Risk in Water Resources Management (Proceedings of Symposium H03 held during IUGG2011
in Melbourne, Australia, July 2011) (IAHS Publ. 347, 2011).

Imprecise probabilities to specify hydrological loads for flood risk management

ANDREAS H. SCHUMANN[1], **MARKUS PAHLOW**[1], **DAVID NIJSSEN**[1] **&**
BASTIAN KLEIN[2]

1 *Institute of Hydrology, Water Resources Management and Environmental Engineering, Ruhr-University Bochum.*
D-44780 Bochum, Germany
andreas.schumann@rub.de

2 *German Federal Institute of Hydrology, Referat M2, Am Mainzer Tor 1, D-56068 Koblenz, Germany*

Abstract Assessments of the performance and the risk of failures of complex technical flood retention systems demand the specification of hydrological loads under a wide range of possible circumstances. The outcome of risk assessments depends on these pre-assumptions. The existing lack of information demands new approaches to characterise uncertainties. As it is more widely recognised that the concept of uncertainty is too broad to be captured by probability theory alone, the application of imprecise probabilities can be useful in this context. In a case study, a wide range of possible flood events was specified by hydrological models, which combine stochastic and deterministic components. The simulation results were analysed by multivariate statistics. With respect to the uncertainties of simulations the plausibility of the flood scenarios was specified by fuzzy sets. This measure of plausibility was incorporated subsequently in a decision support system as basic characteristic of impact assessments of planning decisions.

Key words water; rivers; groundwater

INTRODUCTION

Risk is an indispensable criterion of flood management planning, since an absolute protection from flooding cannot be reached by technical measures. This demands a more comprehensive view on flood risk. In practice, design floods are characterised by the return period of the peak only. However, the performance of technical flood retention facilities depends on multivariate characteristics of floods, which have to be specified by several coinciding random variables such as flood peak, volume, shape and duration. It can be shown that flood protection may be ensured under favourable flood conditions, but in other cases the system may fail, even if a certain flood characteristic, e.g. the flood peak, remains below the value which was assumed for the design flood. Therefore an ensemble of hydrological loads should be applied to demonstrate under which conditions the performance of the planned flood control system may not suffice and to simulate impacts of possible failures. These hydrological scenarios have to be defined probabilistically to describe the effectiveness of flood control measures in the context of risk management. Multivariate statistics can be applied to calculate the probabilities of the relevant hydrological loads (e.g. De Michele *et al.,* 2005; Klein *et al.*, 2010). However, multivariate statistical characterisations demand large data samples. Hydrological time series are often too short, non-homogenous or non-stationary to provide the information regarding the large range of hydrological loads, which is needed in such analyses. To overcome this problem, stochastic-deterministic flood simulation can be applied. It is based on stochastic generation of precipitation events and transformation of these precipitation fields with a deterministic hydrological model into runoff time series (e.g. Blazkova & Beven, 2004; McMillan & Brasington, 2008; Moretti & Montanari, 2008). Modelling approaches are the only way to specify the specific impacts of flood storages and their interactions for large flood retention systems. However, this methodology implies many uncertainties (Cameron *et al.*, 1999; Lamb & Kay, 2004), particularly if such analyses are carried out for a large river basin with spatially-distributed hydrological loads, where many different combinations of influencing factors are possible. To handle these uncertainties, one can use imprecise probabilities instead of probabilities (Klir, 1999). Fuzzy sets are an option to express the imprecision of probabilities of hydrological loads.

METHODOLOGY

Stochastic-deterministic generation of flood scenarios for large river basins

A comprehensive analysis of the performance of flood control measures in a large river basin has to consider a large variety of different flood events. It is proposed to derive these scenarios from simulations. In flood planning these series are not "real" in the sense of being expected, but they are examples of what may occur. Such flood events are treated as being typical for the conditions under which the flood control system is operated. To represent such a system with a hydrological model, a reservoir module for controlled and uncontrolled flood management has to be implemented that describes the operation of dams and reservoirs within the river basin. Assuming a deterministic behaviour of the river basin and of reservoirs and polders located within the basin, the flood scenarios depend on the spatial and temporal distribution of rainfall. Stochastic rainfall generators can be applied to characterise the stochastic behaviour of point measurements as well as the spatial correlation of rainfall. By coupling a rainfall generator with a hydrological model of the river basin, a large number of floods can be simulated. The bottleneck of flood risk assessments by modelling approaches consists in the flood inundation model. With regard to the high damages which can be expected if settlements are affected, a detailed estimation of the inundated areas is needed. Often 2- or even 3-dimensional hydraulic simulation models have to be applied to provide this information. Due to these computationally expensive requirements a complete stochastic risk estimation, based on a statistically significant sample of several thousand flood simulations, is not feasible. Thus an affordable, yet representative number of flood scenarios must be selected.

Categorising hydrological loads with multivariate statistics

The selection of flood scenarios should comprise the range of possible circumstances. The assumption that flood events with similar peaks but different volumes, or caused by different spatial distributions of precipitation within the basin, will have the same probabilities, is not correct. Thus consideration of joint probabilities of relevant flood properties is necessary for risk-based planning. Copulas can be used for the construction of bivariate distribution functions. In hydrology they have been implemented for multivariate analysis of hydrological random variables (see e.g. Salvadori & De Michele, 2004; Klein *et al.*, 2010). A copula is a function which exactly describes and models the dependence structure between correlated random variables, independently of the marginal distributions. The problem arises: which characteristics should be combined in bivariate distribution functions? Problem-oriented multiple combinations are useful, e.g. if a flood reservoir is endangered by floods with unfavourable relationships between peak and volume, or if the effects of polders located downstream depend on coincidences of flood events in tributaries. For complex flood retention systems, each flood scenario can be specified by a multitude of bivariate probabilities.

Characterisation of hydrological loads with imprecise probabilities

The multivariate characteristics of flood scenarios and their interdependencies can be characterised with several copulas. However, the database for these statistical analyses was derived from complex simulations and is affected by many assumptions. Thus the copula-based statistical information about the design flood scenarios can be considered as imprecise probabilities. Imprecise probabilities are a way to handle uncertainties of probabilistic assessments. The imprecision in expressing probabilities, which was very much stimulated by Walley (1999), introduced a new dimension into the formalisation of uncertainty and uncertainty-based information. Here we consider that events which belong to one set of floods according to the return period of the peak may differ in other characteristics. These differences were analysed with copula statistics. The resulting statistical measures are used as additional information to specify the events in a possibilistic way. There are typical events, where the return period of the peak and the return periods of other characteristics are similar, and less typical events, where these probabilities differ significantly. Fuzzy sets are one way to specify such uncertainties. According to fuzzy theory, the membership of a single event within a set of flood events of a certain return period is specified by a membership function. A flood is seen as a fuzzy event which has a probability measure (return period of the peak) and a degree of membership. The highest value of the membership function $\mu = 1$ is attributed to events where the bivariate copula probabilities are nearly the same as the probability of the flood peak. Such flood events seem to be most representative for a certain return period with regard to the agreement of the different statistical characteristics of the flood. If e.g. the return period which was estimated from the joint probability of peak and volume is greater than the return period of the peak, then the event is less probable than expected from the return period of peak alone. If this concordance between return periods is not given, then the assumed probability of the event derived from the flood peak only seems to be less plausible. To consider these differences, a new characteristic value of “plausibility” P is introduced, which is derived from the differences in probabilities. If, for example, the relationship between flood peak and volume is used as an additional characteristic, then the plausibility of a flood event with a return period of the flood peak T_{Peak} can be derived as follows:

$$P_{Plausibility} = \begin{cases} Min\left(\dfrac{T^{\vee}_{Peak,Volume}}{T_{Peak}}; \dfrac{2 \cdot T_{Peak} - T^{\vee}_{Peak,Volume}}{T^{\vee}_{Peak,Volume}} \right), \forall T^{\vee}_{Peak,Volume} \in \left[0; 2 \cdot T_{Peak}\right] \\ 0, \forall T^{\vee}_{Peak,Volume} \notin \left[0; 2 \cdot T_{Peak}\right] \end{cases}$$

Here T_{Peak} and $T^{\vee}_{Peak,Volume}$ are the return periods based on the flood peak statistics or on the copula statistics of flood peak and volume (applied was the logical “or”). Several measures of plausibility can be used according to the location within the river basin and the relevance of different flood characteristics.

CASE STUDY

The Unstrut River basin

The methodology described above has been applied to the Unstrut River basin in the central part of Germany. The Unstrut watershed has an area of nearly 6400 km^2. The catchment has variable topographic structure, with lower regions in the central part, the Harz Mountains in the north and the Thuringian Forest in the south. The current technical flood retention system within this river basin consists of the reservoir Kelbra and the reservoir Straussfurt, several small reservoirs of local importance; a flood channel and a flood polder system with five polders between the cities of Oldisleben and Wangen (see Fig. 1). In total, the flood retention system has a volume of 100 hm^3. A set of planned flood control measures had to be assessed, varying from the optimisation of the existing polders, increase of retention time within polders by additional check dams, creation of new polders, alteration of the polder inlet structures and variation of the inlet regulation (controlled and uncontrolled flooding). These measures were clustered into six different stages of extension of the flood retention system.

Deterministic and stochastic modelling

Due to the construction of reservoirs since the 1960s, the homogenous runoff time series was available only for 32 years of observation. This short observation length was insufficient to estimate multivariate flood probabilities. Therefore, a stochastic model was used to generate daily time series of rainfall at multiple locations (Hundecha *et al.*, 2009). To simulate a long time series of runoff from the stochastically simulated precipitation sequences, a semi-distributed hydrological model, based on an object-oriented framework and following the concept of HBV-96 (Lindström *et al.*, 1997) has been applied. As the runoff conditions within the simulated river basin were affected by reservoirs, a reservoir module for flood management was integrated. This module allows for consideration of the specific hydraulic conditions of reservoirs with regard to the bottom outlet and the spillway. It has been applied to simulate different operation rules. This hydrological model was used to simulate a long time series of discharge values at several locations within the river basin with daily time steps.

To avoid the difficulties of a stochastic simulation of other meteorological variables which were needed for the simulation of the water balance, here a time series of 40 years with measured daily data of air temperature, humidity and radiation were used. These series were repeated (in total 10 times 1000 years were simulated). Considering the large number of possible combinations of precipitation values with other meteorological variables, a realistic representation of the hydrological conditions can be expected. To demonstrate this, the statistics of observed discharge at various gauges were compared with statistics from the simulated series, as shown in Fig. 2.

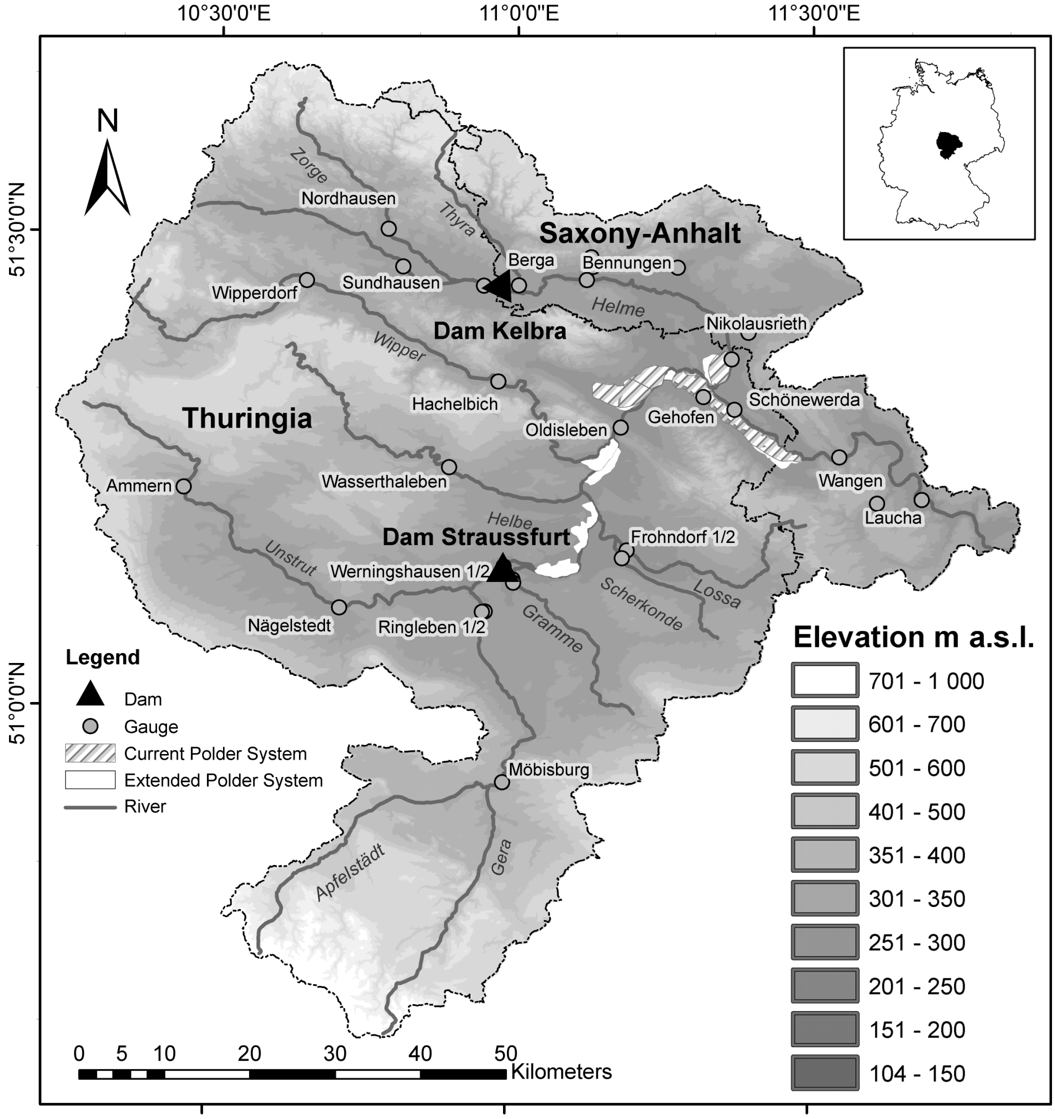

Fig. 1 Topographical map of the Unstrut catchment in the federal states Thuringia and Saxony-Anhalt. Also shown are the technical flood retention system (current and extended) and important gauges within the river basin (Schumann *et al.*, 2010).

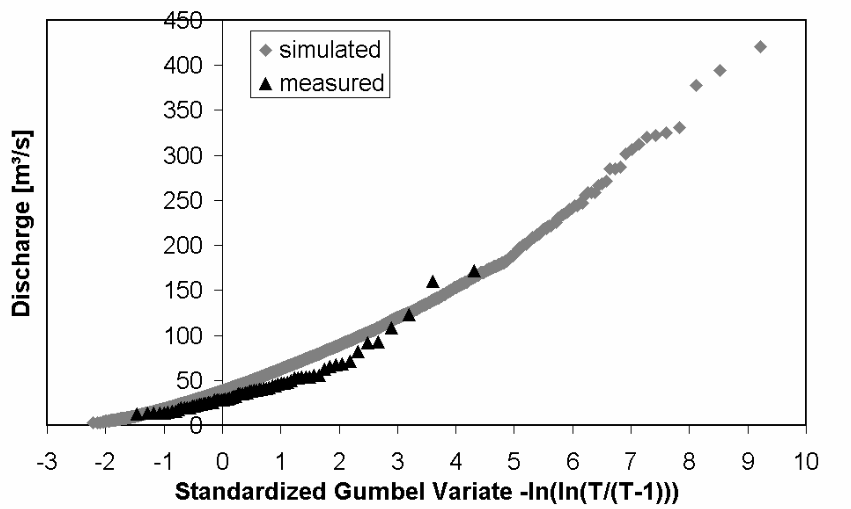

Fig. 2 Comparison of the measured maximum annual daily discharge and the simulated measured maximum annual daily discharge from the coupled daily stochastic rainfall generator and the water balance model at the gauge Moebisburg (Klein *et al.*, 2010).

Selection of flood scenarios

From the simulated 10 000 years of daily runoff data, 30 flood events were selected (five events for each of the six return period classes of $T = 25, 50, 100, 200, 500, 1000$ years) which were used as flood scenarios. The spatial structure and the seasonal variation of historic events which were observed in the past were considered during this selection, but also events that have not been observed, yet seem to be probable were included. These selected events were simulated a second time with an hourly time step to ensure that the flood dynamics and the effects of flood control measures are represented in an appropriate way. To do this, the series of daily precipitation have been disaggregated into hourly time values using the tools HYETOS (Koutsoyiannis *et al.*, 2003) and MuDRain (Koutsoyiannis & Onof, 2001). The return period of the peak was chosen as a basic characteristic of all flood events. The selected events differ in their shape, their volume and the spatial distribution of runoff. The probabilities of these multivariate characteristics were assessed by multivariate statistics. For each of the two reservoirs, the flood control depends on flood peak and flood volume. Therefore, the joint return periods of the values of the annual flood peaks and the corresponding volumes of the inflows to the dams were used to categorise the hydrological scenarios. The performance of the polder system downstream depends on the coincidence of the floods from the two main sub-basins. Here the copula method was used to express the bivariate distribution function of the resulting flood after the point of confluence.

Application of the selected flood scenarios

After level pool routing in the two reservoirs, the propagation of flood waves along the river course was simulated with a coupled 1-D/2-D hydraulic model, which was capable of considering the existing and planned polders (Kamrath *et al.*, 2006). For 180 events (30 hydrological scenarios and six different states of the flood control system) the following characteristics were estimated: (i) inundation areas, (ii) maxima of water levels, (iii) maxima of flow velocity, (iv) the maxima of the products of water level and flow velocity, (v) the total duration of the flood events, and (vi) the time of exceedence for certain water level thresholds. Operation schemes for reservoirs and polders were applied, which were based on analyses of the actual operation of the existing flood storage facilities or assumed for planned polders according to the operation of existing polders. The inclusion of less plausible events demonstrates the ambivalent role of the flood control system. In Fig. 3 the reductions of flood peaks in relation to the current state is compared for three different system states.

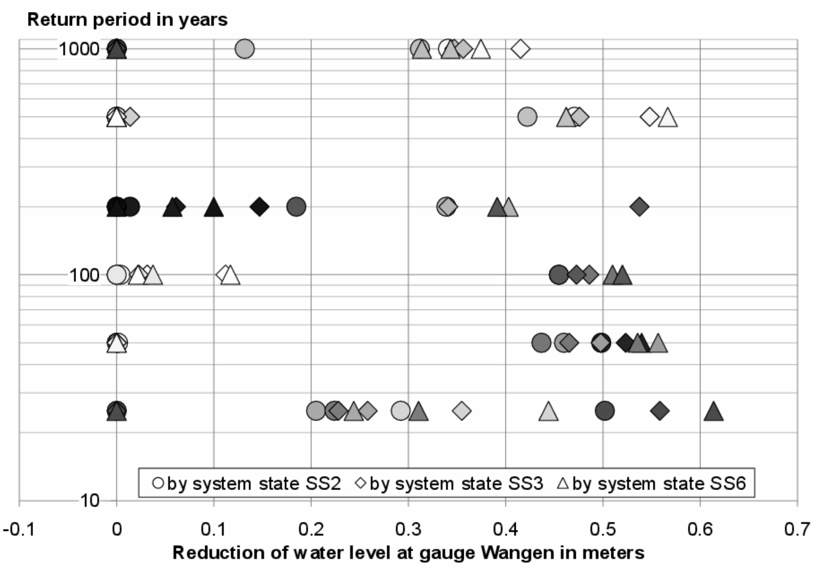

Fig. 3 Reduction of the flood peak at the basin outlet Wangen. Plausibility is depicted in grey scale intensity: plausible events are black, implausible events are white (Schumann *et al.*, 2010).

Under favourable conditions even peaks of very rare floods can be reduced. However, the extended flood control system can have almost no impact on floods with return periods of 50 or 100 years, but the plausibility of these scenarios is low. Such a detailed specification of floods is helpful to characterise the efficiency of the flood control system. An even more important benefit is the characterisation of additional risk, which results from failures of the system under unfavourable conditions. Here it became evident that flood damages may be increased by new polders planned in natural retention areas. Especially for rare flood events the hydraulic conditions may be worsened by additional dykes. The plausibility measures were integrated in two different Multi-Criteria Decision Making (MCDM) frameworks (TOPSIS and F-AHP). It was shown that results with common goals are similar (Schumann *et al.*, 2010). This demonstrates the practical value of the proposed methodology.

CONCLUSIONS

Risk oriented planning depends strongly on the information which can be used to specify hydrological risk. In many cases the flood peak is used as sole characteristic. It is suggested to consider multiple flood characteristics, which are of utmost important for adequately testing the functionality of technical flood retention systems. The combination of these characteristics makes the difference between system failures and effective flood control. Flood scenarios with a probabilistic characterisation through multivariate statistics can be applied to improve flood control planning with special emphasis on possible failures and remaining risks. The application of multivariate statistics demands a large database, which can be derived from simulations by coupling stochastic and deterministic models. However, the results will be uncertain and the derived statistics should be handled as being uncertain as well. This can be done with “imprecise probabilities”. To reduce the information overload for decision makers due to uncertain multivariate probabilities, a methodology was developed based on a plausibility approach. This approach allows the decision makers to retain the classical flood peak based return period and incorporates the other crucial flood characteristics with a plausibility index. The applicability of the methodology was tested in a case study. Because of the inclusion of implausible events, the side-effects of flood protection measures became obvious.

REFERENCES

Blazkova, S. & Beven, K. (2004) Flood frequency estimation by continuous simulation of subcatchment rainfalls and discharges with the aim of improving dam safety assessment in a large basin in the Czech Republic. *J. Hydrol.* **292**(1–4), 153–172, doi:10.1016/j.jhydrol.2003.12.025.

Cameron, D. S.; Beven, K. J. & Tawn, J. (1999) Flood frequency estimation by continuous simulation for a gauged upland catchment (with uncertainty). *J. Hydrol.* **219**(3-4), 169–187, doi:10.1016/S0022-1694(99)00057-8.

De Michele, C., Salvadori, G., Canossi, M., Petaccia, A. & Rosso, R. (2005) Bivariate statistical approach to check adequacy of dam spillway. *J. Hydrol. Engng ASCE* **10**(1), 50–57.

Hundecha, Y., Pahlow, M. & Schumann, A. (2009) Modelling of daily precipitation at multiple locations using a mixture of distributions to characterize the extremes. *Water Resour. Res.* **45**(12), W12412, doi:10.1029/2008WR007453.

Kamrath, P., Disse, M., Hammer, M. & Köngeter, J. (2006) Assessment of discharge through a dike breach and simulation of flood wave propagation. *Natural Hazards* **38**, 63–78, doi:10.1007/s11069-005-8600-x.

Klein, B., Pahlow, M., Hundecha, Y. & Schumann (2010) Probability analysis of hydrological loads for the design of flood control systems using copulas, *J. Hydrol. Engng ASCE* **15**(5), 360–369.

Klir, G. J. (1999) Uncertainty and information measures for imprecise probabilities, an overview. In: *Proc. 1st Int. Symp. on Imprecise Probabilities and Their Applications* (29 June–2 July 1999, Ghent, Belgium).

Koutsoyiannis, D. & Onof, C. (2001) Rainfall disaggregation using adjusting procedures on a Poisson cluster model. *J. Hydrol.* **246**(1-4), 109–122.

Koutsoyiannis, D., Onof, C. & Wheater, H. (2003) Multivariate rainfall disaggregation at a fine time scale. *Water Resour. Res.* **39**(7), doi: 10.1029/2002WR001600.

Lamb, R. & Kay, A. L. (2004) Confidence intervals for a spatially generalized, continuous simulation flood frequency model for Great Britain. *Water Resour. Res.* **40**(7), W07501, doi, 10.1029/2003WR002428.

Lindström, G., Johansson, B., Persson, M., Gardelin, M. & Bergström, S. (1997) Development and test of the distributed HBV-96 model. *J. Hydrol.* **201**, 272–288.

McMillan, H. K. & Brasington, J. (2008) End-to-end flood risk assessment, a coupled model cascade with uncertainty estimation. *Water Resour. Res.* **44**, W03419, doi:10.1029/2007WR005995.

Moretti, G. & Montanari, A. (2008) Inferring the flood frequency distribution for an ungauged basin using a spatially distributed rainfall–runoff model *Hydrol. Earth System Sci.* **12**(4), 1141–1152.

Salvadori, G. & De Michele, C. (2004) Frequency analysis via copulas: Theoretical aspects and applications to hydrological events *Water Resour. Res.* **40**(12), W12511.

Schumann, A. H.; Nijssen, D. & Pahlow, M. (2010) Handling uncertainties of hydrological loads in flood retention planning. *J. River Basin Management*, 8(3), 281–294 © International Association for Hydro-Environment Engineering and Research reprinted by permission of Taylor & Francis Ltd, http://www.tandf.co.uk/journals, on behalf of International Association for Hydro-Environment Engineering and Research.

Walley, P. (1991) *Statistical Reasoning with Imprecise Probabilities*. Chapman & Hall, London, UK.

Zadeh, L. A. (1965) Fuzzy Sets. *Information and Content* **8**, 338–353.