8**, 199–249.