# Floods, droughts and prediction uncertainties

Šárka D. Blažková (ed.)
Patrick Arnaud
Keith J. Beven
Mikhail Bolgov
Libuše Bubeníčková
David A. Jones
Ladislav Kašpárek
Thomas R. Kjeldsen
Adam Kiczko
Alena Kulasová
Michel Lang
Jacques Lavabre
Alberto Montanari
Aurélie Muller
Jarosław J. Napiórkowski
Ondřej Nol
Oldřich Novický
Nadezhda Osipova
Florian Pappenberger
Renata J. Romanowicz
Miroslav Rudiš
Thomas Skaugen
Petr Valenta
Jana Valentová

T. G. Masaryk Water Research Institute, public research institution
Prague, 2011

T. G. Masaryk Water Research Institute, public research institution

Reviewed by prof. Ing. Jiří Zezulák, DrSc.

**Acknowledgements**
The authors thank UNESCO for helping to put together a team of authors by organising the kick-off meeting for writing the book in 2008. Later some new authors joined the team and original ideas have been updated. Some chapters and the printing were supported by the TGM WRI Research Plan MZP0002071101 and colour figures partly by UNESCO.



ISBN 978-80-87402-13-9

## Contents

# 0 The search for better methods for the hydrology of extremes

Šárka D. Blažková, Keith J. Beven

A few decades ago the hydrology of extremes was a statistical problem. Given a set of data, identify an appropriate statistical distribution to represent the data, and extrapolate to the required annual exceedance probability (AEP) (mostly without any concern for the uncertainty in making such estimates). Such methods continue to be used, despite the demonstration 30 years ago that there was a huge sampling realisation effect for periods of the usual length of hydrological records (e.g. Wallis et al., 2007, see also Figure 0.1) and the problem of extrapolation to ungauged sites (e.g. Kjeldsen et al., in press).

One issue that is also often overlooked because standard methods have been defined in engineering practice for different countries is that of whether the extreme tails of the fitted statistical distributions are appropriate. In one sense, of course, we cannot know (for the very same reasons that we have only a small sample of extremes in the observed realisation). In another sense, it does not really matter because if a bigger flood occurs it can always be assigned an AEP that is consistent with a distribution fitted to earlier floods, despite that fact that including a new more extreme event to revise the statistical assessment often changes the tail probabilities (and the uncertainties in those tail probabilities).

What we do know is that there are some important issues associated with estimating the tails. For example:

1. There is an expectation that runoff generation and flood routing processes might change with the magnitude of an event and seasonal antecedent effects (including different types of contributing areas, effects of frozen ground, ice dams, changing vegetation on flood plains etc.).
2. There is an expectation that the observed distribution of floods might represent a mixed distribution of events depending on catchment scale and climatic regime, for example from summer convective events, winter synoptic rainfall events, rain on snow events, hurricane events, ice jam events etc.
3. There is an expectation that the tails should not be infinite (as is often assumed in statistical distributions) because of physical limitations on the potential inputs. Even if the magnitude of a probable maximum precipitation might be difficult to define, many regions show a form of envelope curve for maximum observed rainfalls against duration across all observation sites.

It is the first issue that led to the development of an alternative approach to the estimation of flood frequency, based on derived distributions or continuous simulation, starting with the seminar paper of Eagleson (1972). Runoff generation is, however, a complex of processes that are difficult to simulate in detail in any area since each process representation requires a number of parameters to be estimated. Since this can only be done with some uncertainty the resulting inferences about frequencies will be uncertain (see Chapter by Montanari and Blazkova and Beven).

Thus both statistical and simulation approaches to estimating extremes should be expected to be uncertain, and the question then arises as to what approach should be used to estimate uncertainties. In fitting statistical distributions to flood discharge observations, frequentist

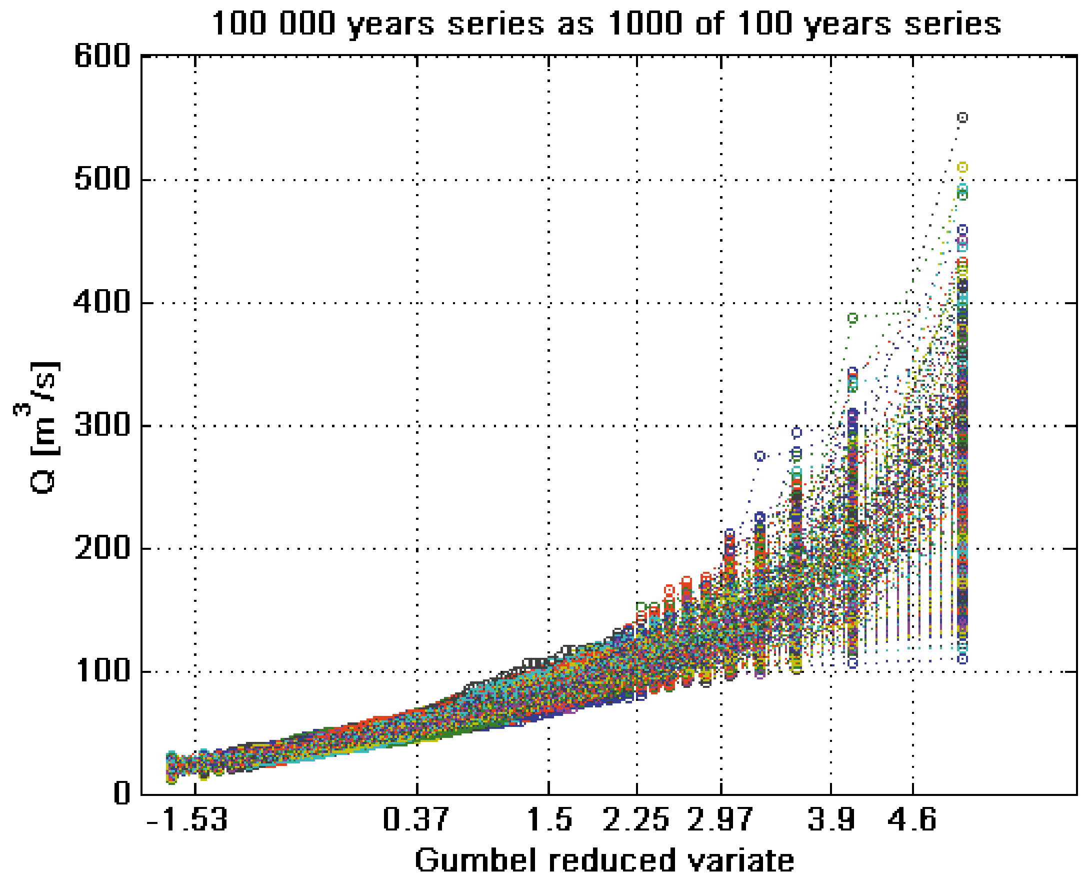



**Figure 0.1** A simulated series of the length of 100 thousand years in hourly time step is plotted as 1000 of 100 years series; the 100 years flood is at the value 4.6 on the x-axis

statistics are often used. But if the nature of the distribution is uncertain, then it might be applicable to use a Bayesian approach to combining results from different assumptions about the nature of the distribution. It may still be necessary to extrapolate with care; Draper (1995) gives a regression example that shows that extrapolations of different regression equation (all with good fits to a limited set of data) to a point beyond the range of data resulted in very different predictions with non-overlapping prediction intervals. His example was an important one: the failure of the o-ring seals on the Challenger space shuttle, but hydrologists have to do a similar degree of extrapolation when estimating very extreme floods for application with similar significance such as evaluating dam safety.

Assessing the uncertainty for simulation models is more problematic still. This is because many of the uncertainties in the simulation process are not statistical in nature but arise from lack of knowledge (epistemic uncertainties). These can be treated within a Bayesian statistical framework if the assumption is made that the epistemic uncertainties behave as if they are statistical in nature. Montanari (this volume) and Bolgov (this volume) demonstrates this approach. Beven (2006 and this volume) explains the reasons why this should only really apply to ideal cases and that most real applications are not ideal in this sense.

In his chapter in this volume, he presents an alternative approach based on the Generalised Likelihood Uncertainty Estimation (GLUE, Beven and Binley, 1992; Beven, 1993) approach which is based on accepting that many different models might be consistent with the available observations (the equifinality principle, Beven, 2006). In the most recent applications of GLUE, limits of acceptability for models are set before making runs of the model. Thus an attempt is made to define acceptability and uncertainty independent of the model residuals. An application of this method to continuous simulation estimation of flood frequency is provided by Blazkova and Beven (2009 and this volume). This approach can also lead to the rejection of all models tried. This might lead to the specification of an improved model, or a model based on different principles (see Romanowicz, this volume) or to the use of quite different modelling concepts (see Skaugen, this volume).

Sometimes it happens that a model is giving good predictions (e.g. flow at the outlet) but for wrong reasons (the inner variables in the model are not in agreement at all with the corresponding states of the catchment, e.g. the wet areas are wrong). We can try to make sure that this is not happening by introducing more criteria, e.g. based on saturated areas mapping (Blazkova et al., 2002ab, Kulasova and Bubenickova, this volume) or the use of groundwater level as in (Kasparek and Novicky, this volume).

By now GLUE has been successfully used in many studies on real catchments. There were, however, also some criticisms from the statistical community or from hydrologists using statistical methods for evaluating simulations (Montanari, 2005; Mantovan and Todini, 2006; Stedinger, 2008). Beven (2006) in a paper called Manifesto for the equifinality thesis has explained the methodology in detail and at the same time presented a new way of evaluating simulations on the basis of limits of acceptability (for a case study see Blazkova and Beven, this volume). The problem of uncertainty in hydrological and more generally environmental modelling is presented in a recent book Uncertainty in Environmental modelling – an uncertain future (Beven, 2009).

Within the uncertainty topic a specific place must be devoted to extremes which are difficult both for statistics and for modelling approaches because the observed series are too short and the errors can be considerably larger than for more usual discharges. Extremes require both forecasting and constructing design events.

Producing design frequencies must cope with uncertainties associated with the lack of data, uncertainty in data, uncertainty in knowledge of distribution and potential non-stationarity due to past and future change. Forecasts require uncertain prediction of future inputs (see Pappenberger, this volume).

The two theoretical chapters of this book (by Montanari and by Beven) are discussing the problems of the treatment of the epistemic uncertainties mentioned above. Can they be treated by statistical methods? What assumptions must be satisfied in this case?

The other chapters are case studies selected in order to illustrate some points presented in this chapter and in the two theoretical chapters.

A successful use of statistical approach is presented by Montanari in his chapter and in the chapter of Kasparek and Novicky.

One difficult practical problem of the assessment of floods is the evaluation of the safety of (especially earthfill) dams. The required return period is 1 or 10 thousand years. Can the continuous simulation replace a statistical analysis? But the continuous simulation approach itself needs distributional assumptions for constructing the series of input. See chapters by Blazkova and Beven and by Lang et al. (this volume).

Catastrophic floods can result in other water management problems the uncertainty of which would probably be even larger. Rudis at al. (this volume) discuss an example of potential re-suspension of contaminated sediments from a river and of sedimentation again in the flood plain with possibly harmful effects on groundwater.

There remain two important topics of uncertainty in hydrology, which are the subjects of very intensive discussion in hydrological community. One is the prediction on ungauged catchments and the other the effect of climate or global change on hydrological design data.

The ungauged catchments are studied in this book in a chapter by Kjeldsen and Jones showing developments of methods of the Flood Estimation Handbook – regression and regionalisation approach. Regionalisation of flood frequency analysis is also discussed by Bolgov and Osipova (this volume).

Many times in the individual chapters of this book the non-stationarity of hydrological data which can be caused by climate change is mentioned. The question to ask is "are the present climate models an adequate basis for assessing future change in extreme events"? At present meteorologists work with ensembles, both of different models and with perturbed parameters but because the models are computationally very demanding the ensembles are rather small and do not cover the parameter space sufficiently. Also, the predictions of precipitation are not so far satisfactory (according to fits on historical periods of data). Beven (2011) suggests a complementary approach for hydrologists, i.e. to run very many simulations of various scenarios as a basis for future adaptive management.

## References


Beven, KJ. (1993) Prophecy, reality and uncertainty in distributed hydrological modelling. *Adv. in Water Resour.*, 16, 41–51.

Beven, KJ. (2006) A manifesto for the equifinality thesis. *J. Hydrology*, 320, 18–36.

Beven, KJ. (2009) Environmental Modelling: An Uncertain Future? London : Routledge.

Beven, K. (2011) I believe in climate change but how precautionary do we need to be in planning for the future? *Hydrological Processes*, Invited Commentary, 25, 9, 1517–1520, DOI: 10.1002/hyp.7963 Published: APR 30, 2011.

Beven, KJ. and Binley, AM. (1992) The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes*, 6, 279–298.

Blazkova, S. and Beven, KJ. (1995) Frequency version of TOPMODEL as a tool for assessing the impact of climate variability on flow sources and flood peaks. *J. Hydrol. Hydromech.*, 43, 392–411.

Blazkova, S. and Beven, KJ. (1997) Flood Frequency Prediction for data limited catchments in the Czech Republic using a stochastic rainfall model and TOPMODEL. *J. Hydrol.,* 195, 256–278.

Blazkova, S. and Beven, KJ. (2002) Flood Frequency Estimation by Continuous Simulation for a Catchment treated as Ungauged (with Uncertainty). *Water Resources Research,* 38, 8, 10.1029/2001WR000500.

Blazkova, S. and Beven, KJ. (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. Hydrology*, 292, 153–172.

Blazkova, S. and Beven, K. (2009) A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation: Skalka catchment, Czech Republic. *Water Resour. Res.*, 45, W00B16, doi:10.1029/2007WR006726.

Draper, D. (1995) Assessment and Propagation of Model Uncertainty. *Journal of the Royal Statistical Society.* Series B (Methodological), 57, 1, 45–97.

Eagleson, PS. (1972) Dynamics of Flood Frequency. *Water Resour. Res.,* 8, 878–898.

Kjeldsen, TR. and Rosbjerg, D. (2002) Comparison of regional index flood estimation procedures based on the extreme value type I distribution. *Stochastic Environmental Research and Risk Assessment*, 16, 5, 358–373.

Kjeldsen, TR. and Jones, DA. (2006) Prediction uncertainty in a median-based index flood method using L moments. *Water Resources Research*, 42, W07414, doi:10.1029/2005WR004069.

Kjeldsen, TR., Lamb, R., and Blazkova, S. Uncertainty in Flood Frequency Analysis. In Beven, K. and Hall, J. (eds) Applied uncertainty analysis for flood risk management. World Scientific/ Imperial College Press, in press.

Lamb, R. (2005) Rainfall-runoff modelling for flood frequency estimation. In Encyclopedia of Hydrological Sciences, Vol. 3, Chapter 125 (eds M. G. Anderson and J. J. McDonnell), Wiley, UK.

Mantovan, P. and Todini, E. (2006) Hydrological Forecasting Uncertainty Assessment: Incoherence of the GLUE methodology. *J. Hydrology*, 330, 368–381.

Montanari, A. (2005) Large sample behaviours of the generalised likelihood uncertainty estimation (GLUE) in assessing the uncertainty of rainfall-runoff simulations. *Water Resources Research*, 41, W08406, doi:10.1029/2004WR003826.

Stedinger, JR., Vogel, RM., Lee, SU., and Batchelder, R. (2008) Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. *Water Resour. Res.,* 44, W00B06.

Wallis, JR., Schaefer, MG., Barker, BL., and Taylor, GH. (2007) Regional precipitation-frequency analysis and spatial mapping for 24-hour and 2-hour durations for Washington State. *Hydrology and Earth System Sciences,* 11, 1, 415–442.

# 1 Uncertainty estimation in hydrological forecasting: should we use statistics?

Alberto Montanari

*Uncertainty assessment in hydrology is briefly revisited by focusing on the opportunity to use statistical approaches. The main pros and cons of statistics in this respect are discussed within a critical view, in the light of recent results proposed by the scientific literature. An application of a statistical methods for assessing the uncertainty of hydrological forecasts is finally presented.*

## 1.1 Introduction

"*Statistics, like veal pies, are good if you know the person that made them, and are sure of the ingredients*". Lawrence Lowell (Harvard President from 1909 to 1933)

"*There are three kinds of lies: lies, damned lies, and statistics*". Benjamin Disraeli (1804–1881)

Uncertainty is an attribute of information (Zadeh, 2005). Uncertainty estimation in hydrology is a wide research field that includes many practical problems, ranging from estimation of parameter uncertainty for hydrological models, estimation of uncertainty in observed data, and many others. In what follows, we will focus on global uncertainty assessment for the output of hydrological models and we will present an application concerning with uncertainty assessment in hydrological forecasting.

Uncertainty estimation in hydrological forecasting is the subject of a growing interest. As a matter of fact, prediction and simulations provided by hydrological models are always affected by a significant uncertainty, much greater than the uncertainty that typically affects the estimation of design variables in other fields of engineering. Therefore uncertainty estimation for hydrological models output is often a necessary prerequisite in order to be able to use the output itself for design purposes. However, uncertainty estimation in hydrology is often a difficult task, mainly because several sources of uncertainty are interacting. Moreover, the sample size of hydrological observed records is often limited, while the variability and heterogeneity of hydrological processes is often much pronounced.

The combination of the above problems makes uncertainty estimation in hydrology highly uncertain. This seems to be a paradox, but actually it is the reality. In order the estimation to be useful, uncertainty should be estimated with a little uncertainty. In other words, at the end of uncertainty assessment the original uncertainty should be considerably reduced. However, by looking at many studies recently presented by the scientific literature it appears that this is not always the case. In fact, in many applications a convincing proof is not provided that the uncertainty in the estimation of, say, the confidence bands of the predicted river flow is lower than the uncertainty that affects the related hydrological model predictions.

One of the most used methods for uncertainty assessment is the Generalised Likelihood Uncertainty Estimation (GLUE). It is often applied because it is very practical and therefore it can be implemented in many situations that are frequently encountered in practice. As a matter of fact, GLUE can always be used, while statistical methods relies on statistical assumptions that are not always satisfied. However, it appears that many users apply GLUE by only performing many model simulations and then deriving so-called "confidence bands" without having any idea of their reliability (and therefore without any idea of their uncertainty). Such uncertainty assessment is probably useless and may be detrimental. In fact, it can induce in the user an illusion of "known uncertainty", that is probably as dangerous as the "illusion of certainty" that might be induced when uncertainty assessment is not carried out (Krzysztofowicz, 2001). The same problem is experienced by users who apply statistical methods for uncertainty assessment without checking the validity of the assumptions underlying the statistical analysis.

One concept that the literature should make clear is that uncertainty assessment can be considered successful, and therefore useful, only when its reliability is assessed. An uncertainty estimation not substantiated by testing could result in an useless information. In the latter years there has been the propensity to suggest that uncertainty assessment should be a necessary requirement in order any hydrological modelling experience to be published. This attitude should probably be revisited, by admitting that a reliable uncertainty assessment is not always possible. Therefore we may say that a hydrological application should be corroborated by a refined analysis aimed at identifying the related sources of uncertainty and checking the possibility to reliably assess uncertainty from a quantitative point of view. The quantitative uncertainty assessment should be carried out only if it is possible to derive indications about its reliability. In any case, the results should be presented in a clear way and terminology (which probably we should agree) in order to clearly convey our degree of confidence on them.

Given this premise, one may conclude that it is extremely important to estimate uncertainty in hydrology by using a method that we can test. This is not a straightforward statement, as hypothesis testing is often complicate within hydrology in general, and within uncertainty assessment in hydrology in particular. Hypothesis testing is a classical and well known branch of statistics. However, the scientific literature has recently argued about the opportunity to use statistical methods for assessing uncertainty in hydrology. Statistics is the most usual method to deal with uncertainty but today many authors are convinced that uncertainty can be efficiently dealt with from a much broader perspective, within which statistical information is one (albeit an important one) of many possible forms of information (Zadeh, 2005; Langley, 2000). This is indeed an interesting perspective, but we believe it is extremely important that we make clear what are the pros and cons of statistics with respect to alternative methods, with regard to hypothesis testing in particular. This paper aims to provide a brief discussion about the opportunity to use statistical approaches for assessing the uncertainty of hydrological model outputs.

### 1.2 Should we use statistics?

It is not easy to provide a definition for "statistical method for uncertainty assessment". A definition of statistics could be: "a mathematical science pertaining to the collection, analysis, interpretation or explanation, and presentation of data". Accordingly to this definition, any uncertainty assessment method could be classified as statistical. However, another definition of statistics reads as: "branch of mathematics dealing with the collection, analysis, interpretation, and presentation of **masses** of numerical data". The presence of the noun "masses" could imply that none uncertainty assessment method in hydrology is statistical. This latter reasoning is not insignificant. In fact, statistics presupposes the availability of extended data sample. Statistical inference based on the analysis of short data samples could be not reliable and even not appropriate.

In hydrology, statistics is frequently applied by using probability theory. Again, it is not easy to provide a definition for "probability". The classical definition of Laplace (1812) says that “the probability of an event is the ratio of the number of cases favorable to it, to the number of all cases possible when nothing leads us to expect that any one of these cases should occur more than any other”. However, this definition is not unanimously accepted today. Actually, a distinction should be made between “frequentist probability”, for which we can use the above definition by *Laplace* (1812) and “bayesian probability”, which needs a different definition that may sound as follows (from Wikipedia): “Bayesian probability is a probability calculus that interprets the concept of probability as the degree of belief in, or the uncertainty about, a proposition. Bayesian methods make extensive use of Bayes’ theorem, which is a consequence of the Bayesian view that probabilities are measures of uncertainty or degrees of belief”.

In practice, on the one hand frequentists talk about probability only when dealing with random experiments that are well defined with given assumptions. The relative frequency of occurrence of an event in repeated experiments is a measure of its probability. On the other hand, Bayesians assign probabilities to any statement, even when no random process is involved, as a way to represent its plausibility in a subjective solution.

At this point, we believe the original questions about the opportunity to use statistical methods for uncertainty assessment should be reformulated in the form: “should we use frequentists methods, or should we prefer subjective methods, including Bayesian analysis?”.

The above question is extremely actual among hydrologists. In fact, the reliability of statistically (frequentist) approaches for uncertainty assessment is often questioned by researchers and practitioners who prefer to use simulation and resampling methods. Indeed, the most used approach for uncertainty assessment in hydrology is GLUE, which is based on the use of a formal likelihood and therefore provides a subjective response, which varies depending on the likelihood and other subjective choices made by the user. GLUE users are much more confident in such kind of subjective assessment, which may be integrated by expert knowledge, with respect to a statistically based approach that is based on questionable assumptions.

In fact, the most relevant advantage of the frequentist approach is its theoretical basis. Many scientists are convinced that the assumptions underlying a statistical analysis are clear and can be tested. For instance, in the context of the meta-Gaussian method illustrated below, the validity of the linear regression on which the uncertainty assessment is based can be tested by checking the homoscedasticity and the Gaussianity of the residuals. However, it is well known that hypothesis testing in statistics is always problematic when dealing with short samples. Moreover, many scientists argue that statistical testing is heavily influenced by the choice of the null hypothesis, that usually is favoured with respect to alternative hypotheses. Many statisticians have pointed out that rejecting the null hypothesis says nothing or very little about the likelihood that the null is true. For this and other reasons, Bayesian statisticians normally do not like the idea of null hypothesis testing.

Moreover, the frequentist approach relies (may be implicitly) on the assumptions of stationarity and ergodicity, whose veridicity is often questioned in hydrology. Looking into the details of uncertainty estimation, the performances of hydrological models are indeed characterized by high variability which could be originated by some non stationarity.

On the other hand, the subjective (may be Bayesian) approach is also affected by some drawbacks. Very often there is no way to test the subjective probabilities and therefore the user is left wondering whether the uncertainty assessment is reliable. Moreover, almost all subjective methods implicitly rely on the assumption of stationarity. Therefore they do not allow us to definitely solve the problem of non stationarity in the hydrological model performances.

So, what might be the conclusion of the above reasoning? We believe that the choice of the most indicated approach necessarily depends on the purposes of the analysis and the behavior of each application. The selection of the most appropriate method should be done by considering that it is very important to test the uncertainty assessment. Within this respect, our feeling is that statistical (frequentist) approaches might be the best choice when dealing with extended data samples, that can allow the users to perform a meaningful hypothesis testing. On the other hand, subjective approaches can provide very useful indications in the context of ungauged basins, when uncertainty assessment is necessarily subjective. But in this case a mean for performing an effective hypothesis testing should be devised.

Ultimately, we believe it is necessary to make clear how the uncertainty assessment is carried out. We should make clear what kind of results we are expecting and their expected reliability. We should recognize that in some cases hypothesis testing could be impossible and therefore uncertainty assessment may be useless. To our opinion this is a fully acceptable conclusion which, however, needs to be adequately justified. When we speak about probability, we should always remember that people still refer to the classical definition of probability. Therefore, when we feel that our uncertainty assessment is subjective, when we suspect for any reason that the estimated confidence bands may fail in including an expected percentage of observed data, we should make it clear and provide an explanation for this result. We believe there are many possible solutions for uncertainty assessment. By revisiting a very famous statement, we may say that “all of them are wrong but some are useful”, depending on how successfully their reliability is proved.

### 1.3 An example of uncertainty assessment in hydrological forecasting by using a frequentist approach

An example of the use a frequentist approach for uncertainty assessment is reported here below, by referring to uncertainty assessment in hydrological forecasting.

In order to estimate the uncertainty of hydrological forecasts, it is assumed here that the forecast error is a stationary and ergodic stochastic process which we will denote with the symbol $E(t)$. We propose to infer its statistical properties by analyzing a past realization $e_{obs}(t) = Q_{obs}(t) - Q_{pred}(t)$ that we assume to be available, where $Q_{obs}(t)$ and $Q_{pred}(t)$ are true and forecasted river flows, respectively. Therefore, this method implies that the hydrological model is preliminarily applied in order to predict past observations by emulating an operational forecasting framework. In this way, the past realization of the forecast error, $e_{obs}(t)$, can be obtained.

To describe the statistical behaviors of $E(t)$, we propose to use a meta-Gaussian model to derive its time varying probability distribution. Basically, the probability distribution of $E(t)$ is inferred depending on $M$ selected explanatory random variables. These are in charge of explaining the variability in time of the marginal statistics of $E(t)$. The statistical inference is performed in the Gaussian domain, by preliminarily normalizing $E(t)$ and the explanatory variables to the Gaussian probability distribution. Normalization is operated through the Normal Quantile Transform (NQT). For more details about the operational use of the NQT see Kelly and Krzysztofowicz (1997) and also Montanari and Brath (2004). In practice, the probabilistic model for $E(t)$ is built as follows.

First of all, it is assumed that positive and negative errors come from 2 different statistical populations $E^{(+)}(t)$ and $E^{(-)}(t)$. Therefore, the probability model for $E(t)$ is given by a mixture of two probability distributions, one for $E^{(+)}(t)$ and one for $E^{(-)}(t)$. The mixture is composed such that the area of the probability distribution of $E^{(+)}(t)$ is equal to the percentage, $P^{(+)}$, of positive errors over the total sample size of the available past realization $e_{obs}(t)$ of the forecast error.

The two realisations $e^{(+)}{}_{obs}(t)$ and $e^{(-)}{}_{obs}(t)$ are normalized through the NQT, therefore obtaining the normalized realizations $Ne^{(+)}{}_{obs}(t)$ and $Ne^{(-)}{}_{obs}(t)$. Then, $M$ explanatory variables, $X^{(i)}(t)$ with $I = 1,\ldots, M$ (which should be readily available at the forecast time) are selected in order to explain the variability in time of the marginal statistics of $E^{(+)}(t)$ and $E^{(-)}(t)$. The values of such explanatory variables for the realizations $e^{(+)}{}_{obs}(t)$ and $e^{(-)}{}_{obs}(t)$ above are estimated and then normalized by using the NQT, therefore obtaining the normalized explanatory variables $Nx^{(i)}{}_{obs}\,(t)$ with $\mathrm{I} = 1,\ldots, M$.

In the Gaussian domain, it is assumed that the forecast error can be expressed as a linear combination of the selected explanatory variables. Let us focus on the positive error. The linear combination can be expressed through the following relationship,

$$Ne^{(+)}(t) = C_1^{(+)} Nx^{(1)}(t) + C_2^{(+)} Nx^{(2)}(t) + \ldots. + C_M^{(+)} Nx^{(M)}(t) + \varepsilon^{(+)}(t) \qquad (1)$$

where $\varepsilon^{(+)}(t)$ is homoscedastic and follows a Gaussian probability distribution. An analogous relationship holds for $Ne^{(-)}(t)$. It is assumed that positive and negative errors are conditioned by the same explanatory variables but with different coefficients of the linear regression (1). Such coefficients are estimated by plugging in (1) the past realizations of normalized forecast error, $Ne^{(+)}{}_{obs}(t)$, and explanatory variables, $Nx^{(i)}{}_{obs}(t)$, and then by identifying the coefficient values that lead to the best fit (for instance by minimizing the sum of squares of $\varepsilon(t)$).

The goodness of the fit provided by (1) can be verified by drawing a normal probability plot and a residual plot for $\varepsilon^{(+)}(t)$. From Figure 1.1 it is clear that unless catchments are located very close together, estimates of the index flood obtained using the original FEH transfer scheme will have prediction $\varepsilon^{(+)}(t)$ as in Montanari and Brath (2004). In the case the goodness of fit is not satisfied, a better result can be obtained by calibrating the regression by using the data points corresponding to the higher river flows only.

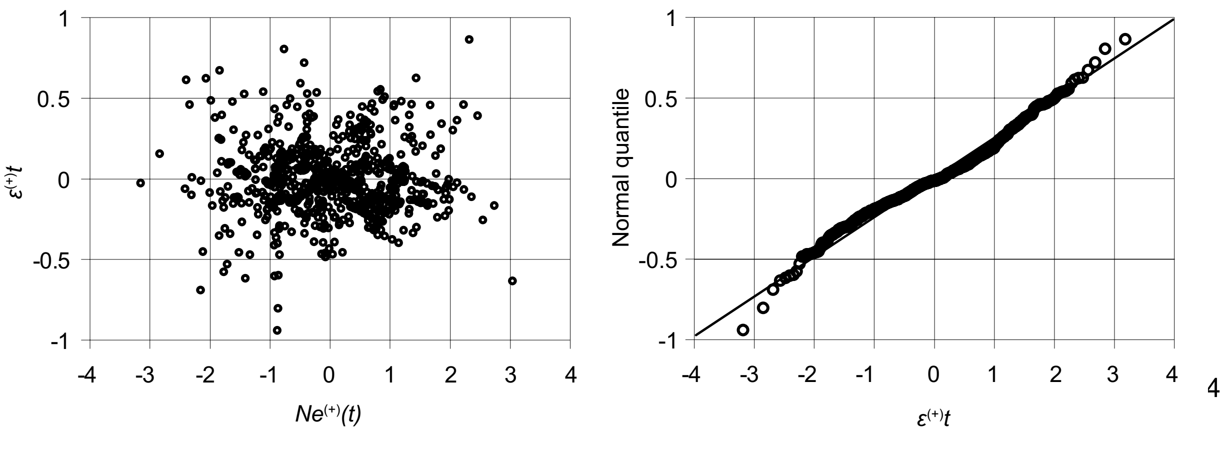



**Figure 1.1** Residual plot (left) and normal probability plot (right) for the residuals of the linear regression (1) applied to the positive 1-hour-ahead forecast errors of the Secchia River synthetic case study

Once the linear regression (1) and its analogous counterpart for negative errors have been calibrated, the probability distribution of the normalized positive forecast error can be easily derived for potential real time and real world applications. Such distribution is Gaussian and is expressed by the following relationship:

$$Ne^{(+)}(t) \sim G\left\{\mu\left[Ne^{(+)}(t)\right], \sigma\left[Ne^{(+)}(t)\right]\right\} \tag{2}$$

where ~ means equality in probability distribution and *G* indicates the Gaussian distribution whose parameters are given by:

$$\mu\left[Ne^{(+)}(t)\right] = C_1^{(+)} Nx^{(1)}(t) + C_2^{(+)} Nx^{(2)}(t) + \ldots + C_M^{(+)} Nx^{(M)}(t)\,, \tag{3}$$

$$\sigma\left[Ne^{(+)}(t)\right] = \sigma\left[\varepsilon^{(+)}(t)\right]. \tag{4}$$

Analogous relationships (from (2) to (4)) hold for the negative error. Therefore, the confidence bands (CB) for the normalized forecast and an assigned significance level can be straightforwardly derived.

In detail, the upper CB of the normalized forecast at the $\alpha$ significance level is given by the $1 - \alpha/(2 \cdot P^{(+)})$ quantile of the Gaussian distribution given by (2), (3) and (4). In the technical computation values greater than 1 of $\alpha/(2 \cdot P^{(+)})$ have to be eliminated, since they have no physical meaning. For instance, if $P^{(+)} = 0.5$ and $\alpha = 10\%$, the normalized upper CB is given by the well known relationship

$$Ne_{90\%}^{(+)}(t) = \mu\left[Ne^{(+)}(t)\right] + 1.96\,\sigma\left[Ne^{(+)}(t)\right] \tag{5}$$

Finally, by applying back the NQT one obtains the CB for the assigned significance level in the untransformed domain. It is important to put in evidence that the $\alpha$ significance level corresponds to the $1 - \alpha$ confidence level. This means that the identified CB of the hydrological forecast are such that there is a probability of $1 - \alpha$ for the true value of the hydrological variable to fall between them.

The reason why positive and negative errors are treated separately was that we would not be able to achieve a good fit through the linear regression (1) if the errors were pooled together. In fact, in this case it appears that the NQT is not effective in making the errors homoscedastic and therefore the assumption of linearity does not hold. The reason for this result is that the NQT is not efficient in assuring homoscedasticity if the mean of the model error is not significantly changing across the range of the error itself, as it often happens when dealing with hydrological models. By treating positive and negative errors separately the problem disappears and the assumptions of the linear regression are met. Finally, it is important to note that the only assumption made about the sign of the future forecast error is that it has a probability equal to $P^{(+)}$ to be positive. Therefore, no inference is made on the sign of the forecast error on the basis of the explanatory variables.

### 1.4 Application to the Secchia River Basin (Italy)

The Secchia River flows northwards across the Apennine Mountains and it is a right tributary to the Po River. The contributing area is 1,214 km$^2$ at the Bacchello Bridge river cross section that is located about 62 km upstream of the confluence in the Po River. The main stream length up to Bacchello Bridge is about 98 km and the basin concentration time is about 15 hours. The mean annual rainfall depth ranges between 700 and more than 2000 mm/year over the basin area. The maximum peak discharge observed at Bacchello

Bridge in the period 1923–1981 was 823 $m^3/s$ (20 April 1960). For the Secchia River, 100 years of synthetic hourly rainfall, temperature and river flow data were generated by using stochastic models (see Montanari, 2005).

The synthetic hydrometeorological data set was then used to perform an extensive forecasting experiment. From the generated data set, some flood events were selected that lasted at least 200 hours, during which the river flow was always higher than 30 $m^3/s$ and reached a peak higher than 200 $m^3/s$. This selection led to picking up 25 floods, with peak flows ranging from about 200 to 672 $m^3/s$.

The above flood events were used to emulate an operational real time flood forecasting, that was performed by using the HYMOD rainfall-runoff model, which was applied at hourly time step. It is a lumped rainfall-runoff model that was introduced by Boyle (2000) and recently used by Wagener et al. (2001), Vrugt et al. (2003) and Montanari (2005). HYMOD was applied to forecast the river flow for each time step of the selected synthetic flood events, with lead time of 1 hour and 6 hours. The precipitation forecasting needed for issuing the 6-hour lead time river flow forecasts was obtained by using the persistent model, which consists of equating the future hourly rainfall to the value observed at the forecast time.

Therefore, for each of the synthetic flood events two forecasted time series of river flows were obtained for the lead times of 1 hour and 6 hours respectively. The corresponding series of forecast errors were also computed.

The first step for the application of the proposed uncertainty assessment procedure for the above forecasts is the identification of the explanatory variables that, once normalized, are plugged in the linear regression (1). These variables should be readily available at the forecast time $t$. By performing extensive tests on the synthetic data, the following explanatory variables were identified: (a) the forecasted river flow, $Q_{pred}(t + \Delta t)$; (b) the average of the absolute value of the past four 1-hour-ahead forecast errors, $E_p$(t); (c) the cumulative rainfall in the six hours preceding the forecast time, $P_p(t)$.

Among the 25 flood events extracted from the synthetic data set, generated by the ADM run, 5 were picked up in order to calibrate the linear regression (1). Calibration was performed by using the data points corresponding to river flows greater than 60 $m^3/s$, therefore obtaining a sample of 683 forecast errors. The percentage of positive errors resulted equal to 79% and 81% for the 1-hour and 6-hour lead times, respectively, which means that HYMOD operated a general underestimation of the river discharge in a potential real time prediction.

Figure 1.1 shows the residual plot and the residual normal probability chart for the linear regression (1) applied to the positive 1-hour-ahead forecast errors. It can be noticed that there is a good agreement with the hypothesis of homoscedastic and Gaussian residuals. A similar result was obtained for the regressions related to the negative forecast errors and the 6-hour lead time (again for both positive and negative errors).

Then, the procedure was applied for assessing the 1-hour-ahead and 6-hour-ahead forecast uncertainty for the other 20 selected events in validation mode. A total of 6654 emulations of a real time forecast were performed for both lead times. Figure 1.2 shows an example of observed and forecasted hydrographs, along with the related 95% CB, for both the 1-hour and 6-hour lead times: the width of the CB increases for increasing lead time, as expected.

The extensive forecasting test performed with synthetic data allows us to perform a systematic testing of the reliability of the identified CB. Laio and Tamea (2007) provided a comprehensive overview of verification tools for probabilistic forecasts of continuous hydrological variables. In particular, they proposed a statistically based technique for testing the accurateness of probabilistic forecasts like the one derived here. However, such method requires the assessment of the whole probability distribution of the forecast in the untransformed domain and not just the CB. Therefore such test, in the context of the technique that is proposed here, would be computationally expensive.

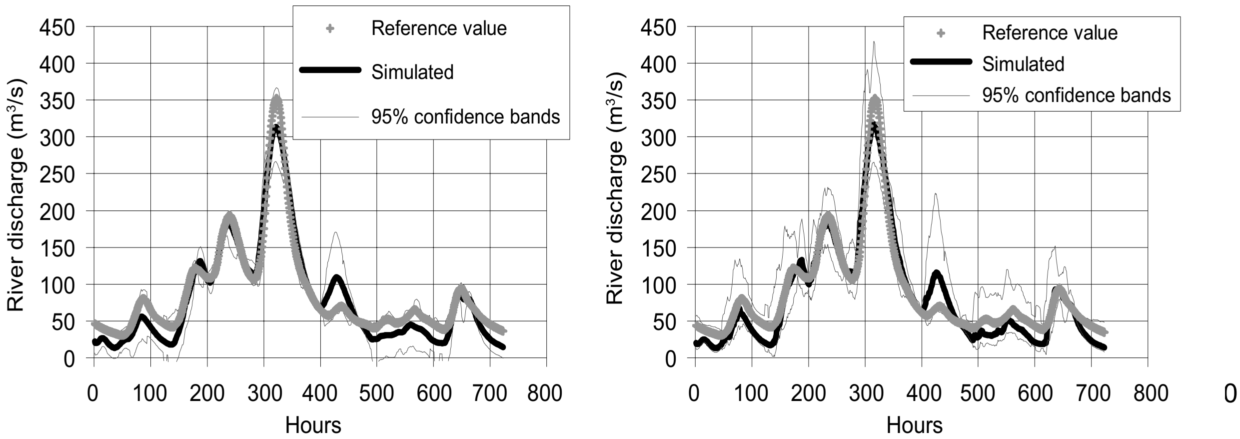



**Figure 1.2** Secchia River synthetic case study. Example of computation of the 95% confidence bands of the forecast with 1-hour (left) and 6-hour (right) lead time. The reference river flow values are the synthetic data that were treated as observed

To avoid the problem, it is suggested that the reliability of the CB is verified by simply computing the percentage *PI* of observations that fall between the CB in validation mode. Ideally, in absence of any non stationarity and sampling variability *PI* should be equal to 95%. Computation has been done separately for positive and negative errors. This test presents the inconvenient that any data located outside the CB is treated as a failure of the uncertainty assessment method, regardless of its distance from the CB themselves. Therefore, we also indicate the *PI* values computed by enlarging the CB of 5%, 10% and 15%.

**Table 1.1** Secchia River synthetic case study. Percentage *PI* of observations lying outside the forecast confidence bands (actual and enlarged). Calibration was performed on 5 flood events randomly selected

| | Actual confidence bands | 5% enlarged confidence bands | 10% enlarged confidence bands | 15% enlarged confidence bands |
|---|---|---|---|---|
| Positive error 1-hour lead time | 1.97% | 0.78% | 0.47% | 0.35% |
| Negative error 1-hour lead time | 5.21% | 2.15% | 0.68% | 0.33% |
| Positive error 6-hour lead time | 1.32% | 0.78% | 0.48% | 0.28% |
| Negative error 6-hour lead time | 9.74% | 6.05% | 3.20% | 1.38% |

Table 1.1 shows the computed *PI* values. It can be seen that the points falling outside the confidence bands are slightly more frequent than expected, summing to a total of 7.18% and 11.06% for the 1-hour and 6-hour lead times, respectively. These values should be

compared with an expected theoretical value of about 5%. This result was induced by using calibration events where the forecast error was slightly more contained with respect to the validation data set. As a matter of fact, the calibration sample has a significant effect on the performances of the uncertainty assessment and therefore it is advisable to pay particular attention to its selection. However, Table 1 shows that almost all the outside observations are located very close to the confidence bands. As a matter of fact, only 0.68% and 1.66% of the points are located outside the 15% enlarged confidence bands, for the 1-hour and 6-hour lead times, respectively.

## 1.5 Conclusions

The hydrologic community is currently animatedly discussing the opportunity to use statistically based methods for uncertainty assessment in hydrology. A personal interpretation of the advantages and drawbacks of statistically based and subjective methods for uncertainty assessment is provided, with the aim to help the user to identify alternative solutions and their peculiarities. Our background is statistics and our research activity within uncertainty assessment has been carried out by applying statistical approaches. Accordingly, this paper presents an innovative method for uncertainty assessment in hydrological forecasting, which is based on inferring the probability distribution of the forecast error by using a statistical approach.

However, we tried to put in light in this paper what are the limitations of statistical approaches in the context of uncertainty assessment and what are the relevant added features of subjective methods.

To draw a final conclusion, we believe that the ultimate scope of uncertainty assessment should be to make clear to the user the magnitude of uncertainty with a little uncertainty. Therefore a proper terminology and a proper hypothesis testing should be used. We believe this is the way to explore for the future research.

**Acknowledgements.** I would like to thank Giovanna Grossi for the cooperation in the development of the uncertainty assessment method described in this paper. I also would like to thank all the colleagues who participated to the 1st and 2nd PUB-IAHS Workshop on Uncertainty Analysis in Environmental Modelling, for the very useful discussions through which I learned about the limitations of statistical approaches in hydrology. The study presented here has been partially supported by the Italian Government through its national grant to the program on “Advanced techniques for simulating and forecasting extreme hydrological events, with uncertainty analysis”.

References

Boyle, DP. (2000) Multicriteria calibration of hydrological models. Ph.D. Dissertation, Dept. of Hydrol. and Water Resour., Univ. of Ariz.

Kelly, KS. and Krzysztofowicz, R. (1997) A bivariate meta-Gaussian density for use in hydrology. *Stochastic Hydrology and Hydraulics*, *11*, 17–31.

Krzysztofowicz, R. (2001) The case for probabilistic forecasting in hydrology. *J. Hydrol.*, 249, 2–9.

Laio, F. and Tamea, S. (2007) Verification tools for probabilistic forecasts of continuous hydrological variables. *Hydrol. Earth Syst. Sci.*, 11, 1267–1277.

Langley, RS. (2000) Unified Approach to Probabilistic and Possibilistic Analysis of Uncertain Systems. *Journal of Engineering Mechanics*, 126, 1163–1172.

Laplace, S. (1812) Théorie analytique des probabilités. Courcier Imprimeur.

Montanari, A. (2005) Large sample behaviors of the generalized likelihood uncertainty estimation (GLUE) in assessing the uncertainty of rainfall-runoff simulations. *Water Resour. Res*., 41, 8, W08406, doi:10.1029/2004WR003826.

Montanari, A. and Brath, A. (2004) A stochastic approach for assessing the uncertainty of rainfall-runoff simulations. *Water Resour. Res*., 40, 1201, doi:10.1029/2003WR002540.

Vrugt, JA., Gupta, HV., Bouten, W., and Sorooshian, S. (2003) A shuffled complex evolution Metropolis Algorithm for optimization and uncertainty assessment of hydrological model parameters. *Water Resour. Res*., 39, 1201, doi:10.1029/2002WR001642.

Wagener, T., Boyle, DP., Lees, MJ., Wheater, HS., Gupta, HV., and Sorooshian, S. (2001) A framework for development and application of hydrological models. *Hydrol. Earth Sys. Sci.*, 5, 13–26.

Zadeh, LA. (2005) Toward a generalized theory of uncertainty (GTU) – an outline. *Information Sciences*, 172, 1–40.

# 2 Uncertainty in hydrological modelling: being positive about equifinality and model rejection

Keith J. Beven

*It seems that my attempt to clarify the justification for continuing to work with the Generalised Likelihood Uncertainty Estimation (GLUE) methodology in the Manifesto for the Equifinality Thesis has not, for many people, been convincing. A number of papers continue to be produced, criticising GLUE for not having a firm foundation in Bayesian statistics and therefore being unscientific or even incoherent. Here the potential for incoherence of formal statistical methods when faced with epistemic uncertainties and the real advantages of recognising the potential for equifinality of models and parameter sets, and learning from model rejections are set out with a view to improving future practice in a positive way.*

## 2.1 Introduction

The Manifesto for the Equifinality Thesis (Beven, 2006a, hereafter referred to as the Manifesto) was an attempt to clarify the justification for continuing to work with the Generalised Likelihood Uncertainty Estimation (GLUE) methodology in terms of the non-statistical nature of many of the errors that we have to deal with in hydrological modelling. Since then, however, a number of papers have continued to criticise GLUE for not having a firm foundation in Bayesian statistics and therefore being unscientific or even (in the case of Mantovan and Todini, 2006) incoherent. Because several of these papers present results from hypothetical examples for which the Bayesian approach can be shown to provide good estimates of model uncertainties and posterior parameter distributions (Montanari, 2005; Mantovan and Todini, 2006; Stedinger et al., 2008) they appear convincing and GLUE appears to be a poor substitute for a formal and apparently objective and scientific methodology. In particular, these authors criticise the subjective choices of a likelihood measure(s) and criteria for rejecting models as non-behavioural. In some cases, they are also clearly bothered by the equifinality concept (that there may be many models that give acceptable simulations within the limitations of the available data) and suggest that if formal Bayesian methods are applied properly it really would not be an issue.

It has also been suggested to me that an expectation of a generic equifinality is a much too pessimistic conclusion and creates difficulties in, for example, engineering applications of hydrological models. Hence, in what follows, I hope both to counter the conclusion that the GLUE methodology might not be useful and to suggest that equifinality is a positive concept on which to base an evaluation of models as hypotheses about catchment response.

The responses to the Manifesto seem to be based on a philosophical (and possibly sociological) response rather than being justified by real experience. It is, in fact, very easy to show that in non-ideal cases (even hypothetical cases for which the wrong assumptions about the errors are made) a formal Bayesian likelihood can over-condition and give the wrong results (see Beven et al., 2007; 2008). In real cases, when unknown input errors and model structural errors might be important, a formal model of the errors might be very difficult to construct, but the wrong choice of error model will over-condition and give incorrect results (as in Figure 2.1). It follows that the choice of a simple formal likelihood measure might actually be a poor choice (even incoherent) for real applications when it is suspected that the structure of the errors is more complex (Beven et al., 2008).

This was explained in the Manifesto but is clearly counter-intuitive to some people. Surely a formal objective methodology should have an advantage over a method that requires subjective decisions. Indeed, some eminent statisticians have argued that formal probability is the *only* way of dealing with uncertainties (e.g. Lindley, 2006; O'Hagan and Oakley, 2004). There is absolutely no doubt that they are right where a formal model of the errors can be found and proven to be valid in both calibration and evaluation periods.

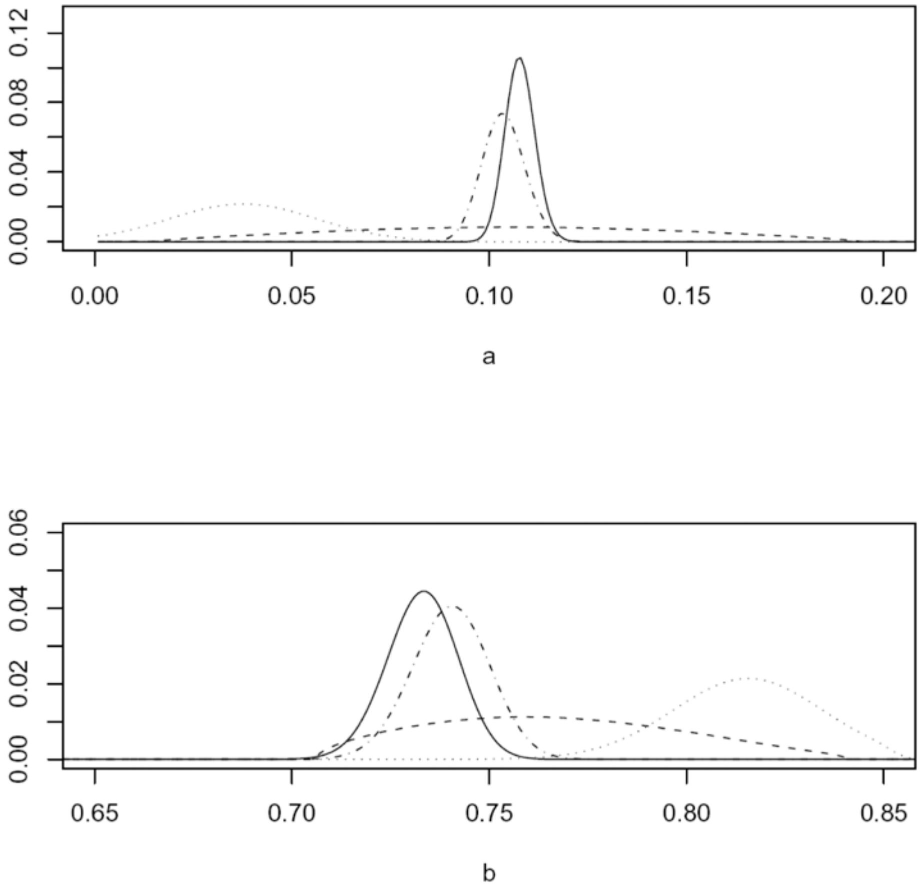



**Figure 2.1** The upper plot shows the marginal posterior distribution of the *a* parameter in the simple *abc* model for a Multiplicative Autoregressive error model (solid), an Additive Autoregressive Model (dotted), GLUE used with a Nash-Sutcliffe likelihood measure (dashed) and a Fitted Box-Cox Transform Autogressive error model (dot-dash) after 500 data points of data simulated with a multiplicative error model and a corrupted input. The lower plot shows the marginal posterior distribution of the *b* parameter under similar conditions. The actual values of the parameters are $a = 0.1$ and $b = 0.75$. Note that in this case the *abc* model structure is known to be correct

But the GLUE methodology is general in this respect (as it says in the name). If we have such a strong belief in a formal error structure then there is absolutely no reason why a formal likelihood cannot be used within the GLUE methodology. In fact, we demonstrated this more than 10 years ago (Romanowicz et al., 1994; 1996), where we even sampled over the error model parameters rather than simply using the expected values calculated from the error series of each hydrological parameter set. It is therefore more than a little frustrating to see criticism of GLUE based on hypothetical examples where there is such a strong belief used in applying the formal Bayes likelihood but where that prior knowledge is then neglected in applying GLUE with a subjective likelihood measure. Of course, if they had assumed the same prior knowledge, their conclusions would not have been so interesting since the resulting posterior parameter distributions and prediction uncertainties would have been almost identical. So it is not really the GLUE methodology that they are objecting to, but their particular choice of informal likelihood measure within GLUE (that did not, in these cases, reflect their prior knowledge). It just does not make such a good story if GLUE gives virtually identical results to the formal Bayes method when equivalent prior knowledge is allowed.

So why not use formal likelihoods always (and thereby avoid being criticised in this way). Because they will not generally work in real applications. In the Romanowicz et al. papers, which were based on real data sets, it seemed that the formal likelihood was over-conditioning the parameter estimates. There are now other examples to real applications in the literature where this is clearly the case (e.g. Thiemann et al., 2001; Feyen et al., 2006). The Bayesian's response to this, of course, would be that although there might be examples of bad practice in formulating a model of the structure of the errors, and that a good error model might indeed be difficult to construct, the formal objective methodology is still to be preferred.

This is evidently true for cases where we can be sure that the errors on the inputs are small, that the model structural errors are small, and that the errors on the observations with which we compare the model outputs have a simple structure (even if not small). But hydrological modelling applications are just not like that. We know that the input errors are often significant, we know that hydrological models often have difficulty in providing good simulations in dry and wetting up periods, or periods affected by snowmelt, we know that errors in discharge measurements can be non-stationary (e.g. when a river scours its bed or goes over-bank in a flood or is subject to macrophyte growth in summer) and that some model predicted variables may be incommensurate with the observations (see the discussion in the Manifesto). This type of errors are often called epistemic errors, in contrast to the aleatory errors that can be treated properly by formal statistical methods. This difference has long been recognised. Knight (1921) for example distinguished between errors that could be quantified in terms of odds, and "real uncertainties" (epistemic errors) that could not (see also Beven, 2009). A formal model of epistemic errors may be very difficult indeed to find, especially when all the errors are assumed to be lumped into a single additive (or multiplicative) error on the discharge. In fact, there is now evidence from hypothetical cases, that such over-conditioning occurs even in only slightly non-ideal cases. Figure 2.1 gives an example; see Beven et al. (2008) for many more cases. In real applications, it may not be appropriate to represent some of these sources of error by probabilities at all; they are much more like the Knightian real uncertainties, although arguments tend to get rather circular at the point in that a statistician can suggest that if an uncertainty cannot be represented in terms of odds then it is outside the domain of statistics (though this is not very helpful in practice of course).

Summarising to this point: there is no reason why, if there is a strong prior belief that a valid formal error model can be found, formal likelihoods cannot be used within the GLUE methodology (although it would not then be a very inefficient way of integrating the likelihood surface to get posterior distributions of parameters, there are much better ways, e.g. Kuczera and Parent, 1998; Vrugt et al., 2003, 2008). But for real applications, it may be very difficult to construct a formal error model; the complexity of the sources of uncertainty are such that the prior belief should be that the error series are likely to be highly complex in structure and possibly non-stationary. It could therefore be argued that (just as all models are wrong and are known to be wrong), all formal likelihood models will be wrong and will be known to be wrong in real applications. Of course the reason for using models in real applications is that models that are wrong might still be useful, and some models might be more useful than others. The same will apply to likelihood measures; even if they are all wrong, some might be useful and some will be more useful than others to achieve the right conditioning of the set of available models.

### 2.2 Likelihood measures and the information content of model errors

So that, in essence, is the question: which likelihood measures might be more useful than others? This question is intrinsically related to the question of the information content of data. One of the real advantages of using a formal likelihood is that it provides a framework for estimating the information content of additional data in conditioning the posterior distributions of parameters and errors (actually this is also possible in principle within the GLUE framework, as demonstrated by the use of empirical information measures in the original GLUE paper of Beven and Binley, 1992).

But, if the assumptions of the error model are false, so will be the assessment of the value of information, and it has already been noted that, very often, the formal likelihood measures will over-estimate the value of information and over-condition the posterior parameter distributions by making the likelihood surface highly peaked. This has the apparent advantage (as claimed by Mantovan and Todini, 2006) of reducing the "problem" of equifinality. It is *always* possible to reduce the apparent equifinality of potential models by making the response surface peakier. This will occur in the formal Bayes methodology, for example, if each model residual is assumed to be an independent random sample from a distribution of errors, conditional on the model being assumed correct (as in Thiemann et al., 2001, and Feyen et al., 2006). The result is evident in Figure 2.1, in comparison with an informal likelihood measure based on the Nash-Sutcliffe efficiency used within GLUE. The informal measure would have to be raised to a high power to have an equivalent peakiness of likelihood to the formal measure.

But think about this. The Nash-Sutcliffe efficiency is a function of the error variance calculated for any individual model (parameter set). Raising this measure to a large power would clearly allow a much greater differentiation between models so that the best models would stand out more clearly. But models that had a very similar error variance before being differentiated by such a transformation would still be giving quite similar degrees of acceptability of model fit (in terms of the Nash-Sutcliffe efficiency). Would we really in actual applications want to differentiate two (or multiple) models, so similar in their error variance performance, in such a way, particularly when we know *a priori* that a different realisation of the input error, or a different period of calibration might reverse their rank order in a list of best models.

When faced with the complex errors of real applications, therefore, some other forms of likelihood measure might be required to represent the "true" information of additional observables in conditioning a model. That is not to say that the type of subjective measures that have been used in past applications of GLUE are a good reflection of the true information content of the calibration data either. In fact tests suggest that they may under-condition, but that this might be an advantage in obtaining more robust ranges of predictions in the face of input and model structural errors (Beven et al., 2008; Smith et al., 2008). Other ways of carrying out model evaluations based on "effective observation errors" limits, rather than residuals, were also suggested in the Manifesto but remain to be tested widely.

It is evident that if we knew a way of assessing the real information content of data for hydrological applications we would also know what form of likelihood weight should be used. The first stage in this is to get people to think about why the real information content of data might be much less than that inferred from simple formal likelihood functions. Input error and the way it can interact with model structural error is the real difficulty in real cases in hydrology, sometimes with a long-lasting carry-over effect. When it appears in rainfall runoff modelling that the model cannot reproduce a hydrograph because there is too much or too little input …is that input error or model structural error, where the structural error might also be conditioned by trying to fit the rest of the record where there may be a quite different form of input error? The complexities of these interactions are the reason why the type of hierarchical Bayes approach allowing for input error, suggested by Kavetski et al., 2005; Kuczera et al., 2006, and others, will not prove ultimately satisfactory.

Such problems are not going to be easily solved by formal likelihood methods unless you are prepared to accept that, like model structures, such a likelihood formulation will be wrong and known to be wrong but possibly a useful approximation the assumptions of which should always be tested post-hoc for validity. It remains difficult, however, to envisage such a formal structure for real, non-ideal, cases (as discussed in the Manifesto) ...and does that make the formal assumptions more scientific than a sensible choice of informal measure, that might require much more data to converge to an answer but which does not over-condition?

And we might yet be able to explore the question of the real information content further. The modifications to the experiments of Mantovan and Todini (2006) reported by Beven et al. (2008) suggest that if single residual errors might be *disinformative* in the face of input and model structural errors, then we will clearly need to average or filter the performance of the model over some blocks of information, where the length of a block should be great enough to integrate over the response time of the system, and any longer time sampling over potential distributions of input error. Clearly, for hydrological models, the potential time scales involved could vary widely depending on the residence times in the system and the nature of the input errors. However, the type of block evaluations that has been used previously in GLUE are clearly of this type.

This problem is related to past discussions of how long a period of hydrological data is required to adequately optimise a hydrological model (and how many parameters might be identifiable given a period of calibration data). Certain pragmatic but ad hoc suggestions have been made in the past but even if there is (as yet) no theory on which to base the choice of block length or appropriate measure, within GLUE those choices have to be made explicit so that they can be discussed, disputed and defended as necessary.

### 2.3 How to be positive about equifinality and uncertainty estimation

This is not a negative conclusion. In fact it is quite possible to be positive about the recognition of equifinality and the effects of different sources of uncertainty in making hydrological predictions. There is, however, one immediate precondition to feeling positive about the equifinality issue and its implications for uncertainty estimation. That is to realise that uncertainty estimation (and the equifinality that might be revealed in estimating prediction uncertainties) should not be the end point of a study. It is, instead, an opportunity that invites the question of what needs to be done to constrain the uncertainty. There are a number of suggestions that can be made about what needs to be done, though some are more positive than others.

The first is to stretch the response surface of the performance or likelihood measure that is used to evaluate a set of parameter values within a particular model structure. This is actually the issue at the heart of the debate about the use of formal statistical methods in parameter estimation for hydrological models noted above. Formal statistical methods are based on strong assumptions about the nature of the modelling residuals. The effect of using a formal likelihood is usually to greatly stretch the apparent difference in likelihood of models that have similar error variance. The question of whether this can be justified or not is discussed in Beven et al. (2008) where it is shown how, for the same hypothetical example as in Mantovan and Todini (2006), even mild departures from the correct assumptions about the error structure can lead to bias in estimates of parameter values. This is precisely because the likelihood surface is stretched so much by the use of the formal assumptions which effectively assume that every residual adds information about the assumed random distribution.

In hypothetical case studies it is clearly possible to construct the problem so that the strong error assumptions are correct (or, in fact, incorrect!). It is also possible to construct the problem so that both the input data and model structure are known to be correct. These are therefore what Beven (2006a) calls ideal cases. Real applications are not ideal in this respect. They involve epistemic (non-random) input and model structural errors; they may also involve commensurability and non-random errors in the observations with which the predictions are compared. Even if input variables had simple random errors (although this is most unlikely) then after processing through a model structure, a much more complex structure would be expected with threshold effects, long-term dependencies, and other non-stationarities. Thus, not every residual will be informative in the sense implied by the formal inference, implying that the stretching of the likelihood surface that results from formal likelihood assumptions will over-condition the parameter estimates.

Thus, since we should expect that residual characteristics in hydrological applications will be generally non-stationary (in bias and/or variance), simple assumptions about the residuals will be difficult to justify. One way round this is to transform the residuals (for example, using Box-Cox or meta-Gaussian transforms) until the simpler assumptions appear more realistic. The effect is, however, similar: if there are a large number of observations, the mathematics ensures that models similar in (transformed) error variance will be given orders of magnitude difference in likelihood even if, visually, their performance does not seem to be better. This is certainly a “solution” to the equifinality issue and theoretically the use of a formal error model allows probability statements to be made about marginal parameter distributions and predicting an observation given the model. The question is whether shoehorning residuals into a strong inference framework based only on randomness is acceptable, when it will certainly have an effect on the parameter inference.

Another approach is to incorporate representations of more sources of uncertainty explicitly. Kuczera et al. (2006), for example, show how (at least in principle) different sources of aleatory uncertainty can be incorporated into a formal Bayesian approach. In particular, they allow for input error in a rainfall-runoff model by allowing rainfall multipliers to be identified for successive rainstorms. This greatly adds to the number of parameters to be identified but results in well defined posterior parameter distributions. However, the whole process is not only dependent on the statistical assumptions being correct but also conditional on an assumption that the model is a correct representation of the catchment dynamics. Since we know this is not the case, then we should expect that the rainfall multipliers will act to compensate for model deficiencies. This might be expect to result in much better predictions in calibration, but there is then the question of how such compensation carries over into prediction when we cannot know what the correct rainfall multiplier for each storm might be, especially in making more extreme predictions.

It appears that for many hydrologists the arguments in favour of the formal statistical approach to uncertainty are overpowering. They would suggest that rejecting this solution is overly pessimistic. But it should certainly be remembered that formal is not always equivalent with correct. The danger of over-conditioning in calibration is well illustrated by the regression analyses of the data on the o-ring seals of the Challenger space shuttle reported by Draper (1995). Different forms of formal statistical regression all gave good fits to the data but quite different predictions at the launch temperature, with non-overlapping uncertainty bounds. The decision to launch was based on a single regression that did not adequately describe the nonlinearity in the failures. This illustrates that formal assumptions can, in some circumstances, over-condition dangerously.

There is a corollary of over-conditioning that if the "perfect" model is driven by input data that are not error-free (and input data errors will also normally be subject to epistemic non-stationarity in hydrology) then it might not prove to have the minimum error variance, and might consequently be given a very low likelihood, relative to the maximum likelihood solution. Thus, we should expect that the maximum likelihood solution should vary with calibration period (or, hypothetically, with different input error sequences for the same calibration period). We would also expect it to vary with different consistent sets of statistical assumptions (the use of the L2 error norm in formal inference is, in effect, an ad hoc choice, albeit mathematically convenient in the pre-computer age, see Tarantola, 2006). So this solution to the equifinality issue does not seem satisfactory, and resulting probability statements might be quite wrong if the formal assumptions about the errors are wrong.

Oops! This is not very positive so far, but I find it difficult to be persuaded that formal likelihood functions are an acceptable solution when nearly all the hydrograph predictions ever published show evidence of non-stationarity in the residual errors. This means that the residuals will be less informative about what is a useful model than if they had a simple random structure. Thus a formal statistical likelihood would always over-condition the likelihood distribution (make the likelihood surface too peaky) in real applications. This therefore suggests that there should be much more hydrologically positive ways of looking at the problem than relying on mathematics to stretch the likelihood response surface.

To avoid the problem of finding a correct statistical likelihood function for the moment, consider a process of model evaluation based on setting prior limits of acceptability as suggested in Beven (2006a). Assume that having gone through an evaluation of different model structures and different parameter sets within a model structure, we find that there is a set of models that appear to fit the data reasonably well in the sense that their predictions always lie within the set limits of acceptability. We could then give each of the models in the set a weight according to how well the models have matched the observations but this is not necessary here. We only need to be happy that each model in the set was acceptable or behavioural in terms of the criteria used. Thus, if the behavioural set is not null (see below if the set is null but we can still be positive for the moment...), then there is equifinality in the modelling process.

We can interpret this as saying that each model in the behavioural set is equivalent to one of many possible hypotheses about how the system is working. Since each model in the set has survived the initial evaluation, they are, as yet, all plausible. We can look at the range of predictions for the set of models to make a (in this case non-probabilistic) estimate of predictive uncertainty under the assumption that since the models were within a certain range of the observations in calibration then they should be within a comparable range in prediction (this is analogous to the statistical assumption that the characteristics of the error model determined in calibration will be similar in prediction). In this approach residual error, in all its complexity and epistemic uncertainty about sources, is handled implicitly. Weighting the output from a particular model run also implicitly weights its associated error series.

It is at this point that we have an opportunity to make further hydrologically relevant tests of the models. If we allow that our set of behavioural models are multiple working hypotheses about catchment response then we can look at their behaviour or predictions to assess whether they continue to be justified as hypotheses. This could be done by further evaluations given new observations of the same type as used in calibration (analogous to a split record test). It could be done by evaluating the predictions of different types of variables to those used in calibration. It could be done by evaluating the predictions of the dominant processes to see whether they are consistent with perceptual model of how that particular catchment responds. This then, provides a framework for doing positive hydrological science, in a way that can use both data and understanding to constrain models as hypotheses (see Beven, 2002a, b).

It is important to recognise, however, that while the construction of predictive models provides a way of formalising our perceptual understanding of hydrological processes, the resulting hypotheses are necessarily simplifications of that understanding. We can perceive more complexity than we can represent mathematically. Thus it may be the case that some compromise will be needed in hypothesis testing. We should not expect that a model will reproduce all the process detail that can be observed in the field. In a very obvious example, we should not expect that a global or catchment average parameter value representing the permeability of the soil will be able to match any local soil hydrological response when we expect soil properties to be highly heterogeneous in other than a very general way. The spatial data might be used to calibrate local parameters (see for example Lamb et al., 1998; Blazkova et al., 2002; or groundwater model applications) but only for those points with observations, leaving degrees of freedom in the spatial pattern. In hypothesis testing, we therefore need to have realistic expectations of what a model can achieve in predicting discharges and other spatial variables using input data that is approximate. What should be considered a realistic expectation (or likelihood measure or range for the limits of acceptability) in reproducing different types of data or process expectation, with different levels of commensurability or observation error? We do not yet know because there have been so few studies of this type.

What our initial experience with this approach has suggested is that it may be rather difficult to find models that fall within reasonable limits of acceptability and all too easy to reject all the models tried. This may be because of model structural error, it may be because the input data are inadequate, it may be because the observations with which the model predictions are being compared are consistently (or even only occasionally) in error. Oops again! This would appear to be a very negative outcome but let me try to convince you that it is, in fact, a positive result if, again, this is not taken to be the end point of a study. The reason why it is positive is that there is the potential to learn a lot from models that fail for all parameter sets. We learn less from models that do not fail, only that they are still feasible hypotheses and can be used with some confidence in prediction. Models that are successful in continuing evaluations will, of course, increase that confidence, but in the case of models that fail then we will need to look for improvements to either the model concepts or the data. Then we may be forced to improve the science.

Naturally, if a model does appear to fail regardless of what parameter sets are tried, then the first thing a modeller will question will be the input data. The second will be the evaluation data. We do not, after all, want to make a Type II error of rejecting a good model just because the data are poor. The limits of acceptability approach will then often reveal some anomalies in a way that a global performance or likelihood measure will not, such as particular time periods when it appears as if no model can get close to the observed data. Is this because of a model limitation, or a data limitation? Resolving this question would be a highly positive outcome, albeit requiring more thought than the normal modelling of calibration and (perhaps) estimating the residual uncertainties.

Another response to model failure is to ensure that the model space has been searched sufficiently to be sure that no good models have been missed. This can be difficult in high dimensional spaces (in the same way as multi-criteria optimisation or importance sampling in high dimensional spaces is difficult), as is the case with many of today's complex process models when it is not possible to securely estimate effective values of parameters independently.

A further response is to relax the rejection criteria or expand the limits of acceptability. We could, for example, in an analogy with formal statistical identification allow that a model should satisfy the limits of acceptability only 90% or 95% of the time rather than 100%. There may be good arguments for allowing a certain relaxation of the rejection criteria, particularly if we have not been able to decide on a good way of representing errors in the inputs (which is often the case, even for variables like rainfall). There is a danger in this, in that the 10% or 5% that is allowed for failure, might be those time periods of most hydrological interest (e.g. around the peaks). Thus any relaxation should be reasoned thoughtfully and carefully.

But finally, if the input data and evaluation observations seem to be reliable, if the model space has been searched thoroughly and if it seems that further relaxation of the limits of acceptability cannot be justified then it will be necessary to try to find a better model. The hypotheses that have been tried have been rejected; a better hypothesis will be needed. We need to go back to the perceptual model and see if something has been left out that is important in representing the real response. This might even mean abandoning well-established theory in favour of something different. Beven (2002a) gives the examples of both the Darcy-Richards equation in modelling subsurface flows in unsaturated soils and the ADE or ADZ models in river dispersion and the use of tracer experiments to fit dispersion parameters.

It is clearly not easy, however, to find a better model. There is an important historical legacy of what is considered reasonable as a representation of hydrological processes, even where this flies in the face of a lot of evidence (which is why the Darcy-Richards equation is still being used in catchment scale models). Why is this? Certainly it is not because the complexities of hydrological processes have not been recognised for a long time (e.g Beven, 1989; 2004; 2007; McDonnell, 2003). It is perhaps because the teaching of hydrology has become too simplistic. This may be for good pedagogical reasons, to suggest to students that there is a body of rigorous hydrological theory, but it does a disservice to the field understanding that underlies our perceptual model of catchment science (e.g. Tromp-van Meerveld and McDonnell, 2006).

So perceptual understanding from field hydrology suggests that we need better process representations, but this will not happen until the classic representations start to be rejected as hypotheses. Model rejection can therefore be positive (e.g. Vaché and McDonnell, 2006); a significant stimulus to improved hydrological theory. The question is whether we have good enough data to be able to distinguish between classical models and new (hopefully improved) process representations. This may not yet be the case; the limitations of input data, characterisation data, and evaluation data may be such that realistic limits of acceptability may be wide enough to accept many different representations as feasible hypotheses (see for example, the different explanations of chloride transport to streams in Kirchner et al., 2001). Equifinality of hypotheses may, with the current state of hydrological measurement, be endemic.

However, we do not have to accept that this will always be the case. We may instead be able to propose ways of testing and differentiating between hypotheses. Indeed analysis of the response characteristics from different models may suggest additional critical measurements that might allow some additional constraints on the range of feasible models (e.g. fractal properties or wavelet characteristics, see e.g. Page et al., 2007). Looking for such hypothesis tests, and new measurement techniques or strategies for doing critical experiments provides a very positive way of doing hydrological science that can combine both observation and theorising.

The question then is whether the hydrological community can come to some agreement about what types of hypothesis should be tested and develop the measurement techniques necessary to carry out the tests at the scales of interest (Beven, 2010). Because hydrology is a difficult science, and given the measurement techniques currently available is likely to remain so, this is not a trivial question. If, for example, we consider the currently fashionable concepts of optimality and maximum entropy production in predicting actual evapotranspiration rates (e.g. Rodriguez-Iturbe and Porporato, 2005; Schymanski et al., 2008), how would we test for optimality given the uncertainties in the energy flux measurements and available water measurements when water availability becomes critical? And when optimality is compromised by some intervention, either a natural weather extreme, fire, or human intervention that might have a long term relaxation time towards a new regime (e.g. Watson et al., 1999) is that a failure of the concept or simply understandable as non-stationarity in boundary conditions? It does make a difference if we want to predict the response at specific places without knowing much about the details of past variability.

That hypothesis testing might be difficult in hydrology should not, however, detract from the advantages of the approach relative to simply thinking up another model structure, calibrating its free parameters using formal statistical inference, and accepting the best model as a good predictor of what might happen in the future. And if this means that there are situations in which we cannot distinguish between different model structures or different sets of parameter values then this should be a positive spur to find ways of doing better hypothesis tests. This is, after all, how the Popperian scientific method is supposed to work. But even Popper allowed for "degrees of verisimilitude". In hydrology, we could interpret that as degrees of success in surviving reasonable hypothesis tests, taking account of the relevant uncertainties. But it is for the community to decide what is considered reasonable for different types of prediction. This suggests a need for communication in developing guidelines for good practice that can only be a positive outcome (see discussions in Pappenberger and Beven, 2006; Faulkner et al., 2007).

Would this be a "solution" to the equifinality issue? Almost certainly not with current measurement technology that leaves such uncertainty in boundary conditions and the detailed nature of storage and flow pathways in hydrological systems. Much of the detail is unknowable and must be inferred from what can be measured, leaving scope for different interpretations. But that is acceptable as long as we do not allow conceptual model representations that are inconsistent with our perceptual understanding (e.g the Darcy-Richards equation, see also Beven, 2006b, or dispersion examples noted above). Rejection of model concepts is a positive result (as long as we are careful not to reject a good model for the wrong reasons). It means that progress is being made. But it also means that rejection must be made respectable. This is not currently the case in the hydrological literature. Model rejection is looked on as model failure and is not easily accepted by referees. Yet, if even the best model found in calibration cannot be shown to be adequate in testing, then it should be rejected if proper account has been taken of the relevant uncertainties. This should be an acceptable result. Referees should then consider if proper account has indeed been taken of the relevant uncertainties in reaching this conclusion, and whether an explanation of the failure has been offered, even if it is not easy to find a better model structure. But that is where creativity is needed, and creativity is where the excitement lies in doing science. Now that is a positive conclusion!

**Acknowledgements.** The thoughts expressed in this comment have arisen out of email and meeting discussions with Rich Vogel, Jery Stedinger, Jasper Vrugt, Alberto Montanari, Günter Blöschl and, at Lancaster, Jim Freer and Paul Smith. Rich and Jery, and Jasper, were good enough to send pre-prints of submitted papers that compared GLUE to formal methods with consequent interesting exchanges that suggested I needed to convey the essence of the GLUE methodology more clearly, while Günter suggested I needed to be more positive because thinking about equifinality was having a bad effect on his students. Mostly we have agreed to disagree about the utility of formal methods but thanks to all!

References

Beven, KJ. (1989) Changing ideas in hydrology: the case of physically based models. *J. Hydrology*, 105, 157–172.

Beven, KJ. (2002a) Towards a coherent philosophy for environmental modelling. *Proc. Roy. Soc. Lond.,* A460, 458, 2465–2484.

Beven, KJ. (2002b) Towards an alternative blueprint for a physically-based digitally simulated hydrologic response modelling system. *Hydrol. Process.*, 16, 189–206.

Beven, KJ. (2004) Robert Horton's perceptual model of infiltration. *Hydrological Processes*, 18, 3447–3460.

Beven, KJ. (2006a) A manifesto for the equifinality thesis. *J. Hydrology*, 320, 18–36.

Beven, KJ. (2006b) The Holy Grail of Scientific Hydrology: $Q_t = H(\underleftarrow{S}\,\underleftarrow{R})A$ as closure. *Hydrology & Earth Systems Science, Discussions*, 10, 609–618.

Beven, KJ. (2008) On doing better hydrological science. *Hydrological Processes* (in press).

Beven, KJ. (2009) Environmental Modelling: An Uncertain Future? London : Routledge.

Beven, KJ. (2010) Preferential flows and travel time distributions: defining adequate hypothesis tests for hydrological process models. *Hydrol. Proces.* (in press).

Beven, KJ. and Binley, AM. (1992) The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes*, 6, 279–298.

Beven, KJ., Smith, PJ., and Freer, J. (2007) Comment on "Hydrological Forecasting Uncertainty Assessment: Incoherence of the GLUE methodology" by Pietro Mantovan and Ezio Todini. *J. Hydrology,* 338, 315–318.

Beven, KJ., Smith, PJ., and Freer, J. (2008) So just why would a modeller choose to be incoherent? *J. Hydrology*, 354, 15–32.

Blazkova, S., Beven, K., Tacheci, P., and Kulasova, A. (2002) Testing the distributed water table predictions of TOPMODEL (allowing for uncertainty in model calibration): the death of TOPMODEL? *Water Resources Research*, 38, 11, W01257, 10.1029/2001WR000912.

Choi, HT. and Beven, KJ. (2007) Multi-period and Multi-criteria Model Conditioning to Reduce Prediction Uncertainty in Distributed Rainfall-Runoff Modelling within GLUE framework. *J. Hydrology*, 332, 216–336.

Draper, D. (1995) Assessment and propagation of model uncertainty. *J. Roy. Statist. Soc.* B37, 45–98.

Feyen, L., Vrugt, JA., O´Nuallain, B., van der Knijff, J., and De Roo, A. (2007) Parameter optimisation and uncertainty assessment for large-scale streamflow simulation with the LISFLOOD model. *J. Hydrology,* 332, 276–289.

Kavetski, D., Kuczera, G., and Franks, SW. (2005) Bayesian analysis of input uncertainty in hydrological modelling: 2. Application. *Water Resources Research*, 42, W03408, doi: 10.1029/2005WR004376.

Knight, F. (1921) Risk, Uncertainty and Profit. Houghton-Mifflin Co. (reprinted University of Chicago Press, 1971).

Kuczera, G. and Parent, E. (1998) Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. *J. Hydrology,* 211, 69–85.

Kuczera, G., Kasvetski, D., Franks, S., and Thyer, M. (2006) Towards a Bayesian total error analysis of conceptual rainfall-runoff models: Characterising model error using storm-dependent parameters. *J. Hydrology,* 331, 161–177.

Kirchner, JW., Feng, X., and Neal, C. (2001) Catchment-Scale Advection and Dispersion as a Mechanism for Fractal Scaling in Stream Tracer Concentrations. *Journal of Hydrology*, 254, 82–101.

Lamb, R., Beven, KJ., and Myrabø, S. (1998) Use of spatially distributed water table observations to constrain uncertainty in a rainfall-runoff model. *Advances in Water Resources*, 22, 4, 305–317.

Lindley, DV. (2006) Understanding Uncertainty. Chichester : Wiley.

Mantovan, P. and Todini, E. (2006) Hydrological Forecasting Uncertainty Assessment: Incoherence of the GLUE methodology. *J. Hydrology*, 330, 368–381.

McDonnell, JJ. (2003) Where does water go when it rains? Moving beyond the variable source area concept of rainfall-runoff response. *Hydrol. Process.,* 17, 1869–1875.

Montanari, A. (2005) Large sample behaviours of the generalised likelihood uncertainty estimation (GLUE) in assessing the uncertainty of rainfall-runoff simulations. *Water Resources Research*, 41, W08406, doi:10.1029/2004WR003826.

O'Hagan, A. and Oakley, AE. (2004) Probability is perfect but we can't elicit it perfectly. *Reliability Engineering & System Safety*, 85, 239–248.

Page, T., Beven, KJ., and Freer, J. (2007) Modelling the Chloride Signal at the Plynlimon Catchments, Wales Using a Modified Dynamic TOPMODEL. *Hydrological Processes,* 21, 292–307.

Pappenberger, F. and Beven, KJ. (2006) Ignorance is bliss: 7 reasons not to use uncertainty analysis. *Water Resources Research,* 42, W05302, doi:10.1029/2005WR004820.

Rodriguez-Iturbe, I. and Porporato, A. (2005) Ecohydrology of water controlled ecosystems: soil moisture and plant dynamics. Cambridge University Press, Cambridge.

Romanowicz, R. and Beven, KJ. (2003) Bayesian estimation of flood inundation probabilities as conditioned on event inundation maps. *Water Resources Research,* 39, 3, 10.1029/2001WR001056.

Romanowicz, R., Beven, KJ., and Tawn, J. (1994) Evaluation of predictive uncertainty in non-linear hydrological models using a Bayesian approach. In Barnett, V. and Turkman, KF. (Eds) *Statistics for the Environment II. Water Related Issues*, Wiley, 297–317.

Romanowicz, R., Beven, KJ., and Tawn, J. (1996) Bayesian calibration of flood inundation models. In Anderson, MG., Walling, DE., and Bates, PD. (Eds) *Floodplain Processes,* 333–360.

Romanowicz, R., Young, PC., and Beven, KJ. (2006) Data assimilation and adaptive forecasting of water levels in the River Severn catchment, UK. *Water Resour. Res,* 42, W06407, doi:10.1029/2005WR004373.

Schymanski, SJ., Sivapalan, M., Roderick, ML., Beringer, J., and Hutley, LB. (2008) An optimality-based model of the coupled soil moisture and root dynamics. *Hydrol. Earth Sys. Sci.*, 12, 913–932.

Smith, P., Beven, KJ., and Tawn, J. (2008) Informal performance measures in model assessment: Theoretic Development and Investigation. *Advances in Water Resources*, 31, 1087–1100.

Stedinger, JR., Vogel, RM., Lee, SU., and Batchelder, R. (2008) Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. *Water Resour. Res.,* 44, W00B06.

Tarantola, A. (2006) Popper, Bayes and the inverse problem. *Nature Physics*, 2, 492–494.

Thiemann, M., Trosset, M., Gupta, H., and Sorooshian, S. (2001) Bayesian recursive parameter estimation for hydrologic models. *Water Resour. Res.,* 37, 10, 2521–2535.

Tromp-van Meerveld, HJ. and McDonnell, JJ. (2006) Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis. *Water Resour. Res.*, 42, W02411.

Vrugt, JA., Gupta, HV., Bouten, W., and Sorooshian, S. (2003) A shuffled complex evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resour. Res.*, 39(8), doi:10.1029/2002WR001642, 2003.

Vrugt, JA., ter Braak, CJF., Gupta, HV., and Robinson, BA. (2008) DREAM: A statistically coherent alternative for generalised likelihood uncertainty estimation. *Water Resour. Res.,* submitted.

Vaché, KB. and McDonnell, JJ. (2006) A process-based rejectionist framework for evaluating catchment runoff model structure. *Water Resour. Res.,* 42, W02409.

Watson, FGR., Vertessy, RA., and Grayson, RB. (1999) Large-scale modelling of forest hydrological processes and their long-term effect on water yield. *Hydrol. Process.*, 13, 689–700.

# 3 Prediction uncertainty in index flood modelling at ungauged catchments

Thomas R. Kjeldsen, David A. Jones

*This paper discusses the prediction variance of an index flood when estimated for an ungauged catchment. Three different methods are investigated: i) using only a newly developed regression model linking median annual flood to a set of four catchment descriptors, ii) an extension using the FEH data transfer method from a nearby gauged catchment to an ungauged catchment, and iii) using a modified version of the data transfer scheme. The results illustrate the link between the structure of the errors of the regression model and the utility of the data-transfer from gauged to ungauged catchments.*

## 3.1 Introduction

Flood frequency analysis based on the index-flood method is the most widely applied method for design flood estimation at ungauged catchments in the UK, as described in the Flood Estimation Handbook (Institute of Hydrology, 1999). The method is based on analysis of annual maximum peak flow data. The index-flood method assumes that data from all catchments within a specified homogeneous region have identical frequency distributions, except for a site-specific scale parameter, the index flood (Hosking and Wallis, 1997). The FEH adopted the median annual maximum flood as the index flood, which differs slightly from the more traditional choice of the mean annual maximum flood. Estimates of the index flood can be obtained using both direct and indirect methods, depending on the availability of data at the site of interest. Direct methods include estimation of the index flood directly from available at-site annual maximum data, whereas indirect methods attempt to estimate the index flood at ungauged sites where no observed flow data are available.

This paper discusses and compares the uncertainty of the prediction errors of the index flood when estimated at an ungauged site using three different methods: i) using a newly developed regression model linking median annual flood to a set of four catchment descriptors, ii) using an extension of this with the FEH data transfer method to incorporate data from a nearby gauged catchment, and iii) using a modified version of the data transfer scheme. Both data transfer methods rely on the regression model, and it will be shown that the correlation structure of errors of the regression model are important when evaluating the uncertainty of the prediction errors of estimates obtained using data transfer.

## 3.2 A hydrological regression model

The estimation of a regression model linking the index flood to a set of catchment descriptors in the UK is described in detail by Kjeldsen and Jones (2008) and only a short summary is given here. To relate the index flood variables from $n$ different catchments to a set of catchment descriptors, consider a vector of sample (log transformed) median annual maximum floods, $\mathbf{y}$, where individual sites are denoted with a subscript $i$. Each sample value is described in terms of a population regression model and two individual error components representing the modelling and sampling errors, $\eta$ and $\varepsilon$, respectively so that

$$y_i = \mathbf{x}_i^T \boldsymbol{\theta} + \eta_i + \varepsilon_i = \mathbf{x}_i^T \boldsymbol{\theta} + \omega_i \qquad (1)$$

where $\boldsymbol{\theta}$ is a vector of regression model parameters and $\mathbf{x}_i$ is a vector of catchment descriptors with a value of one in the first location. Both errors are assumed normally distributed with zero mean values. The covariance matrix of the sampling errors is denoted $\boldsymbol{\Sigma}_\varepsilon$, the corresponding covariance of the modelling errors denoted $\boldsymbol{\Sigma}_\eta$, and the two errors are assumed mutually independent. Further, it is assumed that the elements along the diagonal of the modelling error covariance matrix are identical and equal to $\sigma_\eta^2$. The covariance matrix of the vector of total errors, $\boldsymbol{\omega}$, is defined as

$$\boldsymbol{\Sigma}_\omega = \boldsymbol{\Sigma}_\eta + \boldsymbol{\Sigma}_\varepsilon = \sigma_\eta^2\left(\mathbf{R}_\eta + \boldsymbol{\Sigma}_\varepsilon / \sigma_\eta^2\right) = \sigma_\eta^2 \mathbf{G} \tag{2}$$

where $\mathbf{R}_\eta$ is the correlation matrix of the modelling error. The matrix **G** is introduced for computational convenience and is derived from values of $\sigma_\eta^2$ and $\mathbf{R}_\eta$. In pioneering the use of the Generalised Least Squares (GLS) procedure in hydrology, Tasker and Stedinger (1989) assumed the modelling covariance matrix to be of the form $\boldsymbol{\Sigma}_\eta = \sigma_\eta^2 \mathbf{I}$, i.e. there is an assumption of no cross correlation between the modelling errors. In contrast, the model formulated here is more general and assumes the cross correlation to be represented by the associated modelling error correlation matrix $\mathbf{R}_\eta$.

The sampling and model error components represent two distinctly different sources of error in the regression model. Start by assuming that a 'true' value of the index flood could be estimated for each catchment if an infinite long series of annual maximum peak flow data was available. In practice, the index flood has to be estimated from finite series which introduces a *sampling error* representing the difference between this sample estimate and the notional true value. The *modelling error* represents the inability of a particular regression model to adequately predict the true value of the index flood. For hydrological models such as the regression model studied here, the model error is often much larger than the sampling error if a reasonable number of years have been used to estimate the index flood.

Similarly, the correlations between catchments of the individual error terms have very different interpretations for the two types of error. Correlation between sampling errors is a result of rainfall events causing increased flow in neighbouring catchments at the same time. The existence of correlation in model errors on the other hand, signifies an inability of a particular regression model to adequately represent the true values of the index flood in neighbouring catchments, i.e. the existence of regional clusters of under and over prediction. It can be argued that the existence of model error correlation is a result of an inadequate regression model and should be removed by improving the regression model. However, the approach taken here argues that a simple regression model is unlikely to capture the complex behaviour of real catchments and acknowledges this inability by explicitly allowing model error correlation into the modelling framework.

While the sampling errors are related to the data set used for estimation of the index flood at each individual site, the model errors are specific to a particular regression model, i.e. each choice of a set of catchment descriptors will result in its own specific model error structure. This means that the statistical properties of the sampling error can be estimated once and used in all regression models whereas those of the model error need to be estimated for each regression model tested. Details of the estimation of the sampling error covariance, $\boldsymbol{\Sigma}_\varepsilon$, are not shown here but can be found in Kjeldsen and Jones (2008). Based on detailed investigations, Kjeldsen and Jones (2008) found that model error correlation across sites could reasonably be described for most regression models as

$$r_{\eta,ij} = \varphi_1 \exp\left[-\varphi_2 d_{ij}\right] + \left(1-\varphi_1\right)\exp\left[-\varphi_3 d_{ij}\right]. \tag{3}$$

Here, $d_{ij}$ is the distance between catchment centroids [km] and $\varphi_1$, $\varphi_2$ and $\varphi_3$ are model parameters that must be estimated for each individual regression model.

Having specified the error structure, the regression model parameters can be estimated using a maximum-likelihood procedure which incorporates what are essentially the steps involved in calculating the GLS estimates of the regression parameters. If it is assumed that the regression residuals are normally distributed with mean zero and a total covariance matrix, $\sigma_\eta^2\mathbf{G}$, as described in equation (2), the objective of the overall estimation procedure is to minimise the negative log-likelihood function,

$$-2\ln(L) = \ln\left[\det\left(\sigma_\eta^2\mathbf{G}\right)\right] + \left(\mathbf{y}-\mathbf{X}\boldsymbol{\theta}\right)^T\left(\sigma_\eta^2\mathbf{G}\right)^{-1}\left(\mathbf{y}-\mathbf{X}\boldsymbol{\theta}\right), \tag{4}$$

with respect to the three model error correlation parameters ($\varphi_1$, $\varphi_2$ and $\varphi_3$), the model error variance, $\sigma_\eta^2$, and the regression parameters, $\boldsymbol{\theta}$. The problem is simplified by noting that, for given values of $\sigma_\eta^2$, $\varphi_1$, $\varphi_2$ and $\varphi_3$ (which between them determine **G**) , the value of $\boldsymbol{\theta}$ which minimises (4) is given the GLS estimator

$$\hat{\boldsymbol{\theta}} = \left(\mathbf{X}^T\mathbf{G}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{G}^{-1}\mathbf{y}\,. \tag{5}$$

Thus, estimation by Maximum Likelihood is implemented as a search over the four parameters $\sigma_\eta^2$, $\varphi_1$, $\varphi_2$ and $\varphi_3$. The results are shown in Table 3.1.

**Table 3.1** Summary statistics for regression model describing the (log transformed) median annual maximum flood (Maximum-Likelihood estimates)

| Coefficient | Parameter $\theta_p$ | Standard error | t-value | p-value |
|---|---|---|---|---|
| Intercept ($\theta_0$) | 2.1170 | 0.1172 | 18.06 | 0.000 |
| Ln[AREA] | 0.8510 | 0.0114 | 74.35 | 0.000 |
| $[SAAR/1000]^{-1}$ | -1.8734 | 0.0968 | -19.35 | 0.000 |
| Ln[FARL] | 3.4451 | 0.2654 | 12.98 | 0.000 |
| $BFIHOST^2$ | -3.0800 | 0.1158 | -26.60 | 0.000 |
| $\sigma_\eta^2 = 0.1286$, $df = 597$, $r^2 = 0.945$ (log scale) | | | | |
| $\varphi_1 = 0.4598$ $\varphi_2 = 0.0200$ $\varphi_3 = 0.4785$ | | | | |

Using annual maximum data from 602 rural catchments located throughout the UK, a five parameter regression model was developed, linking the log-transformed median annual maximum flood to a set of four different catchment descriptors. The estimated model parameters are shown in Table 3.1 where AREA, SAAR, FARL and BFIHOST are catchment descriptors describing catchment area [$km^2$], standard average annual rainfall 1960–90 [mm], upstream reservoir attenuation and a measure of the relative baseflow contribution as derived from HOST soil data. These catchment descriptors are available for all gauged and ungauged catchments in the UK larger than 0.5 $km^2$ (Institute of Hydrology, 1999).

The particular choice of catchment descriptors in Table 3.1 and their transformation used here has been based on other analyses which are not described here but which included the examination of the model residuals by plotting them against catchment descriptors.

To estimate the variance of the prediction errors, consider first an estimate of the (log transformed) index flood obtained at an ungauged subject site

$$\hat{y}_s = \mathbf{x}_s^T \hat{\boldsymbol{\theta}} \tag{6}$$

which is considered an estimate of the true (log transformed) index flood, $\xi_s$, defined as

$$\xi_s = \mathbf{x}_s^T \boldsymbol{\theta} + \eta_s \tag{7}$$

where subscript $s$ indicates the ungauged subject site. The prediction error is then defined as

$$\hat{y}_s - \xi_s = \mathbf{x}_s^T \hat{\boldsymbol{\theta}} - \mathbf{x}_s^T \boldsymbol{\theta} - \eta_s = \mathbf{x}_s^T \left( \hat{\boldsymbol{\theta}} - \boldsymbol{\theta} \right) - \eta_s \tag{8}$$

The full variance of the prediction error in equation (8) is given by Kjeldsen and Jones (2007). However, under the assumption that the model error variance is significantly larger than the sampling error in such estimates, the prediction error variance can reasonably be simplified to the model error variance only, i.e.

$$\operatorname{var}\{\hat{y}_s - \xi_s\} \approx \sigma_\eta^2 \tag{9}$$

Table 3.1 also contains estimates of the three parameters describing the correlation between model errors across sites through equation (3). While this correlation does not have a significant effect on the variance of the prediction error when using the regression model at a particular site, it is important in determining the prediction error variance when combining these estimates with data transferred from neighbouring gauged sites, as will be discussed in the next section.

### 3.3 Using data transfer

As the UK has a relatively dense gauging network, the FEH generally recommends using data transfer from 'hydrologically similar' sites for which annual maximum series are available. The data transfer from the gauged to the ungauged catchment is conducted using a scaling factor applied to the non-transformed index flood estimate:

$$m_{s,adj} = m_{s,cds} \frac{m_{g,obs}}{m_{g,cds}}, \quad m = \exp(y) \tag{10}$$

where the subscripts are as follows. *S* and *g*: the ungauged subject site and the gauged sites, respectively, *cds:* catchment descriptor estimates at the gauged and ungauged sites; *obs:* the observed value at the gauged site; *adj:* the adjusted value at the subject site. Kjeldsen and Jones (2007) found the variance of the prediction error of the (log transformed) adjusted index flood, $\hat{y}_{s,adj}$, to be given as

$$\mathrm{var}\left\{\hat{y}_{s,adj} - \xi_s\right\} \approx 2\sigma_\eta^2\left(1 - r_{\eta,sg}\right) + h_{gg} \,. \tag{11}$$

Here $r_{\eta,sg}$ is the correlation between the model errors at the subject site and the gauged site and $h_{gg}$ is the sampling error of (the logarithm of) the median at the gauged site (see Kjeldsen and Jones (2007) for an analytical expression of $h_{gg}$ not shown here). The record length at the gauged site is often sufficiently long that the expression above is dominated by the first term only. Note that if the model error correlation, $r_{\eta,sg}$ , was assumed zero, as done in the GLS model proposed by Tasker and Stedinger (1989), then the prediction error variance becomes almost twice as large as the variance of the error from the regression model alone.

Kjeldsen and Jones (2007) suggested an alternative data transfer scheme

$$m_{s,adj} = m_{s,cds}\left(\frac{m_{g,obs}}{m_{g,cds}}\right)^\alpha \,, \quad m = \exp(y) \tag{12}$$

where the new parameter $\alpha$ is estimated by minimizing the variance of the prediction error for the (log transformed) adjusted index flood $\hat{y}_{s,adj}$ and is given by

$$\alpha = r_{\eta,sg}\frac{\sigma_\eta^2}{\sigma_\eta^2 + h_{gg}} \,. \tag{13}$$

Consequently, the variance of the prediction error for the (log transformed) adjusted index flood is given by

$$\mathrm{var}\left\{\hat{y}_{s,adj} - \xi_s\right\} \approx \sigma_\eta^2\left(1 - r_{\eta,sg}^2\right) + r_{\eta,sg}^2 h_{gg} \,. \tag{14}$$

If a sufficiently long record is available at the gauged site the adjustment factor reduces to $\alpha = r_{\eta,sg}$ and the prediction error variance in equation (14) will be dominated by the first term.

### 3.4 Example

The effect of data transfer on the prediction error variance at an ungauged site as compared to the prediction error variance of an estimate obtained from the regression model only is illustrated in Figure 3.1 by comparing the standard deviation of the prediction error from each of the three methods. Assuming that a long enough record would be available at a gauged site, the only parameter controlling the value of the prediction error variance obtained using data transfer is the distance between catchment centroids, via equation (3) with parameter values listed in Table 3.1.

From Figure 3.1 it is clear that unless catchments are located very close together, estimates of the index flood obtained using the original FEH transfer scheme will have prediction

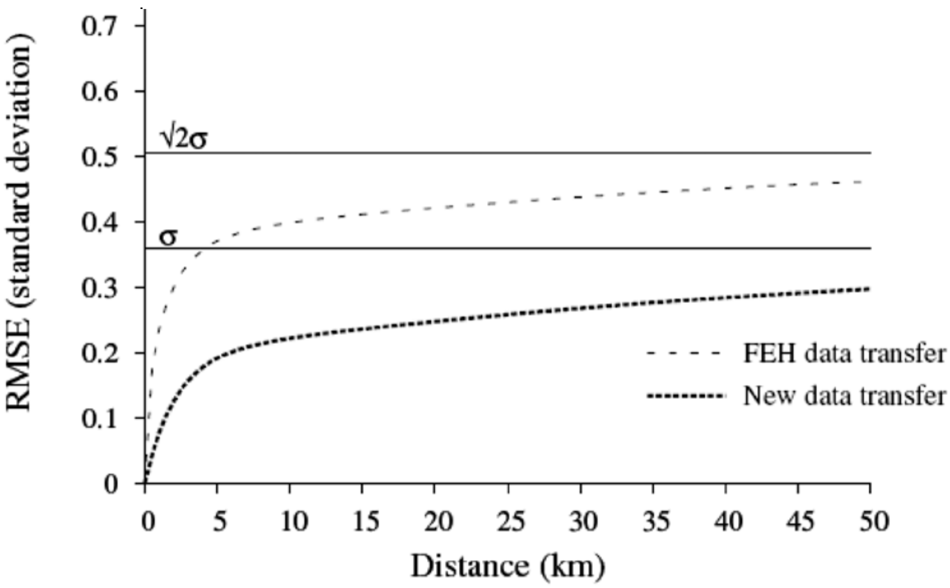



**Figure 3.1** Comparison of standard error of prediction errors at ungauged sites using the regression model only, FEH data transfer and the new data transfer scheme

error variance twice the size of the corresponding estimate from the regression model only. The new transfer scheme, however, ensures that the prediction error variance never exceeds the corresponding variance obtained from the regression model only as the effect of the gauged site is reduced as the distance increases. In Figure 3.1, the standard deviation has been plotted without considering the sampling uncertainty and, hence, the two curves representing the data transfer methods show zero variance at distance zero, where in fact they should show the sampling variance of the gauged site, which would depend on the record-length for that particular record.

## 3.5 Conclusions

By explicitly identifying and estimating the two error sources in a regression model it is possible to derive analytical expressions of the prediction variance of estimates obtained at the ungauged site using data transfer from a gauged site. The results clearly show any improvement to be gained from data transfer arises from correlation in the modelling error. If the model error correlation is not fully taken into account (as in the FEH method) the resulting prediction error of the index flood will have a variance twice as large as the variance of the error from using the regression model only.

References


Hosking, JRM. and Wallis, JR. (1997) Regional frequency analysis: An approach based on L-moments. Cambridge University Press, USA.

Institute of Hydrology (1999) Flood Estimation Handbook. 5 Volumes, Wallingford, UK : Institute of Hydrology.

Kjeldsen, TR. and Jones, DA. (2007) Estimation of an index flood using data transfer in the UK. *Hydrological Sciences Journal*, 52, 1, 86–98.

Kjeldsen, TR. and Jones, DA. (2008) An exploratory analysis of error components in hydrological regression modelling. Submitted to *Water Resources Research.*

Tasker, GD. and Stedinger, JR. (1989) An operational GLS model for hydrological regression. *Journal of Hydrology*, 111, 361–375.

# 4 Regionalization in flood frequency analysis

Mikhail Bolgov, Nadezhda Osipova

*The principal reasons of uncertainty in regionalization in the flood frequency are the following: limited length of samples, complexity of physiographic conditions of runoff formation, irregularity of observational network and small reliability of measurements during catastrophic snow and rain high waters, etc. We have come to the following results: The model of distribution of maximal water runoff of the rivers with rain alimentation can be presented as two densities: GDP for the range of largest values and GEV for other part of the range. These two functions are connected among themselves: GPD is "limiting" case for GEV in consideration of distribution of the large deviations.*

## 4.1 Introduction

A number of recommendations on probability modelling of maximal runoff are well known in hydrology. However because of many reasons a greater uncertainty appears in the solution of this problem. It is possible to mark out the following principal reasons: limited length of processable samples, complexity of physiographic conditions of runoff formation, irregularity of observational network and small reliability of measurements during catastrophic snow and rain high waters, etc.

For the description of distribution of the maximal water runoff three-parametrical distributions of probabilities are most often used: Weibull distribution, Gamma-distribution (Pearson distribution of the third type), Gumbel distribution of extremes, log-normal and logarithmic Gamma-distribution, three-parametrical Gamma-distribution of S. N. Kritskiy and M. F. Menkel, etc. Most of these models are described in detail in Gumbel (1965) and Bobee and Ashkar (1991).

From rather new results it is possible to note stochastic model of maximal runoff, described in Bolgov and Pisarenko (1998), in the form of a mix of two distributions: normal distribution truncated in median and Pareto distribution. Recommended model well reproduced the basic features of the histogram, but a lot of parameters led to computing instability of calculations of them. Besides, in some cases this model gives too high runoff values in the zone of small cumulative probabilities.

Current work describes two problems. The first consists in a choice of suitable parametrical class of applicable models with the possibility of the use of effective statistical methods of estimation of distribution parameters. Further we shall examine two methods of estimations (maximal likelihood and L-moments method) in order to have an opportunity to compare them for small samples. Obtained estimations of parameters should allow to do qualitative conclusions about the character of distribution.

The second problem is connected with the necessity of the increase of accuracy of estimated parameters. It can be overcome by means of data grouping. It is necessary to formulate the assumptions, allowing to pass from individual estimations, obtained for a single river, to estimations of the parameters of aggregated sets.

For the decision of formulated problems we used two distributions, which have appeared rather recently as a result of generalization of the theory of extreme statistics: generalized Pareto distribution (GPD) and the generalized extreme distribution (variation) (GEV) (Embrechts et al., 1977; Singh and Guo, 1995).

Let's discuss the results of estimation of parameters of GPD-distribution for sequences of maximal water runoff of the rivers with rain flood regime. First of all we shall examine properties of different estimations of parameters. For the parameter of distribution form $\xi$ from work (Bolgov and Pisarenko, 1998) Figure 4.1 presents the dependence of maximum likelihood and L-moments estimations. Analysis of the dependence shows that maximum likelihood and L-moments estimations are practically identical at positive values of $\xi$. At $\xi < 0$ correlation between maximum likelihood and L-moments estimations is essentially less. Average value of parameter $\xi$ for the rivers of Primorskiy Kray makes up approximately 0.2 with mean square disorder 0.3 (for samples truncated in median).

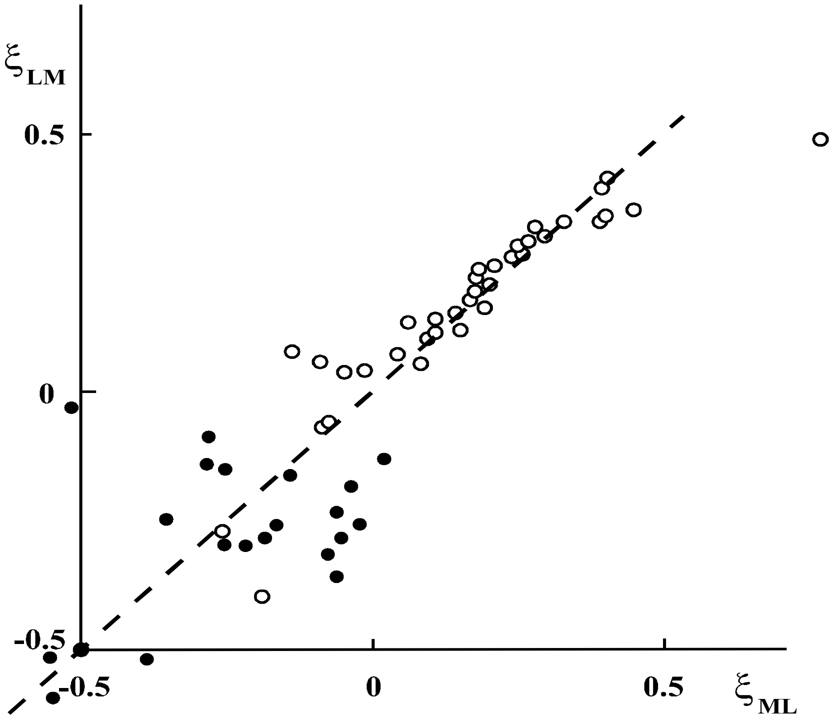



**Figure 4.1** Dependence of sample estimations of the parameter of GPD-distribution form $\xi$ obtained using maximum likelihood and L-moments methods

### 4.2 Observational data of maximal runoff

The data on the maximal annual water runoff of 68 rivers of Primorski Krai were analyzed for the period 1930–1990. As the areas of watersheds of these rivers are essentially various, the processing of such diverse data in the beginning requires the solution of the problem of normalization of observational data and their reduction, to uniform range of changes. This problem is solved as follows. Initial data represent maximal water runoff for one year. Therefore it is reasonable to try to adjust to them distribution GEV. Thus, because for each river the sample duration is insignificant (40–60 annual maximums), it is difficult to expect exact estimations of all three parameters. The purpose of the first stage is to try (even with some mistakes) to estimate the scale parameter $s$ for each river. Then, after normalization of the data for each river on its scale parameter, it will be possible to unite the data according to groups with similar hydrological characteristics. When the number of observations in each group exceeds 100, it will be possible to estimate parameter of the form for each group with enough accuracy.

The other way of estimation of the parameter of form consists in averaging of its individual estimations on set of homogeneous data. Such approach can be in some cases more preferable because the above-described procedure of data normalization can lead to regular increase of dispersion of generalized set because of neglecting many other factors determining individual features of runoff formation.

As a result we obtained the division into districts of the territory of Primorskiy Kray basing on the character of maximal runoff variability estimated by variation coefficient of series. As a criterion of uniformity we have chosen well known in hydrological research ratio of "casual" and "geographical" components of parameter dispersion. Figure 4.2 presents the results of division into districts of the territory of Primorskiy Kray according to the condition of statistical uniformity of variation coefficient of maximal runoff for truncated samples.

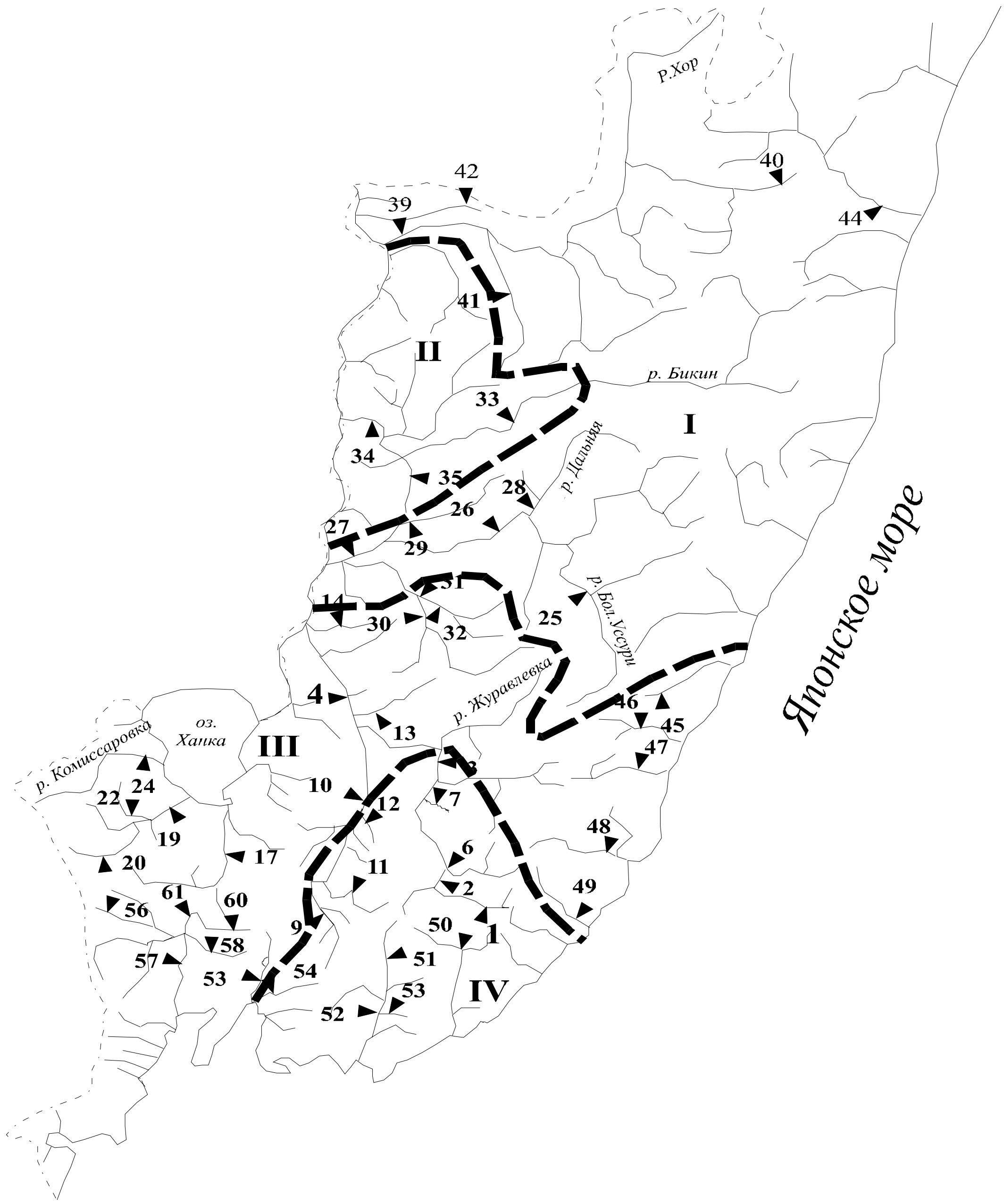



**Figure 4.2** Division into districts of the territory of Primorskiy Kray according to the condition of statistical uniformity of variation coefficient of maximal runoff

So, for each of 68 rivers of Primorskiy Kray three parameters GEV ($c$, $s$ and $\xi$) were estimated. Obtained values of the scale parameter $s$ change on three orders: from 1.19 up to 1484. Parameters of form $\xi$ also noticeably vary, they change from -0.195 to 0.861, but these values are not analyzed at this stage. As it is noted above the purpose of this stage of data processing consists in obtaining of estimations of scale parameter for each river, and further normalizing of data on the corresponding parameter in order to have an opportunity to generalize data on groups of rivers with similar hydrological characteristics.

GEV-distribution describes suitably the histogram for overwhelming majority of observations: the biggest part of points almost ideally lays on the curve of GEV-distribution. But remained largest values also represent the greatest interest for practice. Therefore it is necessary to describe more precisely a tail of distribution of water runoff in the field of rare values. For this purpose further GPD-distribution is used.

So, we shall estimate the parameter of the form $\xi$ of GEV-distribution separately for each of four areas, using normalized values of annual maximal river water runoff. Further, if approximation by means of GEV will lead to unsatisfactory results for maximal values, we shall try to improve approach by means of GPD.

**Table 4.1** Average regional value of the parameter of the form of GEV and GPD-distributions for maximal water runoff of rain floods of the rivers of Primorskiy Kray

| Number of region | Average regional value of variation coefficient $C_v$ | Average regional value of the parameter of the form $\xi$ | |
|---|---|---|---|
| | | GEV | GPD |
| 1 | 0.456 | 0.225 | – |
| 2 | 0.347 | 0.184 | – |
| 3 | 0.733 | 0.733 | 0.199 |
| 4 | 0.641 | 0.530 | 0.153 |

Thus, by means of two distributions GEV and GPD, combined in some point, it is possible to approximate precisely enough the histogram of normalized water runoff on the rivers of Primorskiy Kray in all range of values.

The proposed methodology was applied to the solution of the problem of choice of optimum model of distribution of probabilities for maximal river runoff on the territory of Transbaikalia.

For the solution of this problem we statistically processed data about maximal water runoff of rain floods on 72 rivers in this region with duration of observational period from 30 till 68 years for 1912–1985. At the first stage for each of 72 rivers of region we estimated parameters of GEV-distribution: $s$ and $\xi$. Estimation was carried out using method of maximal likelihood. On the base of obtained results with the purpose of increasing of accuracy of parameters definition for considered class of natural processes all territory was divided into 8 districts (Figure 4.3) according to the conditions of statistical uniformity of variation (variability) coefficient of maximal runoff for using the method of data grouping.

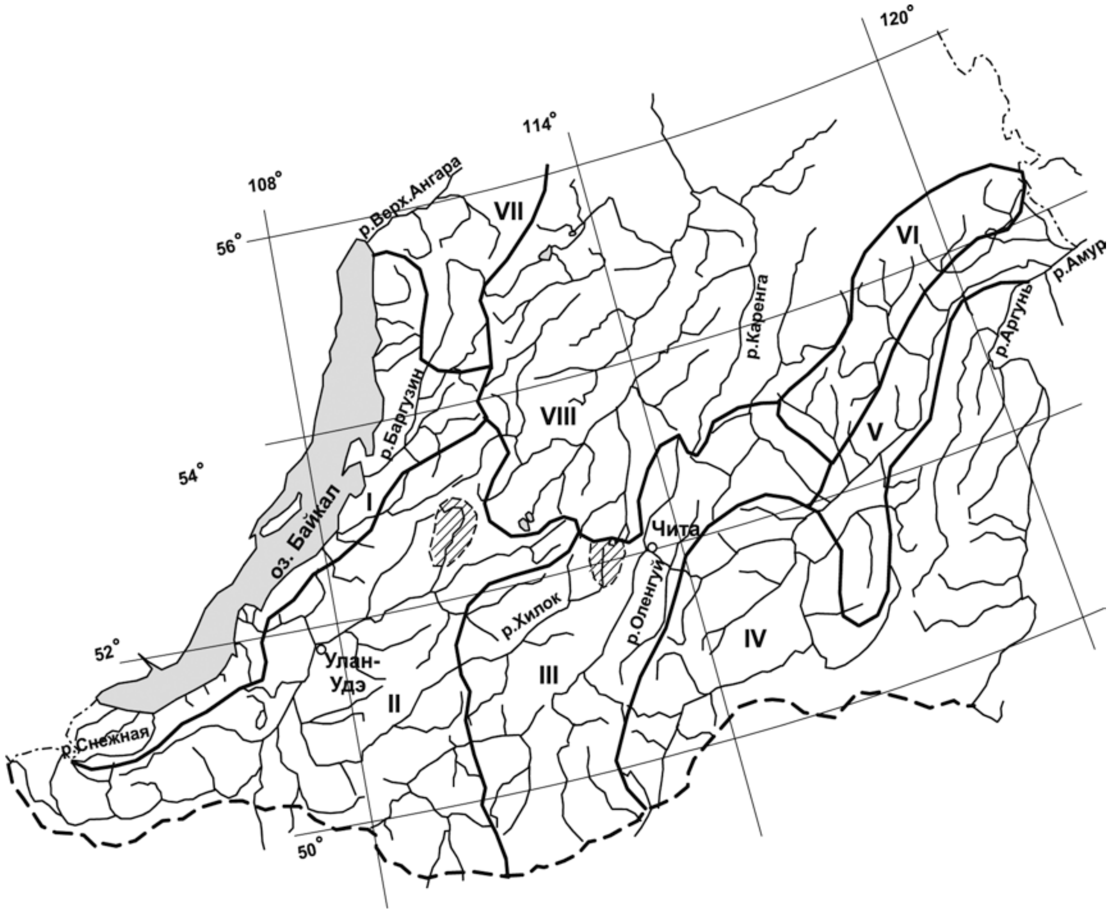



**Figure 4.3** Division into districts of the territory of Transbaikalia according to the condition of statistical uniformity of variation coefficient of maximal runoff

Our problem is to estimate in the best way the top part of the curve of distribution. So all initial series were truncated in median for increasing of the degree of data uniformity and the parameter $\xi$ of GPD-distribution was estimated separately for each of seven regions with enough available data. Values of parameter of the form vary from (-0.24) to 0.37.

**Table 4.2** Average regional value of the parameter of form of GPD-distribution for maximal water runoff of rain floods of the rivers of Transbaikalia

| Number of region | ***Quantity of posts*** | Average regional value of variation coefficient $C_v$ | Average regional value of the parameter of the form $\xi$ |
|---|---|---|---|
| 1 | 10 | 0.43 | 0.16 |
| 2 | 16 | 0.68 | 0.33 |
| 3 | 18 | 0.37 | -0.08 |
| 4 | 11 | 0.93 | 0.37 |
| 5 | 3 | 0.52 | -0.02 |
| 6 | 5 | 0.71 | 0.35 |
| 7 | 2 | 0.23 | -0.17 |
| 8 | 5 | 0.3 | -0.24 |

### 4.3 Discussion of results

It is possible to make the following conclusions:

Application of the theory of extreme statistics in hydrology assumes finding out the character of histogram behaviour in its tail part and the indication of a range of application of this or that model of distribution. It is obvious that GPD model operates starting only from some threshold value and consequently it is necessary to solve rather complicated problem of the search of the optimum value of this threshold (points of truncation as it is accepted to be named in hydrological literature);

The major aspect of statistical analysis is the technology of individual estimation. For "heavy-tail" distributions usual moments can be absent (in the current problem already since the second order) and consequently estimations of maximal likelihood or method of L-moments using ratios between parameters of distributions and serial statistics, estimated on a sample, are recommended;

The model of distribution of maximal water runoff of the rivers with rain alimentation can be presented as two densities: GPD for the range of largest values and GEV for other part of the range. These two functions are connected among themselves: GPD is "limiting" case for GEV only in consideration of distribution of the large deviations.

References

Bolgov, MV. and Pisarenko, VF. (1998) *Water Resources*, Vol. 26, № 6, p. 710.

Gumbel, E. (1965) Statistics of extreme values. Moscow : Mir, 450 p.

Bobee, B. and Ashkar, F. (1991) The Gamma Family and Derived Distributions Applied in Hydrology. Toronto : Water Res. Public., 203 p.

Embrechts, P., Kluppelberg, C., and Mokosh, T. (1977) Modelling of Extreme Events. Berlin : Springer, 645 p.

Handbook of Hydrology (1992) Ed. Maidment, DR. McGraw-Hill, p. 507.

Hosking, JRM. (1990) *J. Roy. Statist.* Series B, 1990, Vol. 52, № 1, p. 105.

Singh, VP. and Guo, H. (1995) *Hydrol. Sci. Journ.,* Vol. 40, № 2, p. 165.

# 5 The realisation effect in the estimation of flood frequency

Šárka D. Blažková, Keith J. Beven

*The method of continuous simulation within the GLUE uncertainty framework has been used on the Skalka catchment (671.7 km$^2$). The study presents evaluation of flood frequency model using the limits of acceptability based on the uncertainty of rating curves and snow water equivalent extrapolation. To get a sufficient ensemble of behavioural models a scoring method of relaxing the limits has been employed. The important effect of randomness in flood frequency simulation is presented. In order to produce reasonable estimates of floods of 10 thousand years return period, series of the lengths of 100 thousand years should be computed.*

## 5.1 Introduction

Continuous simulation in order to estimate flood frequency characteristics has some advantages compared to direct statistical or event based methods. It is possible to compute very long series of precipitation and flow and evaluate the same way as it is done in statistical methods with observed series. The antecedent conditions before extreme floods are continuously modelled so that it is not necessary to decide on antecedent wetness as it is done with event based methods. Examples can be found in Beven (1986a, b, 1987), Blazkova and Beven (1995, 1997, 2002, 2004), Calver et al. (1999), Cameron et al. (1999, 2000, 2001, 2006).

Hydrological models in general and extremes in particular are known to have large uncertainties (e.g. Beven and Binley, 1992; Gupta et al., 1998; Beven and Freer, 2001; Beven, 2008). It is therefore necessary to take them into account and instead of a single value for a flood with a certain return period to produce uncertainty bounds. Following Beven (2006) we are using acceptability limits computed from the data of current metering.

## 5.2 The catchment

The Ohre River down to the Skalka dam has the area of 671.7 km$^2$ and a range of altitudes from 460 to 1041 m above sea level. The annual flood series was available at 6 stations (Table 5.1 and Fig. 5.1). Current metering data were available at 3 sites.

**Table 5.1** Flood statistics for gauged discharge stations in the Skalka catchment (data from Landesamt für Wasserwirtschaft München and CHMI; output of Hosking program (1997) for regionalisation)

| Subcatchment or interbasin | Area | Water gauge Station – Figure 5.1 | No. of years of observation | Mean flood [m$^3$/s] | Estimated 100 yr return period flood [m$^3$/s] |
|---|---|---|---|---|---|
| upper Eger | 114.420 | Marktleuthen on Eger | 67 | 20.28 | 50.37 |
| lower Eger | 209.376 | Hohenberg on Eger | 37 | 29.94 | 74.38 |
| upper Roslau | 130.65 | Lorenzreuth on Roslau | 39 | 22.64 | 56.25 |
| Kossein | 94.82 | Marktredwitz on Kossein | 34 | 15.91 | 39.52 |
| lower Roslau | 91.11 | Arzberg on Roslau | 28 | 53.86 | 133.80 |
| Ohre | 31.3240 | Cheb on Ohre | 60 | 89.98 | 223.52 |

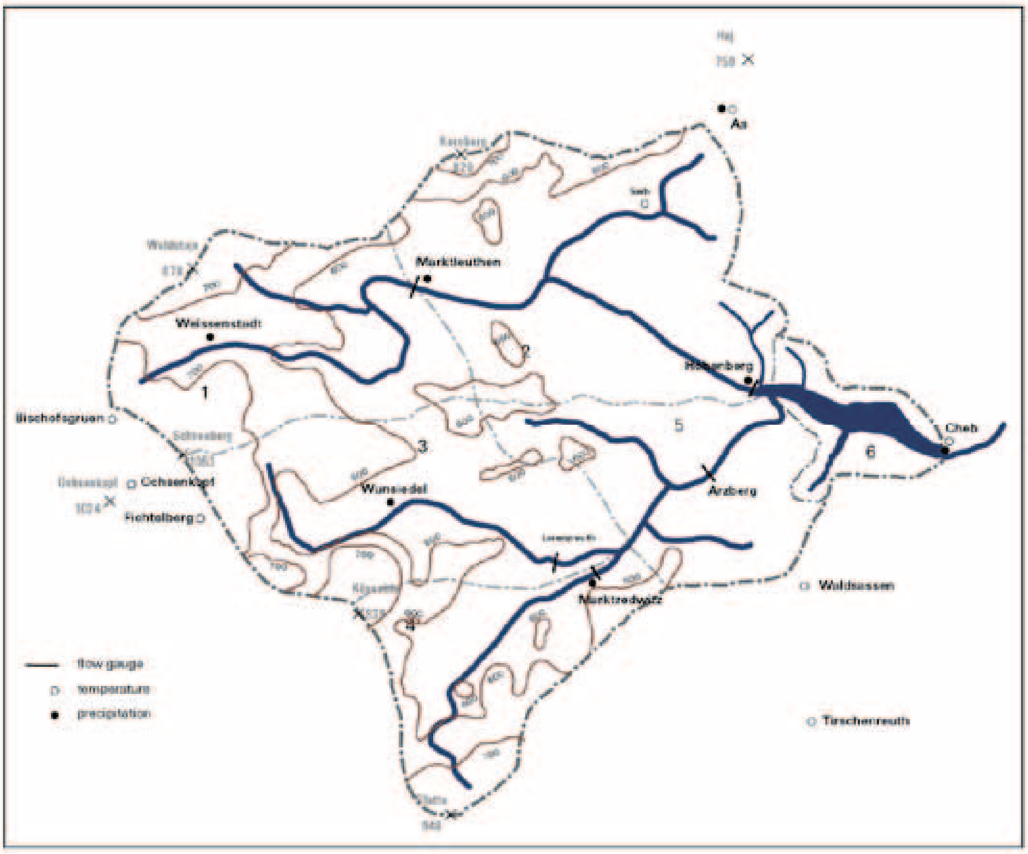



**Figure 5.1** The Skalka catchment on the Ohre (Eger) River with the numbers of subcatchments, flow gauges, temperature and precipitation stations (Blazkova, S. and Beven, K. (2009) A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation: Skalka catchment, Czech Republic, Water Resour. Res., 45, W00B16, doi:10.1029/2007WR006726, reproduced with permission)

## 5.3 The precipitation and runoff model

For this study a frequency version of TOPMODEL has been used, i.e. the input into the runoff model is provided by a stochastic precipitation simulator. There is also a snow accumulation and melt routine. The precipitation events move over the catchment and account is taken of a strong precipitation gradient. Within the GLUE Monte Carlo framework the parameters of both precipitation and runoff models are sampled from uniform distribution with physically reasonable ranges. Simulations are done in two stages. First the same length is simulated as the length of observed series. After selecting the behavioural parameter sets (see below) long series (10 thousand years) are run to estimate the flood with return period of 1000 years.

## 5.4 The acceptability limits

The uncertainty bounds of the rating curve have been computed using fuzzy regression HBS1 (Hojati et al., 2005). The fuzzy bounds are very similar to the statistical bounds (Fig. 5.2). Individual simulation have been evaluated using the acceptability limits of flow duration curve at 2 sites, flood frequency curve up to 10 years return period at 5 sites, and

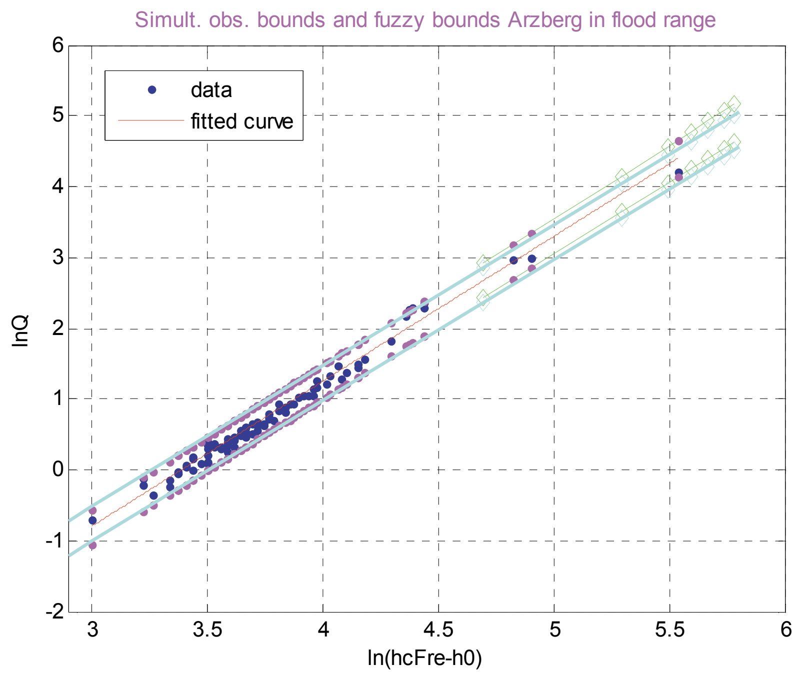



**Figure 5.2** Rating curve; statistical simultaneous observation bounds (green with magenta points) on log log scale and fuzzy bounds using HSJ1 method (cyan); red line – the statistical estimate; diamonds – quantiles of flood frequency curve

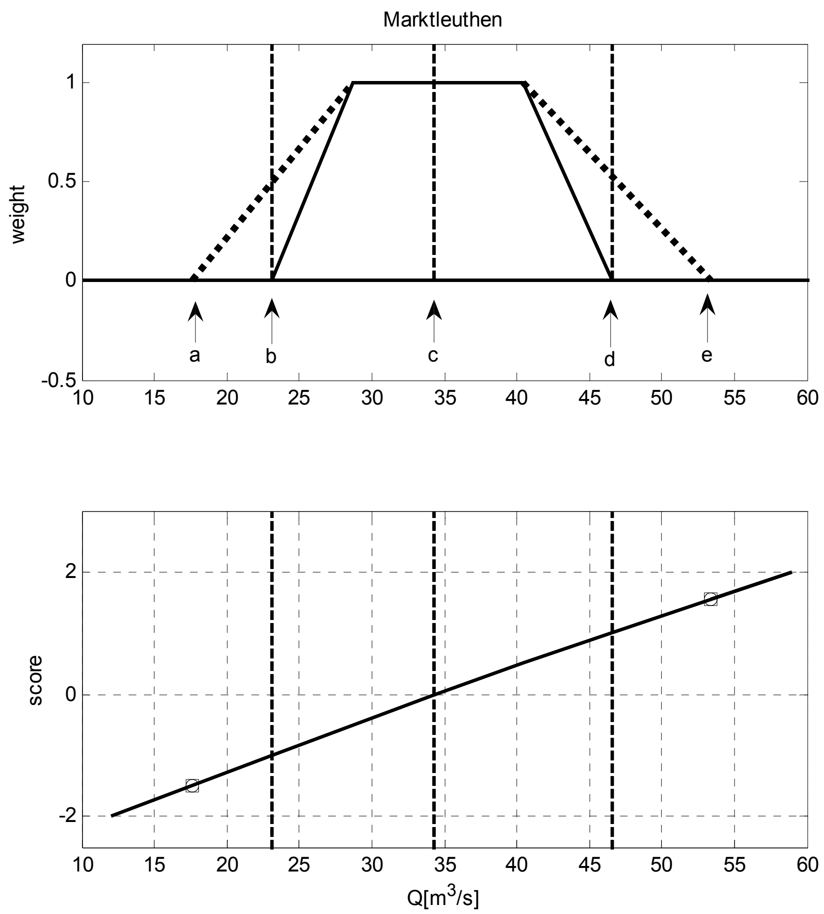



**Figure 5.3** Upper plot: trapezoidal weighting function; full line trapezoid – original limits of acceptability (points b and d), dashed lines – the estimate and acceptability bounds, dotted lines – the expanded trapezoid (points a to e); Lower plot: scores; squares – points to which the bounds have to be expanded

the maximum annual snow water equivalent at 4 sites where the acceptability bounds have been based on the regression with elevation. Only few of the simulations lied within the bounds on every quantile, some of the others however, came pretty close. The acceptability limits were then expanded after putting the different evaluation measures on a common scale by treating each evaluation in terms of a normalised score that has the value -1 at the lower limit, 0 at the observed value and +1 at the upper limit (Fig. 5.3). Trapezoid in Fig. 5.3 was then used as a weighting function on the individual quantiles for computing uncertainty bounds. An example of the original acceptability limits and of the expanded bounds is in Fig. 5.4.

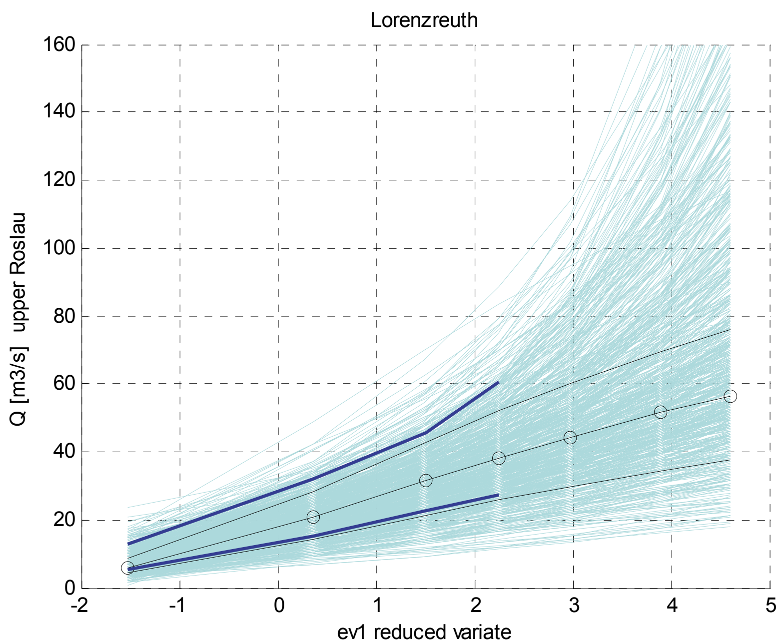



**Figure 5.4** Flood frequency curve at Lorenzreuth; black lines – frequency curve from observed data with uncertainty bounds, o – flood quantiles, cyan – the simulations, blue lines – uncertainty bounds from the trapezoidal weighting

## 5.5 Discussion and conclusions

A methodology for estimating the frequency of flood extremes is described based on continuous simulation. Model predictions from a large number of Monte Carlo simulations using different parameter realisations are compared against summary information of the flow duration curve, and the frequency characteristics of flood discharges and snow water equivalent. Model evaluation is carried out within the extended GLUE limits of acceptability approach. Since the parameters of the stochastic input model are part of the evaluation exercise in continuous simulation of this type, it is not possible to make an explicit assessment of input error. However, observational errors have been estimated in both discharges (using rating data) at 5 sites within the catchment, and snow water equivalent in 13 snow zones, 4 of which have observed data.

The uncertainty limits set before the simulations had to be expanded to get enough simulations for the computation of uncertainty bounds. One justification for this is the effect of different input realisations on acceptability. We have taken one of the behavioural parameter sets and generated 10,000 input sequences of the same length as the observed flood series. The realisation effect is obvious from Fig. 5.5. It follows that other parameter sets with relatively low critical values for acceptability might be fully behavioural given other input realisations (Blazkova and Beven, 2009).

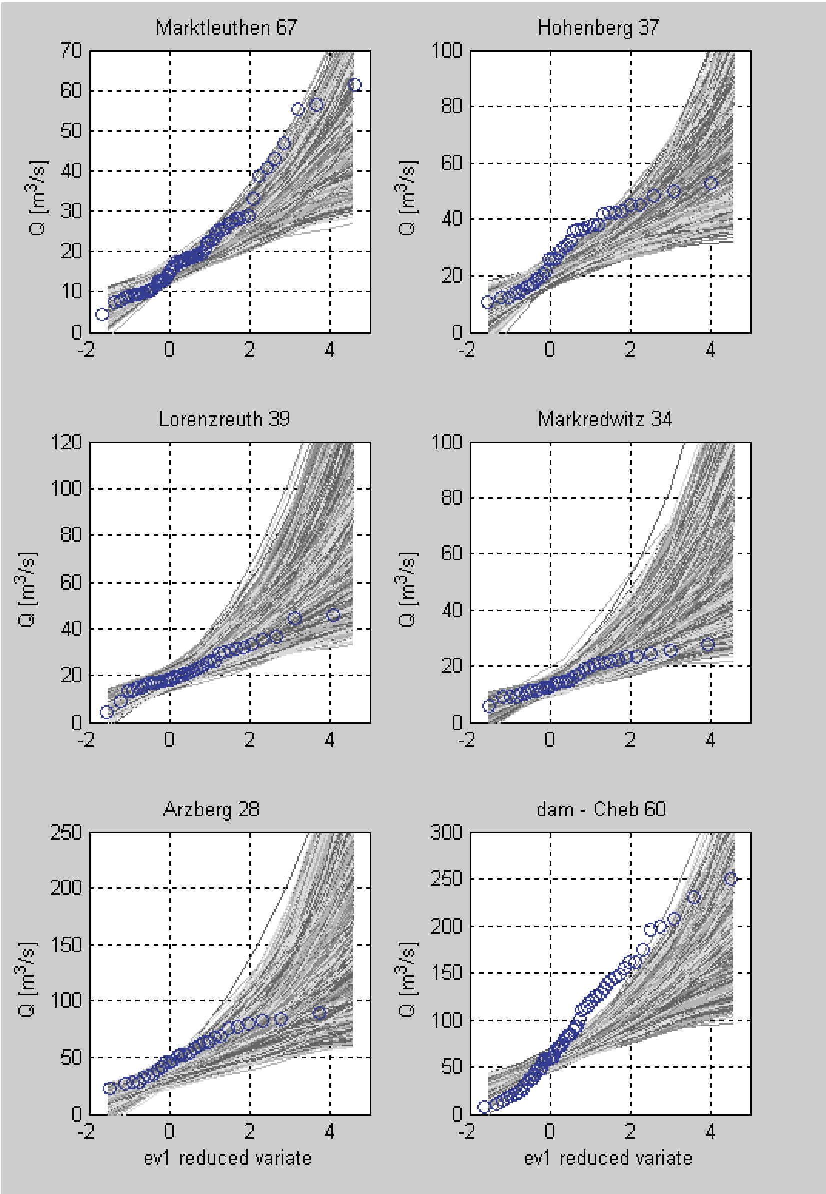



**Figure 5.5** Ten thousand realisations of flood frequency curves with the same parameter set; grey – flood frequency curves, blue circles – observed annual floods, numbers in the title are lengths of observation in years and also lengths of the simulation in the particular subcatchment (Blazkova, S. and Beven, K. (2009) A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation: Skalka catchment, Czech Republic, Water Resour. Res., 45, W00B16, doi:10.1029/2007WR006726, electronic supplement, reproduced with permission)

With the simulations within the expanded limits over 4000 simulations of the length of 10 thousand years have been computed in the paper of Blazkova and Beven (2009). This gave an uncertainty estimate which can be considered reasonable up to the return period of thousand years. For the design and assessment of safety of some important dams

10 thousand years return period is required. The logical way is to model series of the length of 100 thousand years which is computationally extremely demanding. The reason for such an exercise can be seen in the following Figures 5.6 and 5.7. Series a1 in Fig. 5.6 looks like having all the points (annual peaks) approximately in the right position on the ev1 probabilistic paper. But series a2 (Fig. 5.7) shows a flood obviously larger than what corresponds even to the 100 thousand years. Such a flood, however, could have been modelled in any 10 thousand years series and if taken into account it would grossly distort the flood peak distribution.

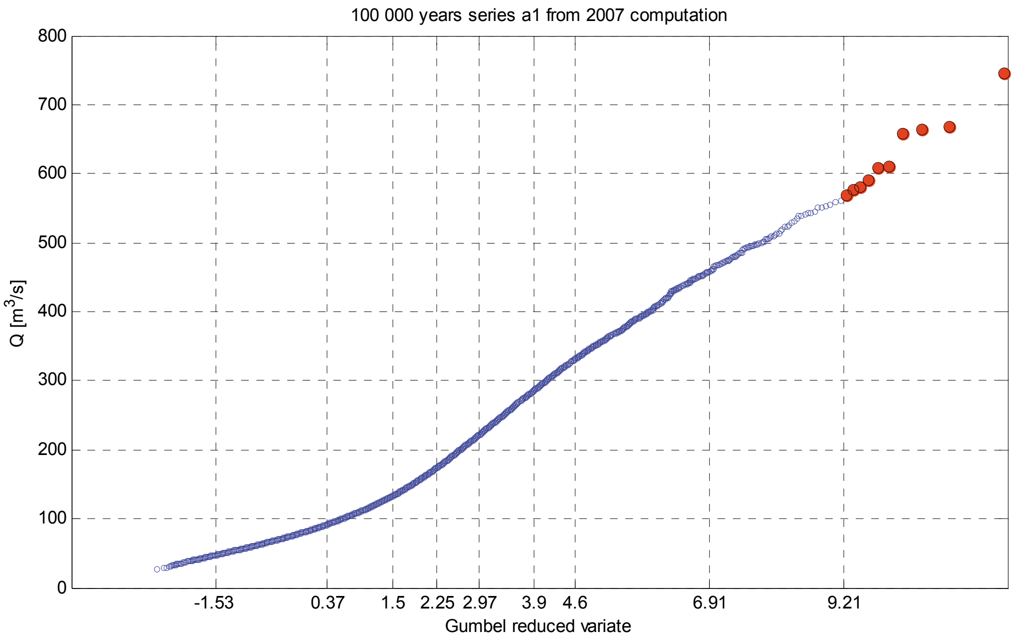



**Figure 5.6** Annual peaks from series a1 of the length of 100 thousand years

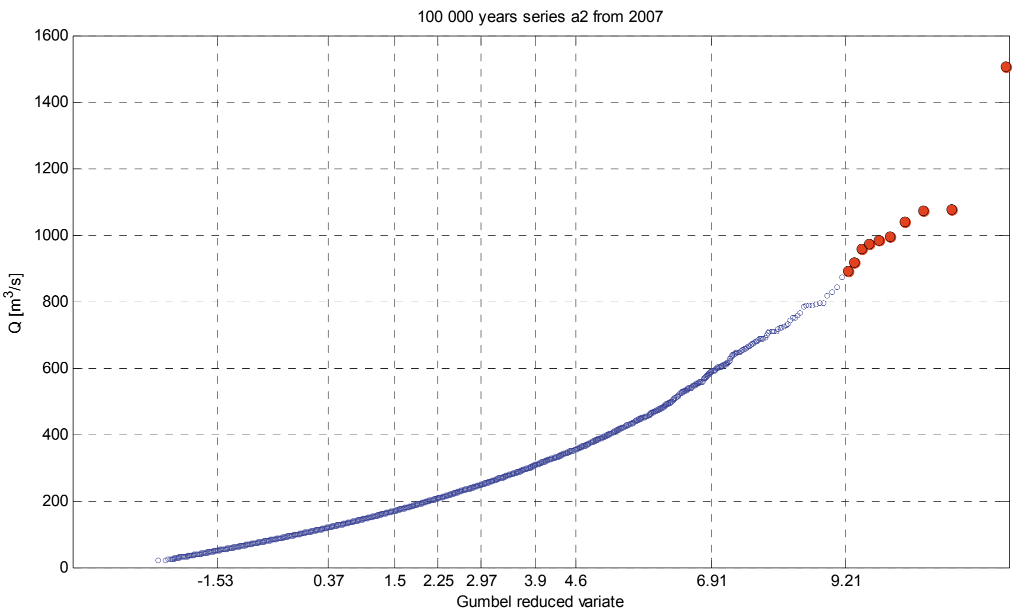



**Figure 5.7** Annual peaks from series a2 of the length of 100 thousand years

**Acknowledgements.** Czech data was provided by the Czech Hydrometeorological Institute (CHMI). An important part of the simulations has been carried out at the Lancaster University parallel system. Data from Eger and Roslau have been provided by Landesamt für Wasserwirtschaft München partly through the Commission for boundary streams. German meteorological data was provided by Deutscher Wetterdienst, München. The study was supported by the Ministry of Environment of the Czech Republic under the grants SP/2e7/229/07, MZP0002071101 and its previous grants and by the Ministry of Education, Youth and Sports which partly supported WATCH project (grant 7A08036). The continued development of GLUE has been supported by NERC long term grant NER/L/S/2001/00658. KJB was supported for a year as Konung Carl XVI Gustafs Gästprofessur I Miljövetenskap at Uppsala University. SDB is grateful to Mehran Hojati for the help with the fuzzy regression implementation.

## References

Beven, KJ. (1986a) Hillslope runoff processes and flood frequency characteristics. In Abrahams, AD. (Ed.) *Hillslope Processes*, 187–202, Allen and Unwin, Boston.

Beven, KJ. (1986b) Runoff production and flood frequency in catchments of order n: an alternative approach. In Gupta VK., Rodriguez-Iturbe I., and Wood EF. (Eds) *Scale Problems in Hydrology*, Reidel, Dordrecht, 107–131.

Beven, KJ. (1987) Towards the use of catchment geomorphology in flood frequency predictions. *Earth Surf. Process. Landf.,* V12(1), 69–82.

Beven, KJ. (2006) A manifesto for the equifinality thesis. *J. Hydrology*, 320, 18–36.

Beven, KJ. (2008) Environmental Modelling: An Uncertain Future? Routledge : London, in press.

Beven, KJ. and Binley, AM. (1992) The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes*, 6, 279–298.

Beven, KJ. and Freer, J. (2001) Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems. *J. Hydrology,* 249, 11–29.

Blazkova, S. and Beven, KJ. (1995) Frequency version of TOPMODEL as a tool for assessing the impact of climate variability on flow sources and flood peaks. *J. Hydrol. Hydromech.*, 43, 392–411.

Blazkova, S. and Beven, KJ. (1997) Flood Frequency Prediction for data limited catchments in the Czech Republic using a stochastic rainfall model and TOPMODEL. *J. Hydrol.,*195, 256–278.

Blazkova, S. and Beven, KJ. (2002) Flood Frequency Estimation by Continuous Simulation for a Catchment treated as Ungauged (with Uncertainty). *Water Resources Research,* 38, 8, 10.1029/2001WR000500.

Blazkova, S. and Beven, KJ. (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. Hydrology*, 292, 153–172.

Blazkova, S. and Beven, K. (2009) A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation: Skalka catchment, Czech Republic. *Water Resour. Res.*, 45, W00B16, doi:10.1029/2007WR006726.

Calver, A., Lamb, R., and Morris, SE. (1999), River flood frequency estimation using continuous runoff modelling. *Proc. Instn. Civ. Engrs., Water, Maritime, Energy*, 136, 225–234.

Cameron, D., Beven, KJ., Tawn, J., Blazkova, S., and Naden, P. (1999) Flood frequency estimation by continuous simulation for a gauged upland catchment (with uncertainty). *J. Hydrology*, 219, 169–187.

Cameron, D., Beven, K., and Naden, P. (2000) Flood frequency estimation under climate change (with uncertainty). *Hydrology and Earth System Sciences*, 4, 3, 393–405.

Cameron, D., Beven, KJ., and Tawn, J. (2001) Modelling extreme rainfalls using a modified random pulse Bartlett-Lewis stochastic rainfall model (with uncertainty). *Adv. Water Resour.,* 24, 203–211.

Cameron, D. (2006) An application of the UKCIP02 climate change scenarios to flood estimation by continuous simulation for a gauged catchment in the northeast of Scotland, UK (with uncertainty). *J. Hydrol.,* 328, 1, 212–226.

Gupta, HV., Sorooshian, S., and Yapo, PO. (1998) Towards improved calibration of hydrologic models: multiple and incommensurable measures of information. *Water Resour. Res.*, 34, 751–763.

Hojati, M., Bector, CR., and Smimou, K. (2005) A simple method for computation of fuzzy linear regression. *European Journal of Operational Research,* 166, 172–184.

Hosking, JRM. (1997) Fortran routines for use with the method of L-moments, Version 3.02. *IBM Res. Rep. RC 20525 (90933)*, IBM Research Division, Almaden.

# 6 A new model for the dynamics of runoff derived from information on river network and shape of catchment area

Thomas Skaugen

*A new principle of discharge formation is put forward. The dynamics of discharge is considered as a function of river network density and shape of catchment. Total discharge is the sum of discharge from several layers each with different, and fixed, velocities. For a fixed time interval, velocity and for each layer, a certain fraction of the catchment is drained into the river network. This fraction is estimated as an area of a zone, whose width, perpendicular to the river network, is determined for the time interval of interest, by the velocity. Such zones, or contributing areas, are considered to be horizontally adjacent to each other and will eventually cover the whole catchment. The dynamical properties of discharge can thus be parameterised from information derived from maps. A model is developed using the soil water and snow routine from the HBV model, but using the proposed principle for modelling the dynamics of discharge. The model performs as well as the traditional HBV model and modest testing indicates promising properties for PUB applications.*

## 6.1 Introduction

In this study we aim at describing the dynamics of discharge as a function of river network and the shape of the catchment. If we consider Figure 6.1, we have a simple square catchment with a river going through in which the water flows with a velocity $k$. Lateral inflow with velocity $k0$ occurs along the river. If we consider the lateral inflow to be of a Darcian type, as we do for real catchments, the differences in velocity between $k$ and $k0$ is by several orders of magnitude. The foundation of the method presented in this study is that the dynamics of discharge measured at the outlet is determined by the velocity of water in the soils and the distance water needs to travel in order to reach the river.

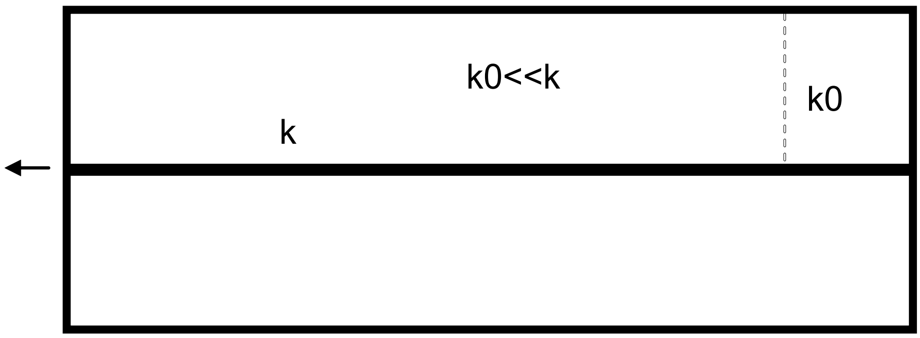



**Figure 6.1** Rectangular catchment with a river at the centre. The velocity of water in the river is $k$ and the lateral inflow from the soils is with velocity $k0$. $k0 << k$

This line of reasoning is partly inspired from constructal theory (Bejan, 2000) which discusses optimal shapes and structures of natural system such as river networks, ventilation channels in bee-swarms etc. One of the principles taken into account in developing the theory presented in this study is: *The flow path in a volume to point problem is optimized in such a way that resistance to flow is minimal (Bejan, 2000, p. 52)*. This implies that the river network we find in natural catchments is the result of an optimization process of the volume to point problem. We should thus find important information on the rainfall runoff system of the catchment by investigating the river network. This approach has led to the finding of a new type of catchment characteristic, which, in the authors view, has a very promising potential.

### 6.2 Principles of discharge formation

Let us assume that a uniform rainfall occurs over a catchment and water starts to flow towards the river network with a uniform velocity. The water that drains into the river during a time interval has travelled a distance which defines a zone around the river network. For several time intervals we have several horizontally adjacent zones, which eventually cover the entire catchment area.

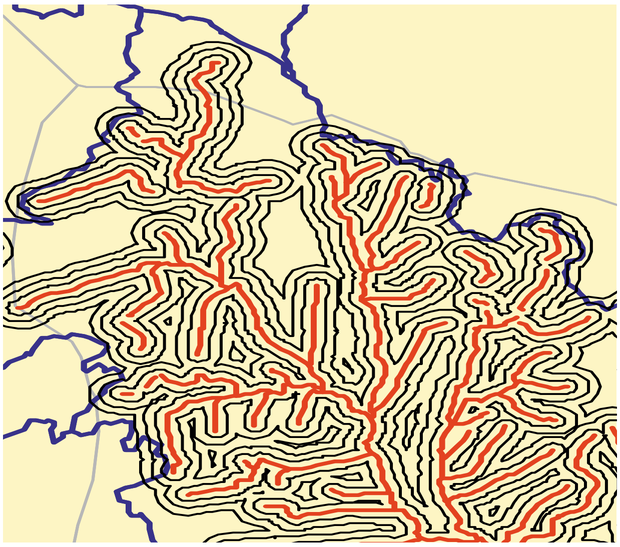


**Figure 6.2** The river network is shown in red and we see three zones with a width of 200 metres each. The width corresponds to a velocity of $0.0023\ m/s$ for the duration of one day (86,400 seconds). We find that the area of the zones decrease as the time intervals increase (zones further away from the river). Blue line represents catchment divide

Each of these zones defines an area which we consider the active drainage area for the specific time interval. When considering all the consecutive zones we have different active drainage area for each time interval, and this active area tends to decrease as the time interval increases until the whole catchment area has drained. We denote the areas contributing area, and Figure 6.3 shows how the cumulative distribution (with respect to time) of contributing area for two catchments Eggedal (308 km$^2$) and Narsjø (119 km$^2$). The functions in Figure 6.3 inform us of two features: 1) By fixing the time interval and the velocity, thus fixing the width of the zones, the number of time intervals before the whole catchment has contributed is found, and 2) the contributing area pr. zone and time step. From Figure 6.3 we observe how much faster the contributing area increases, i.e. how the water escapes much more easily to a river, for the catchment Eggedal than at Narsjø. The drainage density at Eggedal is about twice to that of Narsjø.

Let us further assume that the process described in the previous paragraph takes place in several horizontal layers, but with different velocities. The total discharge released to the river at a certain time interval will be the sum of discharge from the different layers.

Figure 6.4 illustrates the discharge from two layers, one rapid and one slow. Note that the catchment is drained for the rapid layer in about 10 time intervals whereas more than 30 time intervals is used for the slow layer.

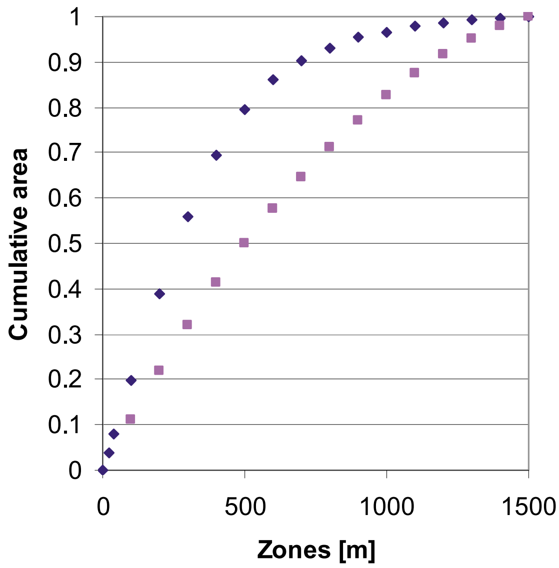



**Figure 6.3** Cumulative distribution of contributing area as a function of length of zones. Black diamonds for the catchment Eggedal and grey squares for the catchment Narsjø

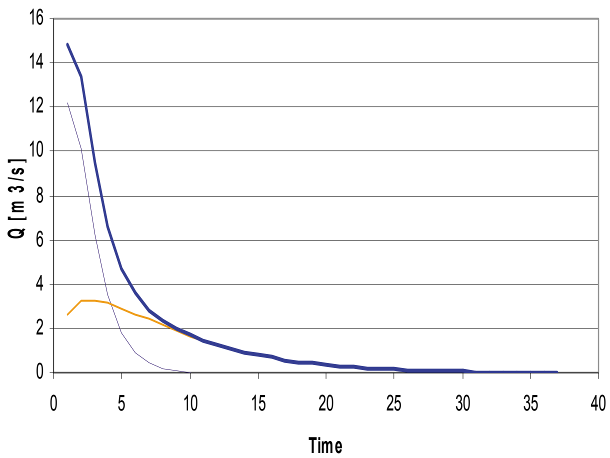



**Figure 6.4** Total discharge (thick black) and contribution from rapid layer (black thin line) and slow layer (grey thin line)

## 6.3 The model

Input to the model is precipitation and temperature. The snow routine consists of a degree-day model (Sælthun, 1996) for accumulation and melting of snow, and the spatial distribution of snow and the evolution of snow free areas are modelled according to Skaugen (2007).

The wetness in the catchment, $sm / fc$, where $sm$ is the soil moisture and $fc$ is field capacity, is calculated using the soil moisture routine of the HBV model (Bergstrøm, 1992). A fraction of the water from the snow routine is routed, non-linearly if preferred, through the soil moisture routine as a function of $sm / fc$. Here, unlike in the HBV model, a fixed percentage (2 %) of the soil moisture reservoir is continuously drained.

When modelling the dynamics of discharge, we assume, as do many authors, a fixed exponential distribution of velocities $k$ (Beven, 1982; Beldring et al., 2000). The velocity profiles stay fixed in the model and are thus, in principle possible to assume a priori from measurements.

The rainfall (snowmelt) that passes the soil moisture routine is distributed to the different layers according to an exponential distribution with a parameters determined by the wetness conditions ($sm / fc$). Figure 6.5 shows an example of how the parameter of the exponential distribution is determined. A negative value of $D\_dist$ indicates a dry situation and most of the water is distributed to the slow layers, whereas a positive $D\_dist$ indicates wet conditions and the rapid layers are favoured.

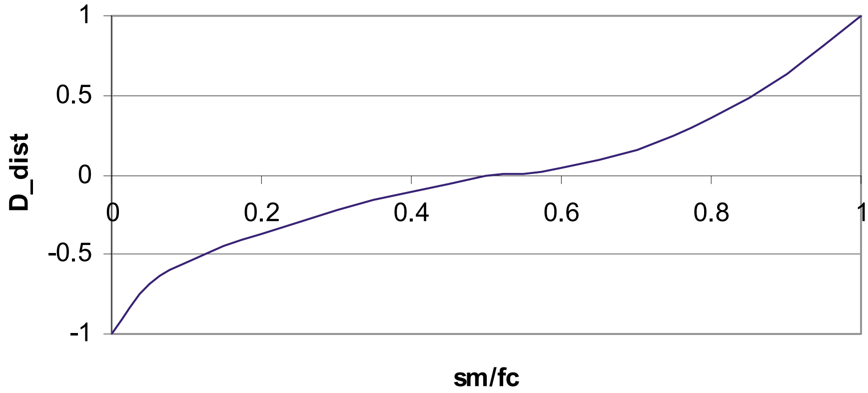



**Figure 6.5** Parameter for distributing rainfall/snowmelt to layers. A negative value of $D\_dist$ favours the slower layers, whereas a positive $D\_dist$ favours the rapid layers. $D\_dist$ equal to zero gives an uniform distribution of water to the layers

Discharge from each layer is calculated for a number of time intervals ahead, appropriate to the number of zones found to cover the whole catchment for the slowest layer, and total discharge for a specific time interval is the sum over all the layers.

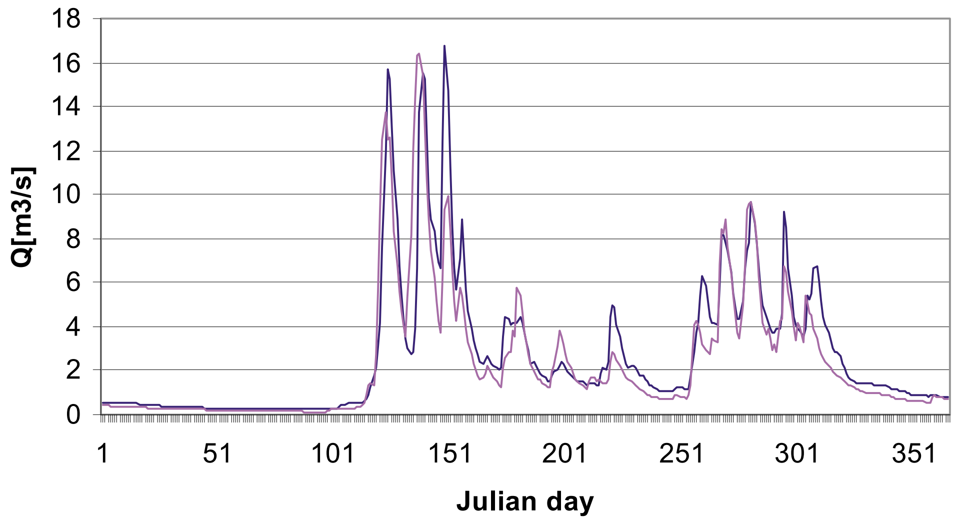



**Figure 6.6** Observed (black) and simulated (grey) discharge, for the Narsjø catchment

### 6.4 Results

We have tested the model for two catchments in Southern Norway, Eggedal (1978–1984) and Narsjø (1981–1986). The Nash-Sutcliffe criterion R2 was similar to that of the traditional HBV model used operationally in the national flood forecasting service (R2 = 0.76 for Eggedal and R2 = 0.86 for Narsjø). Figure 6.6 shows an example of a simulation for Narsjø (by no means the best simulation, but the complexity of the annual hydrograph is well captured). It can be noted that approximately 450 days was used to drain the slowest layer for the Eggedal catchment.

## 6.5 Discussion and conclusions

The model has, of course, a set of parameters used for calibrating the model, besides those who are catchment specific (like the function in Figure 6.3). There are a potential of 18 parameters to tune but very few are used. When the parameters calibrated for the simulation of the Eggedal catchment was used for simulating for the Narsjø catchment, we obtained a respectable R2 of 0.69. This is promising with respect to the PUB potential of the model.

The runoff dynamics in the version of the HBV model wich is used operationally at NVE is controlled by two linear reservoirs. One of the reservoirs has two outlets in order to better simulate quick responses, which makes the resulting runoff depart from that of a classical linear reservoir. The number of parameters is thus five, three time constants, one threshold for quick response, and finally one parameter distributing water between the two reservoirs. In the proposed model several parameters are associated with the dynamics of runoff, but only three parameters are subjects to tuning. The first parameter is the maximum velocity, the second parameter is that of the exponential distribution of velocity and finally there is the parameter that determines the function that distributes the water to the different layers conditioned on the wetness. The reduction in model parameters without loss of precision is an obvious improvement, and a step towards better predictions in ungauged basins.

The proposed principle of discharge dynamics questions the concept applied in some distributed models using grid cells as a semi-independent hydrological units. Due to the different velocities of the layers the water at a given grid cell and at a given time will consist of water entering the catchment at different locations upstream from the grid cell. The presented model concept has the potential for a very detailed mapping of moisture in a catchment. Figure 6.7 represents a slow layer in a strip of catchment perpendicular to the river. $A$ represents the area of the cells in the strip, and $D$ represents the input which is precipitation or snowmelt.

It takes 10 time steps to drain the slow layer. A faster layer will, of course have less but larger cells. The slow layer defines the smallest spatial resolution parallel to the river whereas the resolution perpendicular to river can be chosen arbitrarily (the change in area for the grid cells has to be taken into account if the perpendicular resolution becomes too large). For an arbitrary cell, the water content is the sum of the previous events which entered the catchment in cells upstream. The water in, for example, cell A2 for the slow layer (see Figure 6.7) is the sum represented by the diagonal indicated in the figure and is thus:

$Q(A2) = A(D1 + 0 + D3 + 0 + 0 + D6 + D7 + D8 + D9)$

For the $j$'th cell we have:

$$Q(A_j) = A(\sum_{i=1}^{N-j+1} D_i) \tag{1}$$

Here we have assumed that the size of the areas, $A$, are all equal. If they are not, the input $D_i$ is scaled by $\frac{A_i}{A_j}$ where $i$ is the cell in which the input entered the catchment, and $j$ is the catchment for which are estimating the water content. If different sizes of $A$ is taken into account, the expression (1) will take the form:

$$Q(A_j) = A_j(\sum_{i=1}^{N-j+1} D_i \frac{A_i}{A_j}) = \sum_{i=1}^{N-j+1} D_i A_i \tag{2}$$

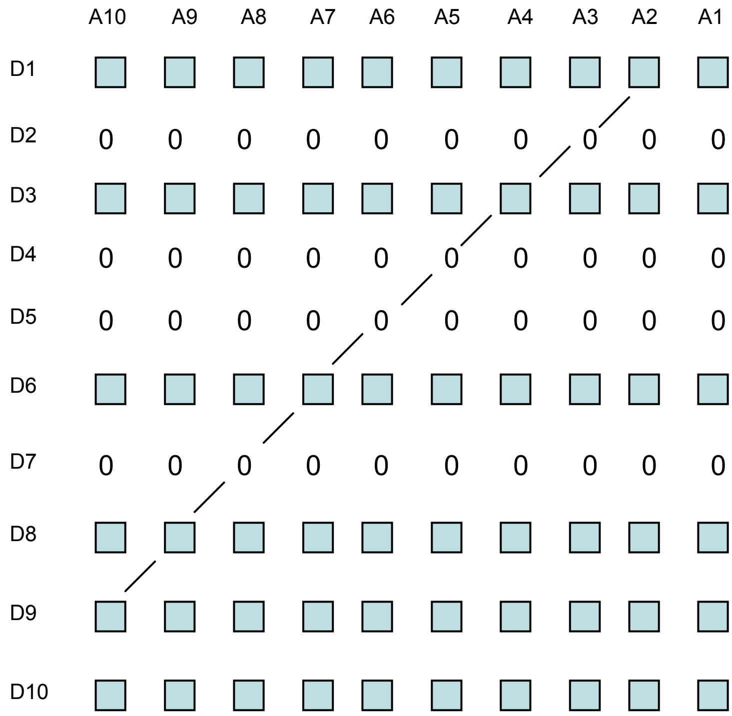



**Figure 6.7** The input D for 10 time steps for a strip of the catchment perpendicular to the river. Cell $A1$ is nearest to the river whereas $A10$ is farthest away from the river

The area of the target cell of the slowest layer represents fractions of the areas of that particular time step for the faster layers. These fractions can then be summed and the water content at the resolution of the slowest layer can then be estimated. These fractions can easily be calculated with the estimated distribution of velocities. With a fixed distribution of velocity the spatial distribution of water can be estimated for the catchment once at a resolution of the slowest layer (or aggregated up to a more rapid layer of our choosing). The input D (precipitation, snowmelt) may of course be allowed to vary in space, and thus we have the potential of a very detailed spatial map of runoff.

Different temporal resolutions should not represent a problem with respect to the dynamics of the model. This feature remains, however, to be investigated.

There might be some water quality applications for the model as it has the potential to time the arrival of water from the different layers to the river network. This requires a distributed version of the model where each layer is mapped. Also this feature remains to be investigated.

## References


Bejan, A. (2000) Shape and Structure, from Engineering to Nature. Cambridge : Cambridge University Press, UK.

Beldring, S., Gottschalk, L., Rohde, A., and Tallaksen, LM. (2000) Kinematic wave approximations to hillslope hydrological processes in tills. *Hydrological Processes*, 14, 727–745.

Bergström, S. (1992) The HBV model–its structure and applications. SMHI Hydrology, RH no. 4, Norrköping, 35 p.

Beven, K. (1982) On subsurface stormflow: an analysis of response times. *Hydrological Sciences Journal*, 4, 505–521.

Skaugen, T. (2007) Modelling the spatial variability of snow water equivalent at the catchment scale. *Hydrology and Earth System Sciences (HESS*), 11, 1543–1550.

Sælthun, NR. (1996) The "Nordic" HBV model. Description and documentation of the model version developed for the project Climate Change and Energy Production. NVE Publication no 7-1996, Oslo, 26 p.

# 7 Comparison of uncertainties in extreme rainfall distribution using a GP distribution and the SHYPRE regional stochastic rainfall model

Michel Lang, Aurélie Muller, Patrick Arnaud, Jacques Lavabre

*The paper presents a comparison of sampling distribution and model parameter uncertainty on a long rainfall series, in Marseilles, between 1882 and 2003. Two probabilistic models have been used: a Generalized Pareto distribution (GPD) on rainfall events larger than 20 mm, and a stochastic rainfall model SHYPRE. The results show that SHYPRE uncertainties are smaller, due to the fact that the rainfall model is parameterized with a larger number of observed data.*

## 7.1 Introduction

Extreme rainfall analysis can be developed through extreme value theory (Beirlant et al., 2004), or the use of a stochastic rainfall generator. Several methods have been proposed to quantify modelling uncertainties (see Engeland et al., 2005 for a review). By combining prior knowledge on parameters with a likelihood function of data, the Bayesian approach provides an *a posteriori* distribution of the parameters. Simulation procedures based on the Markov Chains Monte Carlo (MCMC) method are then used for uncertainty analysis, by sampling parameters through their posterior distribution. The GLUE method, introduced by Beven and Binley (1992), formulates the hypothesis that several parameters can produce similar results depending on the calibration data, and defines a "subjective" likelihood measure, not necessarily the result of a probabilistic model. Cameron et al. (2000) presented an application of the GLUE method to a stochastic rainfall model of the Bartlett-Lewis type. In this paper, we focuss on sampling distribution and model parameter uncertainty, using the long data series from Marseilles, France.

## 7.2 The data and the models

The comparative study is based on a 122 year daily rainfall series in Marseilles between 1882 and 2003. The rate of missing values is lower than 10% for each year. We consider the summer season between June and November, when rainfall is highest in Marseilles. No significant monotonic trend was detected by the Mann-Kendall test at the 10% level (Muller et al., 2008). Such stationary results are in agreement with the European project IMFREX in the same field (Dubuisson and Moisselin, 2006). A Generalized Pareto Distribution (GPD):

$$P(X \le x \mid X \ge u) = 1-[1-k(x-u)/\alpha_u]^{1/k}, \quad \alpha_u>0, \ k \ne 0, \ x \text{ such as } 1-k(x-u)/\alpha_u>0 \qquad (1)$$

has been fitted to threshold daily rainfall values, separated by at least one day with a daily rainfall of less than 4 mm (requirement of independent events). Based on a test on the scale parameter (Lang et al., 1999; Coles, 2001), the threshold value $u$ = 20 mm provided 444 events in the 122 year-period.

The SHYPRE hourly rainfall generator is of the aggregation type (Cernesson et al., 1996; Arnaud and Lavabre, 1999, 2002). A rainfall event is first defined by a daily value larger than 20 mm, and consecutive daily rainfall events greater than 4 mm. Each event is then analyzed hourly, and described by storms. A storm is an elementary hyetograph with only one maximum. A rainfall event is then defined by several descriptive variables: the number of storms (clustered or not), the dry period between storms, and the duration, depth and profile of each storm. The persistence of rainstorms within a rainy event was observed to increase with an increase in the depth of successive storms, and consequently empirically modelled. Hourly hyetographs are then simulated from the descriptive variables and their respective distributions. The model generates long series of independent events (for example, 1,000 series of 500 years). The distributions of the annual maximum rainfall depth and of the excesses of a high threshold (20 mm) are then deduced for durations ranging between one hour and 72 hours. The year is separated into two seasons: summer (from June to November) and winter (from December to May). The seasons were defined for hydrological and climatic reasons (there are more convective events in summer and more frontal events in winter). In this paper, we use a simplified version of the model with only three local parameters taken from daily rainfall data, and regional values for the other parameters. The three local parameters are: *Ne,* the mean annual number of rainy events; *Pjx,* the mean of maximum daily rainfall in an event (in mm); *Dtot,* the mean duration of an event (in days).

### 7.3 Sampling uncertainty analysis

**Case of GP distribution.** The parameters of the GPD have been assessed by the weighted moments, with a covariance matrix approximated by (Hosking and Wallis, 1987):

$$1/[n(1+2k)(3+2k)]\begin{pmatrix} \alpha^2(7+18k+11k^2+3k^3) & \alpha(2+k)(2+6k+7k^2+2k^3) \\ \alpha(2+k)(2+6k+7k^2+2k^3) & (1+k)(2+k)^2(1+k+2k^2) \end{pmatrix} \tag{2}$$

where $n$ is the number of exceedances used in the estimation. The sampling distributions of quantiles are deduced from those of the parameters. In practice, the parameter estimates are simulated according to their sampling distribution, using the following method:

$$\hat{\alpha} = \hat{\alpha}_0 + Z_1\sqrt{Var(\hat{\alpha}_0)} \quad and \quad \hat{k} = \hat{k}_0 + (\rho Z_1 + (1-\rho^2)^{1/2} Z_2)\sqrt{Var(\hat{k}_0)} \tag{3}$$

where $\hat{\alpha}_0, \hat{k}_0$ are the weighted moment estimates of the parameters $\alpha$ and $k$; $Var(\hat{\alpha}_0)$, $Var(\hat{k}_0)$ are the estimator variances, estimated in $\hat{\alpha}_0, \hat{k}_0$ ; $Z_1$, $Z_2$ are standard Gaussian independent distributions and $\rho = Cov(\hat{\alpha}, \hat{k}) / \sqrt{Var(\hat{\alpha}_0)Var(\hat{k}_0)}$. This simulation method avoids the computation of the square root of the covariance matrix, which can sometimes lead to numerical error due to near zero values.

**Case of the SHYPRE hourly rainfall generator.** Since the SHYPRE parameters are means of a high number of variables, their sampling distributions can be assumed to be Gaussian. In the Marseilles series, the estimated correlation between parameters is 0.22 between *Pjx* and *Dtot*, and lower than 5% between *Ne* and the other two parameters. The sampling distribution of the parameters is then modelled by a three dimensional Gaussian distribution with independent margins centered on the estimates of the parameters, and with variances equal to the sampling variances of the parameter estimators. Rainfall events are then simulated by SHYPRE with different parameters (*Ne, Pjx*, *Dtot*) sampled in their sampling distribution. Then after a large number of parameter simulations, the median and confidence intervals of quantiles are deduced. Here, the quantiles of interest are those of the daily rainfall above a defined threshold (20 mm).

**Application to the Marseilles series.** Figure 7.1 shows that the confidence intervals of the simulated quantiles of SHYPRE are slimmer than those of the GP distribution. This is due to the fact that SHYPRE parameter estimators are sample means, whereas the GPD parameter estimators depend on the weighted moments of order 0 (which corresponds to the sample mean) and 1 (which corresponds to the weighted mean). Figure 7.2 gives a comparison of the medians and confidence intervals of the quantiles estimated by GPD and simulated by SHYPRE, considering six successive sub-series of 20 years each. SHYPRE appears to be less sensitive to sampling uncertainties than GPD. The next section deals with the parameterization uncertainties of both the SHYPRE and GPD models.

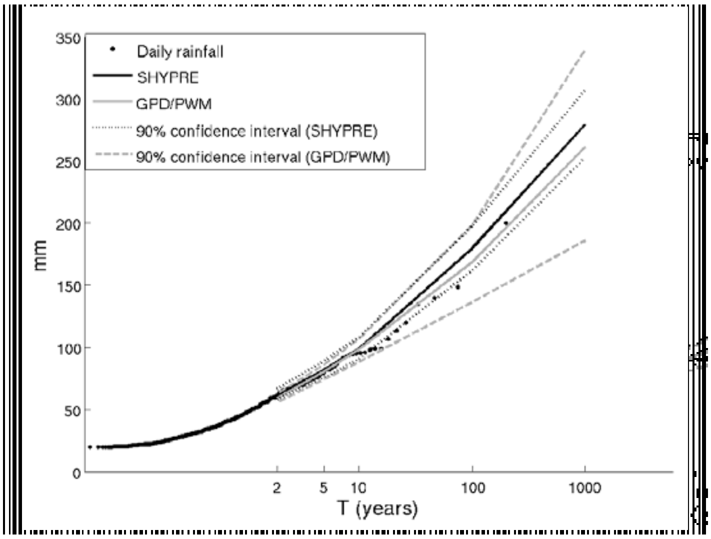



**Figure 7.1** 90% confidence intervals of the quantiles of daily rainfall exceedances of 20 mm, for the June–November season, with the probabilistic GPD model and the stochastic SHYPRE model

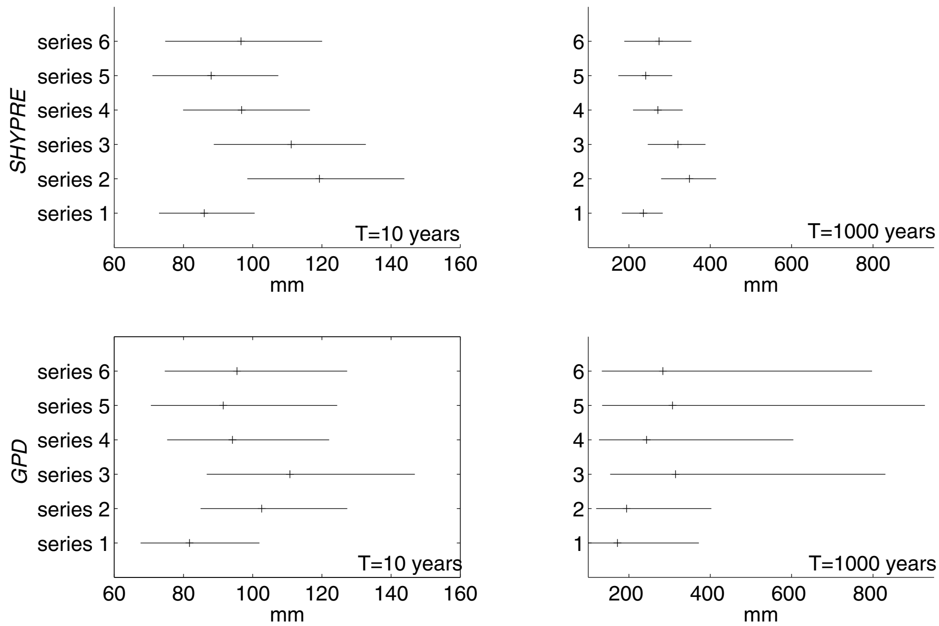



**Figure 7.2** Estimates and 90% confidence intervals of the quantiles of daily rainfall exceedances of 20 mm (10 and 1000-year return periods) for the GP distribution and the SHYPRE model, for each 20-year sub-series.

## 7.4 Parameterization uncertainties with a Bayesian analysis

**Bayesian approach.** In the Bayesian framework, the confidence intervals of the parameters are called credibility intervals. Unlike frequency analysis, which supposes fixed parameters and random estimators, Bayesian analysis supposes that parameters are random (with a prior distribution, independent of the observed data, and provided by expert knowledge); it gives *a posterior* probability to each parameter set, providing a useful inference on the model uncertainties. The simulated and observed data used for comparison are daily data higher than 20 mm. Heuristically, since the sample size is large (444), the posterior distribution depends much more on the data than does the prior distribution of the parameters. If this heuristic value is true, the posterior distribution should be near the sampling distribution obtained in a frequency analysis (Le Cam, 1953).

The prior distribution $\pi(\boldsymbol{\theta})$ of the parameters $\boldsymbol{\theta}$ should be defined without using data $\mathbf{X}$ for inference. The prior distributions used for the GPD parameters $\boldsymbol{\theta} = (\alpha, k)$ are the lognormal distribution with parameters 0 and 100 for $\alpha$ and the uniform distribution between –1 and 1 for $k$. These prior distributions are similar to those of Coles and Pericchi (2003) for the Generalized Extreme Value distribution parameters. The regional study of the local SHYPRE parameters $\boldsymbol{\theta}$ = (*Ne, Pjx*, *Dtot*), in a range of 50 km around Marseilles, enabled the definition of a uniform prior with the following ranges: $Ne \in [3; 6]$; $Pjx \in [33; 45]$; $Dtot \in [1.4; 2.1]$. Therefore, the posterior distribution of the parameters is obtained by combining the prior distribution $\pi(\boldsymbol{\theta})$ with the likelihood of the observations $p(\boldsymbol{X}|\ \boldsymbol{\theta})$:

$$p(\boldsymbol{\theta}|\boldsymbol{X}) = \pi(\boldsymbol{\theta}) p(\boldsymbol{X}|\boldsymbol{\theta}) / \int p(\boldsymbol{X}|\boldsymbol{\theta}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta} \qquad (4)$$

**MCMC Simulation.** In practice, the posterior distribution resulting from equation (4) is estimated by simulation using an adaptative Metropolis sampler, by combining the Metropolis-Hastings algorithm (Hastings, 1970) with the Gibbs sampler (Geman and Geman, 1984). The different steps of the algorithm are clearly described in a paper by Renard et al. (2006). The convergence of the algorithms was checked with Gelman's $R$ statistic (Gelman et al., 1997), calculated on six parallel sequences with starting points sampled in the prior distribution of the parameters. $R$ is equal to 1.000 after the 50,000 first iterations of the Metropolis algorithm with both the SHYPRE and the GPD model. The last 30,000 simulations are then supposed to be representative of the result of the posterior distribution of the parameters. The predicted quantiles are then calculated for every simulated $\boldsymbol{\theta}$ parameter set, giving an estimation of the predicted distribution of the quantiles. The likelihood measure is expressed by using a GP density:

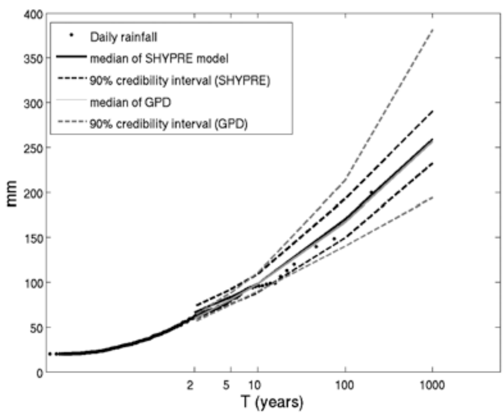
$$p(\boldsymbol{X}|\boldsymbol{\theta}) = \prod_{i=1}^{n} \left[1 - k(x_i - 20)/\alpha\right]^{1/k-1} / \alpha \qquad (5)$$

In the case of the SHYPRE model, as the parameter set $\boldsymbol{\theta}$ = (*Ne, Pjx*, *Dtot*) is not directly related to a ($\alpha$, $k$) pair, a preliminary correspondence was investigated. The parameter space of $\boldsymbol{\theta}$ values was sampled using a uniform grid for the parameters (*Pjx*, *Dtot*) with a step ($\Delta Pjx$ = 2mm, $\Delta dtot$ = 0.5 day), and $Ne$ = 20. Consequently, a SHYPRE data series was simulated for each $\boldsymbol{\theta}$ value of the grid, providing 35 quantiles $y_i$ . A corresponding GP distribution was fitted, with a ($\alpha_s$, $k_s$) pair, by minimization of sum $S$ in equation (6):

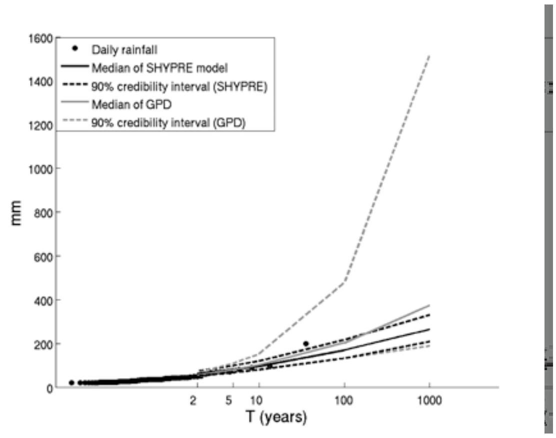
$$S = \sum_i \left\{ \left[ F_{th}(y_i) / (1 - F_{th}(y_i)) \right] \left[ F_{th}(y_i) - \hat{F}(y_i) \right] \right\}^2 \quad (6)$$

where $F_{th}(x) = 1 - \left[ 1 - k_s(x - 20)/\alpha_s \right]^{1/k_s}$ is the probability distribution of the GP distribution with parameters ($\alpha_s$, $k_s$), $\hat{F}(x)$ is the empirical non-exceedance probability associated to the value *x,* and ($y_i$, $i$ = 1 to 35) are the simulated SHYPRE quantiles: nine quantiles have a return period greater than 100 years, 21 quantiles have a return period lower than 10 years. The influence of parameter *Ne* is straightforward as there is a simple translation between two simulated SHYPRE distributions with the same (*Pjx*, *Dtot*) parameters. Finally, the MCMC simulation is carried out using equation (4), where the likelihood measure is expressed by equation (5). Every sampled $\boldsymbol{\theta}$ value of the MCMC algorithm is converted into a ($\alpha$, $k$) pair, using an interpolation between the nearest points of the uniform grid.

**Results with the whole data series.** As shown in Figure 7.3a, in the case of GPD, the credibility intervals are nearly the same as the confidence intervals obtained by frequency analysis in the section 3, confirming the heuristic hypothesis about the asymptotic behavior of the posterior distribution of the quantiles. In the case of SHYPRE, the marginal posterior distributions of the *Ne* and *Dtot* parameters, obtained after MCMC simulations, are similar to their uniform prior distributions. The posterior distribution of *Pjx* is closer to a normal distribution with a mean of 40.3 mm and a standard deviation of 1.00 mm, than to its sampling distribution described in the previous section. This result shows that the SHYPRE calibration is not sensitive to *Ne* and *Dtot* parameters and depends only on the parameter *Pjx,* which displays slight variability. The credibility intervals of the quantiles are then slim. The estimated median is close to observed data which demonstrates the ability of SHYPRE to reproduce observations. It was important to check this point, because SHYPRE is neither fitted nor calibrated to the daily data. SHYPRE is an hourly hyetograph generator, through parameters (*Ne,Pjx,Dtot*). The simulated hourly hyetographs allow estimation of rainfall distribution for time periods of between 1 and 72 hours. Our study was not extended to hourly data, but these can be reproduced using the same method.




(a) Case of a long series from Marseilles 1882–2003




(b) case of a short series (1983–2003 in Marseilles)

**Figure 7.3** 90% credibility intervals of quantiles, simulated by SHYPRE and estimated by the GPD

**Results with the 20 last years.** Here, the aim is to consider parameterization uncertainty and sampling effect together. To this end, the Bayesian study is reproduced with the 20 last years of the series (1983–2003, actually 21 years but the year 2001 contains no exceedance of 20 mm), with the following modification: with 20 years of data, the tail of the simulated distribution of SHYPRE can be very uncertain. The method of fitting a GPD to SHYPRE distribution was thus modified: the sum $S$ in equation (6) takes only quantiles $y_i$ whose return period is lower than 10 years. The quality of the fit was checked and was found to be satisfactory (except for return periods of more than 10 years since this part of the distribution was not considered in the fitting procedure). The convergence of the MCMC algorithms was also checked and was found to be satisfactory. The credibility intervals of the SHYPRE quantiles obtained here (cf. Figure 7.3b) are larger than the former since the data are less numerous. In the case of a short series, such intervals contain the data of the 20 last years and the former credibility intervals, which contain 443 of the 444 data of the whole series. In the case of GPD, the credibility interval is very large, giving unrealistic values (the observed values in the region of Marseilles are lower than 500 mm, and the daily rainfall record observed in France is 1,000 mm and the world record is 1,800 mm, observed in the Réunion Island). Figure 7.3b shows that the median estimate of the 1000 year quantile (374 mm) is outside the credibility interval of the SHYPRE quantiles. This proves that, with the prior distributions considered here, the GPD is highly variable and SHYPRE has low variability. However these results are highly dependent on the prior distributions of the parameters. For example, if the prior distribution of the GPD parameters is more precise, the credibility intervals are reduced.

## 7.5 Discussion and conclusions

This study assessed sampling uncertainties and model uncertainties of two models: the stochastic SHYPRE hourly rainfall generator and the probabilistic GPD model applied to a 122-year series of daily rainfall data from Marseilles. The sampling uncertainties were studied by frequency analysis, while the parameterization uncertainties were studied by Bayesian analysis, taking into account the randomness of the parameters. Both frequency and Bayesian analyses showed that SHYPRE is faithful to field observations, is relatively insensitive to sampling effect and has low variability. In contrast, GPD was shown to be highly sensitive to sampling effect, and has a high variability due to its sensitivity to the shape parameter. Future studies should compare the accuracy of quantile estimates from regional GPD analysis and SHYPRE simulations, in order to use regional a priori information in both cases.

**Acknowledgements.** Météo-France is gratefully acknowledged for providing the Marseilles rainfall series: the long daily rainfall series was reviewed by Météo-France in the framework of the European IMFREX project. The present work was financially supported by Cemagref, within the framework of PhD research conducted by A. Muller. This extended abstract will be published as a full paper in Hydrological Sciences Journal: “Uncertainties in extreme rainfall distribution using a stochastic rainfall model”, by Aurélie Muller, Patrick Arnaud, Michel Lang, Jacques Lavabre.

References

Arnaud, P. and Lavabre, J. (1999) Using a stochastic model for generating hourly hyetographs to study extreme rainfalls. *Hydrol. Sci. J.*, 44, 3, 443–445.

Arnaud, P. and Lavabre, J. (2002) Coupled rainfall model and discharge model for flood frequency estimation. *Water Resour. Res.*, 38, 6, 1075–1085.

Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. (2004) Statistics of Extremes, Theory and Applications. Wiley Series in Probability and Statistics.

Beven, KJ. and Binley, AM. (1992) The future of distributed models: model calibration and uncertainty prediction. *Hydrol. Processes*, 6, 279–298.

Cameron, D., Beven, K., and Tawn, J. (2000) Modelling extreme rainfalls using a modified random pulse Bartlett-Lewis stochastic rainfall model (with uncertainty). *Adv. Water Resour.*, 24, 2, 203–211.

Cernesson, F., Lavabre, J., and Masson, JM. (1996) Stochastic model for generating hourly hyetographs. *Atmos. Res.*, 42, 1–4, 149–161.

Coles, S. (2001) An introduction to statistical modelling of extreme values, London : Springer, 210 p.

Coles, S. and Pericchi, L. (2003) Anticipating catastrophes through extreme value modelling. *J. Royal Statist. Soc.* Series C : Applied Statistics, 52, 4, 405–416.

Dubuisson, B. and Moisselin, JM. (2006) Evolution des extrêmes climatiques en France à partir des séries observées. *Colloque SHF, Lyon, France,* 113–121.

Engeland, K., Xu, CY., and Gottschalk, L. (2005) Assessing uncertainties in a conceptual water balance model using Bayesian methodology. *Hydrol. Sci. J.*, 50, 1, 45–63.

Gelman, A., Carlin, JB., Stren, HS., and Rubin, DB. (1997) Bayesian Data Analysis. London : Chapman & Hall.

Geman, S. and Geman, D. (1984) Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6, 721–741.

Hastings, WK. (1970) Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57, 97–109.

Hosking, JRM. and Wallis, JR. (1987) Parameter and quantile estimation for the Generalized Pareto Distribution. *Technometrics*, 29, 3, 339–349.

Lang, M., Ouarda, T., and Bobée, B. (1999) Towards operational guidelines for over-threshold modelling. *J. Hydrol.*, 225, 103–117.

Le Cam, L. (1953) On some asymptotic properties of maximum likelihood estimates and related Bayes estimates. *Univ. California, Publications in Statistics,* 1, 11, 277–330.

Muller, A., Bacro, JN., and Lang, M. (2008) Bayesian comparison of different rainfall depth-duration-frequency relationships. *Stochastic Environmental Research and Risk Assessment,* 22, 1, 33–46.

Renard, B., Garreta, V., and Lang, M. (2006) An application of Bayesian analysis and Markov chain Monte Carlo methods to the estimation of a regional trend in annual maxima. *Water Resour. Res.,* 42, W12422, doi:10.1029/2005WR004591, 17 p.

# 8 Uncertainty in water balance simulation

Ladislav Kašpárek, Oldřich Novický

*The chapter describes approaches that can be used to reduce uncertainty in estimating parameters of water balance model by optimisation of its parameters in two steps. The first step includes parameters that affect substantially the total runoff while parameters which affect mainly the distribution of the total runoff among its components are calibrated in the second step. The calibration in the individual steps uses different optimisation criteria. In addition to fit between observed and simulated total runoff, it is suitable to include in the second step also fit between the simulated base flow and base flow derived from groundwater observation.*

## 8.1 Introduction

Water balance models are tools for assessing balance among water cycle components of a basin. The output of these models includes base flow, which is one of the components of total runoff. It is determined from balance relationships among the components of water cycle, which include precipitation and regional evaporation. The boundary between the base flow and interfow is uncertain and standard methods for model parameters estimation can provide less reliable estimates of base flow for some basins, e.g. for those having poor hydrogeological conditions.

The use of the information on base flow that has been derived independently on hydrologic balance from groundwater and flow observations can significantly improve reliability of the calibrated parameters of the model and the simulated water balance components.

## 8.2 Method

The proposed method was applied by its incorporation into algorithm of Bilan water balance model that has been developed by T. G. Masaryk Water Research Institute.

The model (described in detail e.g. in Tallaksen and Lanen, 2004) is a tool for assessing balance between water cycle components of a basin in a monthly step. The structure of the model is formed by a system of relationships describing basic principles of water balance on land surface, in the zone of aeration, including the effect of vegetation cover, and in groundwater. Air temperature is used as a main indicator of energy conditions, which affect significantly the water balance components.

The input data of the model include time series of basin precipitation, air temperature and other meteorological variables, which affect substantially the basin evaporation, such as relative air humidity. In calibrating the parameters of the model, runoff series at the outlet from the basin is used or, alternatively (in presented modification), we can also use results of observations of groundwater, i.e. groundwater levels at observation boreholes or spring yields.

The model generates time series of basin potential evapotranspiration, evaporation, infiltration to the zone of aeration, percolation of water towards the groundwater aquifer (groundwater recharge), and water storage components in the snow cover, zone of aeration (soil) and groundwater aquifer. The total runoff is generated as a sum of three components, which are direct (surface) runoff, interflow and base flow. It is assumed that the base flow is the groundwater component which outflows from the basin at its closing site together with the outflow from surface runoff.

In addition to several parameters, whose values are determined by physical conditions or are otherwise considered to be constants, the model has eight free parameters. The structure of the model ensures minimum overlapping of the individual parameters. The parameters of the model are calibrated by using an optimisation algorithm, which is aimed at attaining the best fit between the observed and simulated runoff series, for which several optimisation criteria are available.

The optimisation algorithm originally used the standard error of estimate (standard deviation between the observed and simulated runoff series) as an optimisation criterion. A drawback of this criterion is in the fact that its application does not ensure good fit between the observed and simulated runoff series in the area of low flows. This can be substantially improved by using a sum of relative deviations between the observed and simulated runoff series (relative means that individual deviations are divided by the mean runoff in the respective month) instead of the standard error of estimate. However, this criterion frequently deteriorates the fit in terms of the mean runoff and, therefore, an optimisation procedure combining these two criteria was developed.

The calibration of the parameters was executed in two steps. In the first step, the standard error of estimate is used as the optimisation criterion to calibrate four parameters that affect significantly the mean runoff. The remaining four parameters affecting the runoff distribution into its individual components are then calibrated by using the sum of relative deviations. It has been demonstrated by experimental calculations that this calibration procedure ensures mostly an acceptable fit in terms of both mean runoff and low flow runoff, which is formed predominantly by base flow.

Another method for reducing uncertainty of base flow simulation that has been developed and tested uses base flow series that has been derived independently on balance relationships among water cycle components by using results of groundwater observation. This additional information can be utilised in calculating the optimisation criterion, which applies fit between observed and simulated total runoff series together with fit between base flow calculated from the observed data and simulated by the model.
In calibrating the parameters of the model, it is suitable to calibrate firstly the parameters affecting substantially the total runoff by using the criterion of the best fit between the observed and simulated runoff, while the remaining parameters, involving those affecting mainly the distribution of the runoff between base flow and interflow, are calibrated by using some of the variables, which have been derived from the results of groundwater observation.

## 8.3 Results

The above method has been tested by using data from long-term observation, which has been carried out by T. G. Masaryk Water Research Institute in the Upper Metuje basin (Polická Cretaceous basin).

The results of groundwater level observation at VS3 borehole (under the Adršpach) in the period 1972–2000 were used for derivation of series of mean daily groundwater levels. Main part of the series was derived from water-level records, while the data from some short periods after 1997 were calculated by using linear interpolation of weekly observations.

The series of groundwater levels in VS3 borehole is illustrated in Figure 8.1, which shows that this series involves two periodical components, one representing an annual cycle, while the period of the other component is 8–10 years.

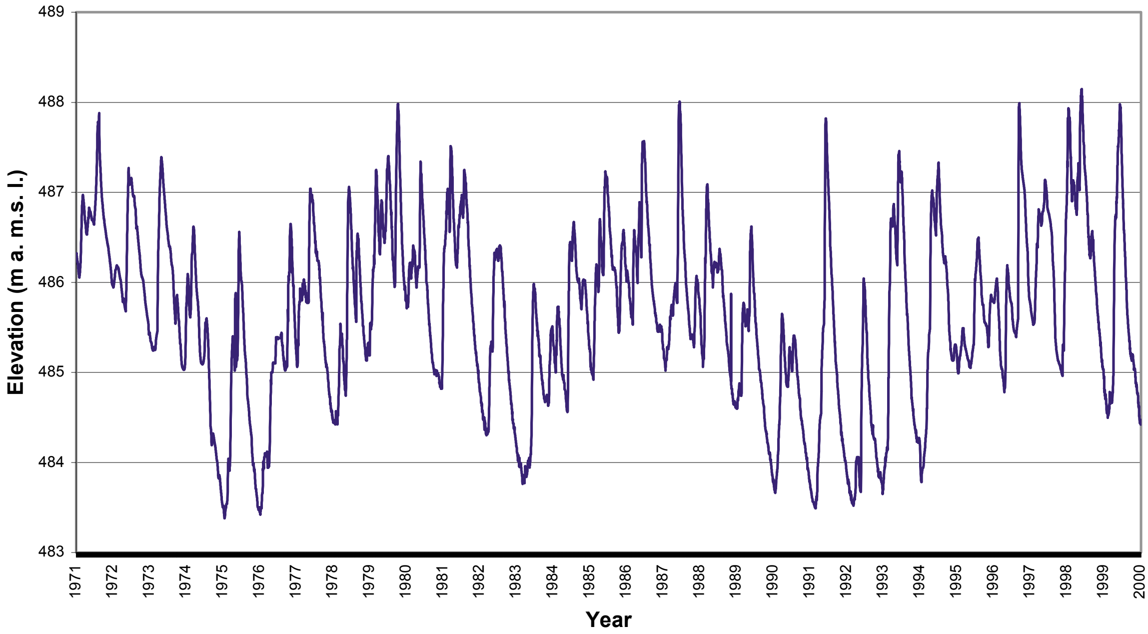



**Figure 8.1** Groundwater level in V3S borehole

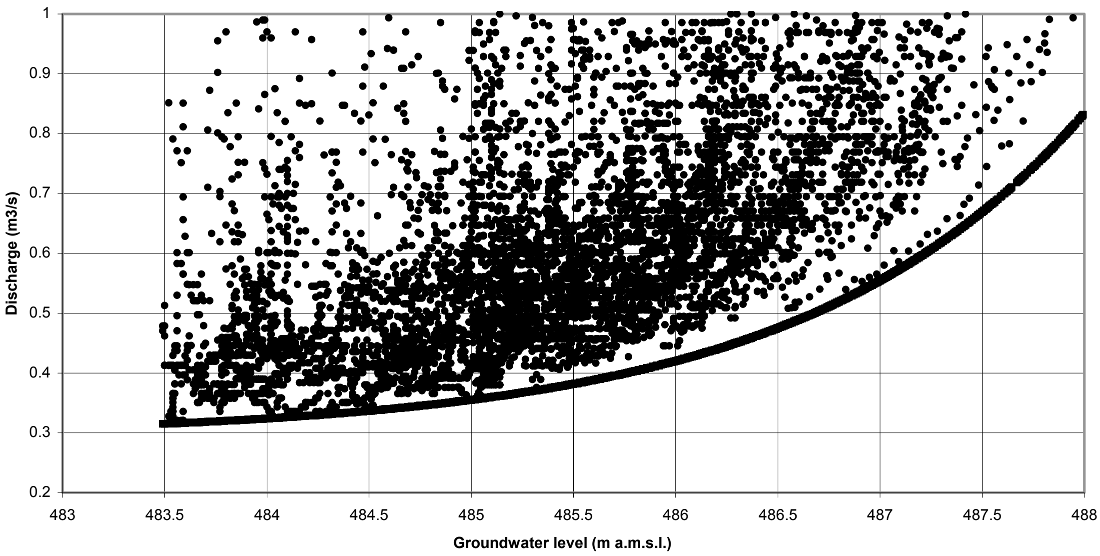



**Figure 8.2** Relationship between the groundwater level (VS3) and discharge of the Metuje River

Figure 8.2 shows a relationship between the groundwater level in the VS3 borehole and the discharge at M XII water gauging station, which closes 73.63 km$^2$ of the area of the Metuje River basin and monitors the flows from the Polická Cretaceous Basin. Data from the period 1977–2000 were used to derive a non-linear relationship, which was subsequently applied in accordance with the method developed by Kněžek (1974) for derivation of base flow from the results of the groundwater level observation. This equation has the form

$$Qz = 0.295 + 0.014 * e^{0.73 * (H - 483)} \tag{1}$$

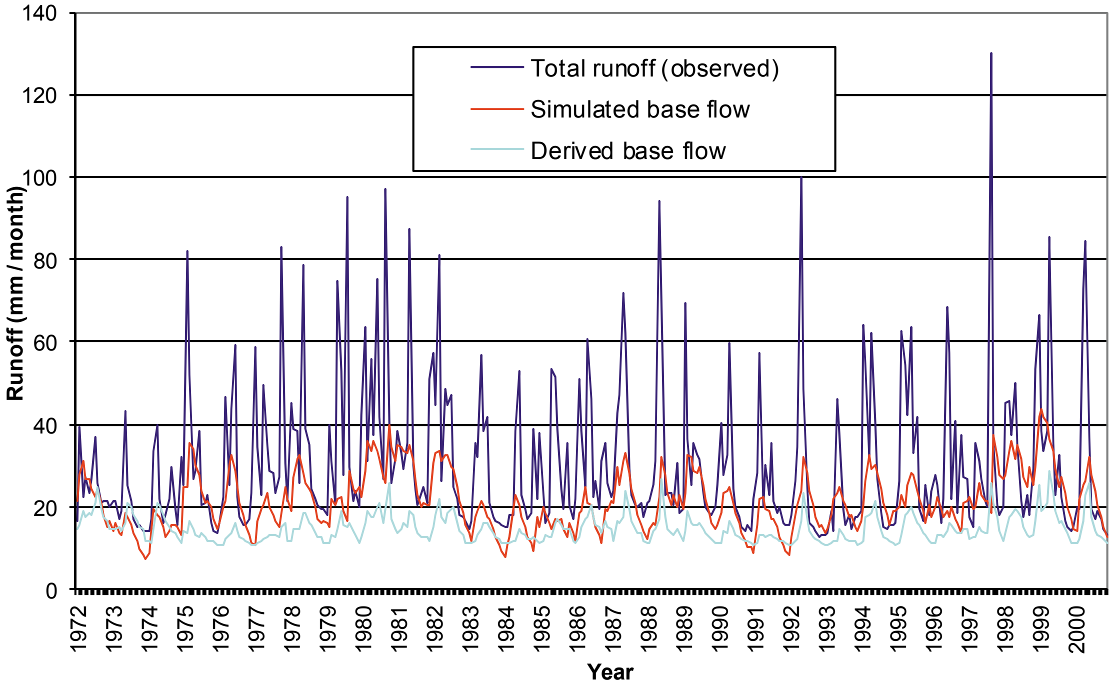



**Figure 8.3** Observed total runoff and base flow derived from groundwater level and simulated (Bilan) by using standard optimisation algorithm

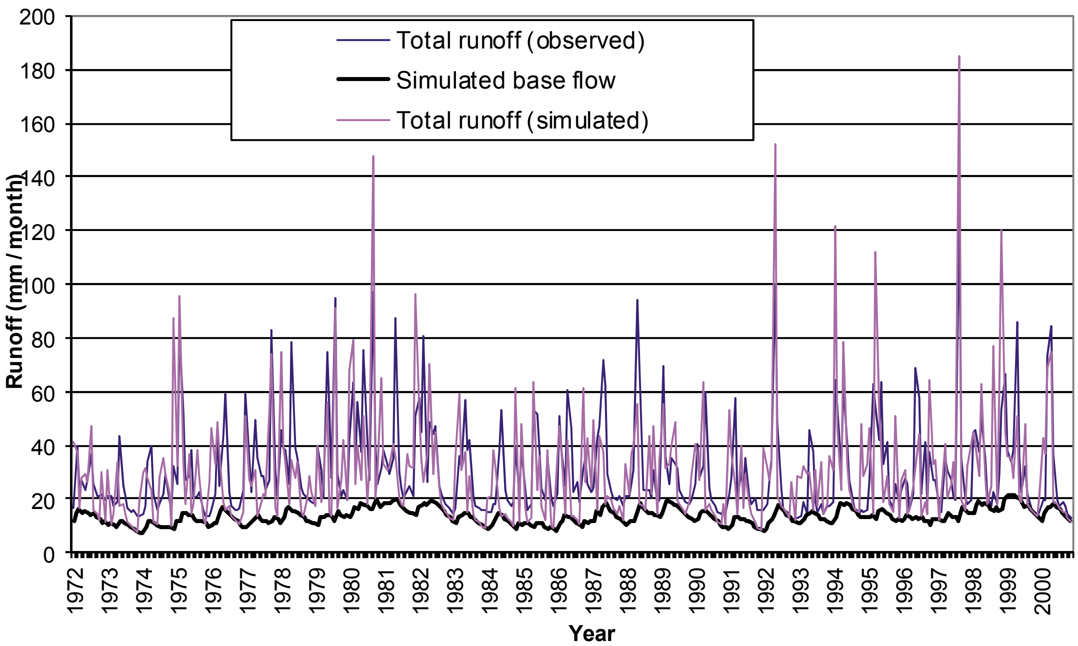



**Figure 8.4** Observed total runoff and base flow derived from groundwater level and simulated (Bilan) by using the developed method

For the simulation of the series of water cycle components, standard version of Bilan water balance model, involving calibration of the model parameters on the basis of optimisation of the fit between simulated and observed total runoff, was used in the first step.

Two base flow series (base flow series simulated by using the water balance model and the series derived from the groundwater level observation in the VS3 borehole) are shown in Figure 8.3 together with the observed total runoff. The figure shows that these approaches provide different results. The simulated base flow exceeds that derived from the groundwater level and its seasonal variability is higher, exhibiting particularly high values in the spring months. Some of the estimates exceed the observed total runoff.

The results indicate that the relationship, which was used to derive the base flow from the groundwater level observation, provides estimates which are mostly below the estimates from the standard version of the water balance model.

In a subsequent step, some of the parameters of the model were calibrated on the basis of the best fit between the base flow series which was simulated by the model and that derived by the developed method.

## 8.4 Discussion

The results of the application of the proposed method are illustrated in Figure 8.4, which shows the observed total runoff together with base flow that has been derived from groundwater level observation and base flow simulated by Bilan model by using the developed method. The figure illustrates that the number of months in which base flow exceeds the observed runoff decreased and therefore the simulated series of the base flow was substantially more reliable.

## 8.5 Conclusion

The method and its application described in the paper shows that model parameter uncertainty can be substantially reduced by developing approaches that use and integrate information from all observations available in the basin.

Results of the application of the proposed method for several basins in different hydrogeological conditions also indicated that its efficiency depends highly on accuracy of the observed series (groundwater level observation).

**Acknowledgements.** The work was supported by the WATCH project of the 6th framework and the Ministry of Education, Youth and Sports of the Czech Republic (7A08036).

References

Kněžek, M. (1974) Výzkum metod členění podzemních vod v povrchovém odtoku (Development of methods for separation of base flow from surface runoff). Research report, T.G.M. WRI, Prague.

Tallaksen, LM. and van Lanen, HAJ. (Eds) (2004) Hydrological Drought – Processes and Estimation Methods for Streamflow and Groundwater. Amsterdam, Developments in Water Sciences 48, Elsevier B.V.

# 9 Bayesian decision for low flow evaluation in non-stationary conditions

Mikhail Bolgov

*The reasons for the application of Bayesian methods are that it is desirable to define runoff characteristics more precisely taking into account information of territorial character and a necessity of obtaining estimations for non-stationary case. The result of using this method in the present contribution was that forecasted cumulative probabilistic curve of minimal water runoff, based on setting of various weights to estimations of average value for conditionally homogeneous periods and calculation of forecasted frequency distribution using formula of total probability, was obtained.*

Necessity of application of Bayesian methods is caused by two reasons. First, characteristics of runoff are desirable to be defined more precisely taking into account information of territorial character, i.e. to involve data about objects-analogues within the limits of homogeneous region. The second reason is a necessity of obtaining the estimations for non-stationary case when essentially different behaviour of river runoff process in different time intervals (periods) is observed and the forecast is characterized by high uncertainty.

Data of monitoring testify that essential changes of runoff occur in some cases on the territory of Russia ace, basically interannual values. Significant increase of runoff during winter low water period and reduction of runoff of spring flood is observed, i.e. redistribution of runoff within a year occur there. On the one hand, it simplifies the decision of water-economic problems, and on the other hand, it demands the development of methods of processing of non-stationary series of river runoff.

A lot of new problems demanding new methods and approaches for their decision arise in the non-stationary case.

First, it is necessary to find out "the character" of non-stationarity and to offer the general scheme of modelling. If process (sample) is characterized by the increase of fluctuation range it is possible to offer multiplicative model. If the average value changes it is advisable to use model with linear or nonlinear trend. In all cases there is a question to which today it is complicated to give the unequivocal answer: whether obtained trend reflects the unidirectional changes in development of process or it is a fragment of some low-frequency fluctuation in hydro-climatic system which would be observed in the form of the return tendency after some time.

It is necessary to note that available hydrological time series are too short for the construction of complex probability models using only statistical methods. It is impossible to construct a matrix of probabilities of transition from one stationary condition continuing during tens of years into another, which was observed also during decades, because we deal only with one case of condition change. Only if we study several tens of such cases it would be possible to speak both about distribution of probabilities of duration of system stable periods and about the estimation of matrix of transitive probabilities. As a rule, we have two-three conditions available, which compel us to apply the elementary hypotheses. We shall return below to their discussion.

Secondly, the division of long-time series into homogeneous parts leads to reduction of volume of independent data for short intervals and, correspondingly, to the increase of sample dispersion of estimations. This problem can be solved by grouping of data within the limits of homogeneous territories (method of generalization of sets or regionalization).

The analysis of changes of the minimal runoff on the rivers of Upper Don river basin shows that approximately from the beginning of 1980s there were essential changes in alimentation regime of the rivers. Minimal runoff has grown on all rivers of the region, and character of its fluctuations in time allows to make a conclusion that new conditionally stationary regime was generated there. It is possible to obtain statistical characteristics of runoff only for a new condition but it is desirable to make it under the condition of taking into account all available data of hydrological observations on the territory studied.

For more precise results of statistical processing it is advisable to use all information containing in the data set of the region. It is possible to use the Bayesian methods of estimation.

Let's discuss the problem of estimation of statistical characteristics of time series on the base of the Bayesian approach. We shall remind that we speak about estimation of parameters of runoff process for new conditionally stationary conditions. As a whole the stochastic model of runoff describing transitions from one level to another is rather complex therefore at the beginning we shall study only estimation of parameters of one-dimensional distributions of runoff.

Frequency distribution of runoff values is characterized, as a rule, by positive asymmetry and is approximated by the two- or three-parametrical Gamma-distribution. We shall limit our investigation to the two-parametrical Gamma-distribution in the following form:

$$P(x_0,\gamma,x)=\frac{\gamma^{\gamma}}{\Gamma(\gamma)}\left(\frac{x}{x_0}\right)^{\gamma-1}e^{-\gamma\frac{x}{x_0}} \tag{1}$$

where $x_0$ – average value, $\gamma = 1/C_v^2$, $C_v$ – variation coefficient, $\Gamma(\gamma)$ – Gamma-function.

The standard method of estimation of parameters using formula (1) supposes the search of the maximum of likelihood functions or applications of the method of moments. The Bayesian ideology is based on the assumption that some data about the estimated parameter, obtained before sample processing, is available for the statistician.

These data is available a priori, i.e. prior to the beginning of processing, and it is represented in the form of so-called aprioristic distribution of the parameter. An important point of Bayesian analysis is the formation of aprioristic distribution, and proposed approach consists in construction of using data of spatial distribution of the studied parameter, in other words using estimations of parameter for observational points located in homogeneous hydrological region. Generally the dispersion of spatial distribution of parameter should be corrected taking into account spatial correlation of observational data. However, it is necessary also for combined analysis to use the methods of S. N. Kritskiy and M. F. Menkel. At the current stage we shall neglect the spatial correlation of estimations of average values.

Let's accept ratio $\theta = \frac{\overline{Q_2}}{\overline{Q_1}}$, where $\overline{Q_1}$ and $\overline{Q_2}$ are average values of the studied characteristic for the first and second periods correspondingly, as the parameter for which Bayesian estimation is searched. In our case it will be the average values of minimal runoff for the two parts of the time series: from the beginning of the observations over river runoff to 1980, and from 1980 till the present time. The assumption of ratio $\theta$, which is a normalized characteristic, allows to bypass, in our opinion, a problem of the dependence of runoff on the area of drainage basin and on the other azonal factors.

Further calculations consist in the estimation of so-called aprioristic frequency distribution of parameter $\theta$ using the theorem of Bayes:

$$p(\theta / x) = \frac{P(x/\theta) \cdot p(\theta)}{\tilde{P}(x)} \tag{2}$$

where

$$\tilde{P}(x) = \int_{\theta} P(x/\theta) \cdot p(\theta) d\theta \tag{3}$$

*p(θ)* – aprioristic distribution of estimated parameter $\theta$, *P(x/θ)* – likelihood of data *(x)* at the present value of parameter $\theta$, and $\tilde{P}(x)$ calculated by the equation (3) plays a role of normalizing multiplier. In the equations (2) and (3) likelihood *P(x/θ)* is calculated as

$$P(x/\theta) = \prod_{i=1}^{n} p(x_i, \theta, \gamma) \tag{4}$$

where $p(x_i, \theta, \gamma)$ – two-parametrical Gamma-distribution (formula 1), for which parameter $\gamma$ is supposed to be known (for example, moment individual estimation), $n$ – length of processable part of time series, and $\theta$ – estimated parameter for which aprioristic frequency distribution is searched.

Aprioristic distribution of the parameter *p(θ)* is obtained by statistical processing of values $\theta$ for each observational point dividing time series into two parts (before 1980 and after 1980). Average value $\overline{\theta}$ and variation coefficient $Cv_\theta$ are estimated. Two-parametrical Gamma-distribution is recommended as aprioristic frequency distribution because the values $\theta$ cannot be negative.

Thus, all functions in the equations (2) and (3) are defined and it is possible to execute necessary calculations. The result of the calculations using the theorem of Bayes is a posteriori distribution of parameter $\theta$.

Further processing can be continued in two ways. First, we can accept one of the values $\theta$ as an estimated parameter, for example, position measure or the average of distribution. In some cases, for introduction of reserves into results of calculations it is possible to assume the quantile of a posteriori distribution $p(\theta / x)$.

The second way consists in the construction of forecasted frequency distribution of probabilities of water runoff on the base of the formula of total probability. In that case forecasted frequency is estimated as follows:

$$\pi(x) = \int_{\theta} P(x, \theta) \cdot p(\theta / x) d\theta \tag{5}$$

where $p(\theta / x)$ – a posteriori frequency distribution of parameter $\theta$, $P(x,\theta)$ – two-parametrical Gamma-distribution with preset parameter $\gamma$. As a result of numerical integration (5) we shall obtain Bayesian forecasted frequency distribution of probabilities of investigated runoff characteristics.

It is necessary to note that a posteriori frequency distribution of parameter $\theta$, obtained at current stage, not strongly differs from corresponding sample distribution of maximal likelihood estimation. It means that in that case observational data in a point prevail on group estimations. Besides the dispersion of aprioristic frequency distribution is insignificant, which also explains the obtained result. Probably in further research, it will be necessary to increase dispersion of aprioristic distribution $\theta$ taking into account spatial correlation of data.

Let's examine the problem of taking into account several conditionally stationary conditions in forecasting of runoff characteristics for the period of operation of projected object. First of all it is necessary to discuss the character of the stochastic model allowing to reproduce the moments of the change of the conditions of the process.

In the applied sections of the theory of casual processes stochastic models with several stationary conditions are known but in the case of the long-term fluctuations of runoff characteristics such models can be hardly applied. Physically proved hypotheses confirming presence of such stationary conditions are not proposed in up-to-date hydrology.

In our case only two conditions of the process are available and we can assume only that the system with the probabilities both $n_1/N$ and $n_2/N$ can be in one of them. Here $n_1 + n_2 = N$, where $N$ is total length of non-uniform sample.

Concerning future fluctuations of runoff we can only approve that with corresponding probabilities the system can be in one of the two conditions. In that case it is possible to assume that the average of distribution of the forecasted process is a combination of two distributions

$$\tilde{p}(\theta / x) = \frac{n_1}{N} \cdot \eta(\theta, n_1) + \frac{n_2}{N} p(\theta / x) \qquad (6)$$

where $\eta(\theta, n_1)$ – sample distribution of average value (average of distribution) for the first conditional-stationary part with length $n_1$, and $p(\theta / x)$ – a posteriori frequency distribution of Bayesian estimation $\theta$ for new climatic conditions. It is easy to notice that frequency distribution (6) is two-modal because the sample frequency of the average of distribution is Gaussian and a posteriori frequency $p(\theta / x)$ at small dispersion also is close to the normal law.

The final stage of Bayesian estimation (forecasting) consists in calculation of forecasted frequency distribution using the formula of total probability. In this case combination of distributions (6) will be presented as a posteriori distribution, and modelling distribution will be two-parametrical Gamma-distribution with parameter $\gamma$, equal to average value for both samples:

$$\pi(y) = \int_{\theta} P(y, \gamma, \theta) \cdot \tilde{p}(\theta / x) d\theta \qquad (7)$$

Obtained forecasted frequency distribution will not any more be Gamma-distribution and will be calculated by numerical integration of the equation (7).

In future if the system passes into the third condition differing from the two previous it would be possible to repeat the procedure once again and to transform the equation (6) for three components.

It is clear that forecasted frequency distribution (7) – presented on Figure 9.1 – is not the full task of casual process because it represents only one-dimensional law of distribution and can be used only for the purpose of estimation of the values of parameters (quantiles) taking into account the possible non-stationary behaviour of the process.

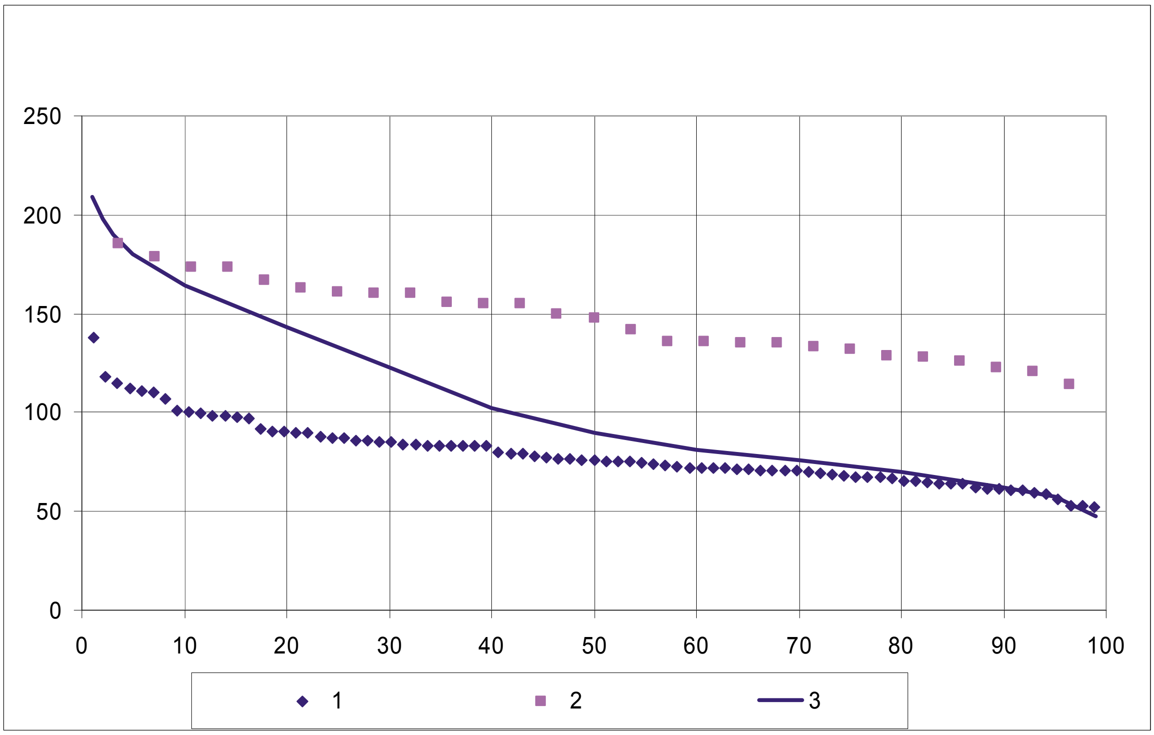



**Figure 9.1** Cumulative probabilistic curves of minimal runoff, river Don – town Liski
1 – minimal runoff before 1979, 2 – minimal runoff after 1980, 3 – Bayesian estimations of minimal water runoff

## Conclusion

As a result of application of Bayesian ideologies, forecasted cumulative probabilistic curve of minimal water runoff, based on setting of various weights to estimations of average value for conditionally homogeneous periods and calculation of forecasted frequency distribution using formula of total probability, was obtained. As a result of application of described procedure minimal water runoff value was estimated. Thus, regular application of the theorem of Bayes allowed to precise estimated characteristics of minimal river runoff in the Upper Don river basin and to evaluate reliability of water supply.

# 10 Options for drought analysis

Ladislav Kašpárek, Oldřich Novický

*The paper reviews methods that are used for temporal and spatial analysis of hydrological drought. Definition is firstly given of hydrological drought and subsequently the paper focuses on the statistical methods which are used for the drought analysis (mainly methods developed for assessments of drought duration and drought severity) and describes selected results of the application of these methods.*

## 10.1 Introduction

The experience used in this paper originates mainly from the studies carried out by the authors within the framework of their participation in the Low Flow group of the Northern European section of FRIEND (Flow Regimes from International Experimental and Network Data) project (a component of UNESCO's fifth International Hydrological Programme).

The Low Flow group activities included the ASTHyDA (Analysis, Synthesis and Transfer of Knowledge and Tools on Hydrological Drought Assessment through a European Network) project, which was focused mainly on the development of a textbook on Hydrological Drought (Tallaksen and Lanen, 2004), which describes processes and estimation methods for streamflow and groundwater droughts.

## 10.2 Methods

### *10.2.1 Definition of hydrological drought*

Different authors specified a variety of definitions of drought. Detailed review was made in Hisdal et al. (2000). In general, drought is deficiency of water to meet some need in given time and in given area. Water can be used by natural systems or it can meet the needs of mankind and its activities. The drought is therefore relative conception dependably on the need that is taken into account.

Hydrological drought is associated with a decrease in streamflows and subsequently with a decrease in groundwater levels and storages. The hydrological drought develops gradually because the streamflow is always fed from groundwater storage, which depletes slowly even if not augmented by percolation of water from precipitation. Hydrological drought can occur or continue also in winter periods if the precipitation is accumulated in the basin as a snow cover and low air temperature prevents snow melting.

Figure 10.1 shows time sequence of meteorological, agricultural and hydrological drought.

### *10.2.2 Characteristics of minimum flows*

Hydrological drought is always coupled with occurrence of low flows and the simplest methods for assessing extremity of drought events use basic statistical characteristics of flows.

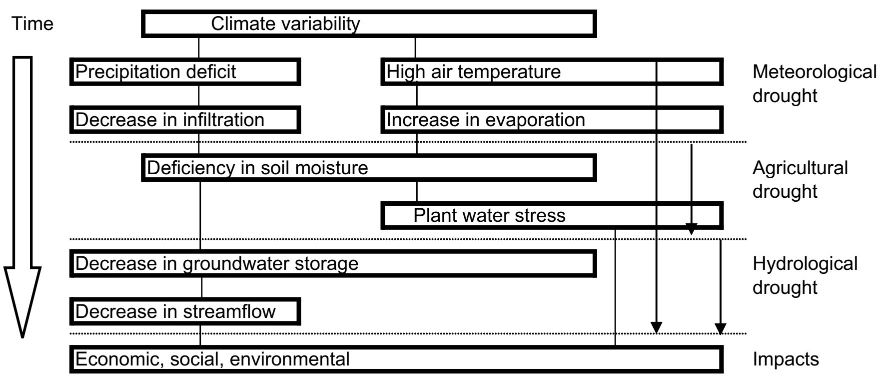



**Figure 10.1** Time sequence of meteorological, agricultural and hydrological drought, according to National Drought Mitigation Centre, University of Nebraska-Lincoln (simplified)

The most common method uses quantiles of daily flow duration curves, which are used in the Czech Republic in the form of m-day flows (flows which are exceeded for m days in a mean year). A lack of low flow assessment based on daily flows is that the low flows in daily series are mostly inaccurate and can easily be affected. In a number of European countries, the flow duration curves are derived from mean flows calculated for longer time interval such as 7, 10 or more days. Statistical characteristics of these series are less randomly affected and the information on hydrological drought from the longer time interval is more useful for assessments of drought impacts on natural ecosystems and water resources.

### *10.2.3 Analysis of deficit volumes*

Characteristics of low flows series can be insufficient for description of hydrological drought and therefore more complex approach based on analysis of deficit volumes and drought duration is increasingly used. This approach uses a threshold discharge (Figure 10.2), which can be interpreted as a flow need. This discharge is therefore predominantly selected at a level between $Q_{50\%}$ and $Q_{95\%}$.

Analysis of deficit volumes (volumes between the threshold and real discharge) and drought durations is focused on description and statistical analysis of time series of these characteristics within periods when the flow is below the threshold discharge.

For studies whose objectives include comparison of deficit volumes from different basins, it is suitable to use relative values, which express the volumes as percentages of long-term annual mean runoff. Drought durations do not have to be standardised.

Computer programs that have been developed for the analyses of deficit volumes and drought durations are mostly based on daily flow series. Algorithms of these programs include components for joining dependent deficit volumes, i.e. those which are interrupted by short period only when the flow exceeded the threshold, and for omitting negligible events, i.e. those whose duration or volume is small. Results of a study of these criteria are described in Hisdal and Tallaksen (2000). Effective methods for the assessment of daily flow series include the use of moving averages calculated e.g. for duration of 10 days.

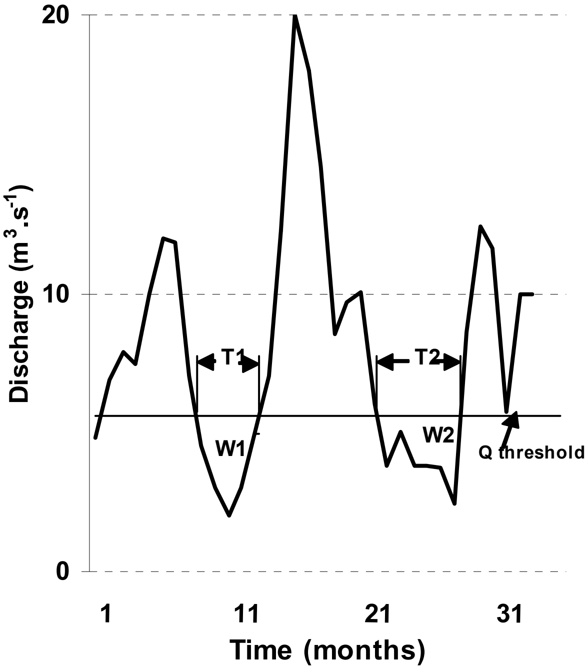



**Figure 10.2** Duration of a period (T) when flow is below threshold discharge and relevant deficit volume (W)

Methods and computer programs available in the Czech Republic for deficit volume analyses were compared in Kašpárek and Novický (2000). The methods were also reviewed with the aim to assess their suitability for routine analysis of network data.

In 1994, Czech Hydrometeorological Institute contributed to the activities carried out by Low Flow group of FRIEND project by developing EXDEV (Experiments with deficit volumes) computer program, which is a basic tool in use in the Czech Republic for calculation of deficit volumes (Řičica and Novický, 1995).

In 2000, Agricultural University in Wroclaw developed ROZKLAD NIZOWKI computer program (or Low Flow Estimation computer program), which was also available for members of the Low Flow group. These computer programs include similar components for analyses of deficit volumes by using their empirical distributions. Low Flow Estimation program provides wider options for the application of theoretical probability distribution functions. The output includes parameters of the probability distributions as well as tables and diagrams of the non-exceedance probability curves.

## 10.3 Results

The methods that were developed for assessment of hydrological drought have been applied in several regional studies in the Czech Republic.

These studies include an assessment of deficit volumes and drought durations derived from daily flow series (Kourková, 2002 and 2003), which was performed on 30 water gauging stations from the basin of the Upper Elbe River upstream from the Chrudimka tributary and for 15 stations from the whole territory of the Czech Republic as a component of draught analysis in the Czech Republic, Slovak Republic and Poland. These studies used ROZKLAD NIZOWKI (Low Flow Estimation) computer program.

Spatial variability of drought in the Elbe River basin was analyzed in Kašpárek and Novický (1996). For this study, the Elbe River basin was divided into 15 subbasins and 3 intermediate basins. Monthly flow series at respective water gauging stations were used for calculation of deficit volumes and duration of periods when the flows were below threshold discharges on a level of 50% of a long-term mean flow. In the period 1932–1990 used for the assessment, 13 significant droughts occurred in all of the basins (with one exception). Hydrological drought can therefore occur simultaneously on the whole basin but its individual parts are mostly affected miscellaneously. The largest deficit volumes (volumes standardised by long-term annual mean runoff) do not occur in mountain areas (Jizerské, Orlické or Šumava Mountains) neither in basins that have high storages of groundwater (basins of Loučná, Metuje and Orlice Rivers) but usually in basins of tributaries of the Middle and Lower Elbe and less frequently in the basins of the Sázava, Lužnice and Lower Vltava Rivers. An exception was period 1958–1963, when the most severe drought events occurred in Western Bohemia. Spatial distribution of characteristics that describe drought duration is similar to that of deficit volumes. The drought durations are less spatially variable for some of the drought events.

### 10.4 Discussion

It has been shown in several studies that precipitation deficit with duration of several months only may cause hydrological drought in catchments where water storage capacity is small compared to precipitation. Basins with high groundwater storage capacity are less vulnerable and hydrological drought in these basins occurs consequently to precipitation deficit with duration of several years.

An increase in air temperature in periods when it is above zero can contribute to evaporation and thus also to severity of drought. In catchments where winter air temperature is permanently below zero the most severe droughts occur consequently to dry summer and to subsequent prolongation of the dry period over to the cold winter to spring season.

Similar conclusions can be applied also in the Czech Republic whose variability in climate and hydrogeological conditions forms such combinations when flows in individual basins can drop to negligible values after only several dry months or when the flows in cretaceous basins are fed from high storages of groundwater which depletes slowly.

Hydrological droughts under Czech conditions can be divided into three groups:

- long-term drought (consequently to a long-term decrease in precipitation) lasting for several years being interrupted only by a short flow increase in spring,
- short and severe drought periods lasting from spring to autumn consequently to precipitation deficit in a period of several months,
- drought periods lasting from growing season to the whole summer and winter (mostly the most severe droughts) – see an example from 1953–1954 in Figure 10.3.

In general, hydrological drought in conditions of the Elbe River basin in the territory of the Czech Republic can occur when precipitation in a period of several months is smaller than potential evapotranspiration and water storage in the basin at the beginning of the period is below its mean value.

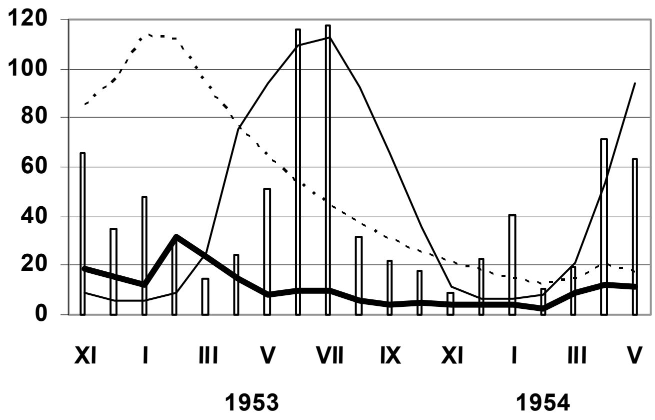



**Figure 10.3** Time distribution of main hydrological variables in the basin of the Elbe River at Děčín during hydrological drought in 1953–1954. Summer 1953 drought continued also in winter months in 1954, mean basin temperature was below zero from January to March (columns – monthly basin precipitation in mm, full light line – monthly potential evapotranspiration in mm, full heavy line – monthly runoff in mm, dashed line – simulated groundwater storage)

## 10.5 Conclusions

Development of methods and tools for assessment of hydrological drought in Europe has been substantially supported by the activities of Low Flow group of North European section of FRIEND (Flow Regimes from Experimental and Network Data) project. This project was a platform for verification and exchange of methods, computer programs, data and also knowledge from low flow studies. The Czech Republic has contributed to this cooperation by providing computer programs and results of the studies that are partially summarised in this paper. Profit of the Czech Republic includes free access to results of the projects carried out in other countries in the area of low flows and hydrological drought.

Some of the methods that have been developed in other countries and examined under conditions of the Czech Republic should be put into practice. These methods include mainly the use of flow duration curves and low flow characteristics derived from mean flows calculated for time interval that is longer than 1 day (e.g. 7 or 10 days). Tools from abroad that are well applicable in the Czech Republic include ROZKLAD NIZOWKI (or Low Flow Estimation) computer program, which is intended for analysis of deficit volumes and drought durations.

**Acknowledgements.** The research described in the chapter was carried out in the framework of the project No. VG20102014038: Proposal of a system for managing emergency situations associated with drought and water scarcity in the Czech Republic. Support is given by Ministry of Interior of the Czech Republic under the Security research program in 2010–2015 (BV II/2-VS).

References

Hisdal, H. and Tallaksen, L. (2000) Drought Event Definition, Technical report no. 6, ARIDE project. Oslo, Department of Geophysics, University of Oslo, 41 p.

Kašpárek, L. a Novický, O. (1996) Regionální studie nedostatkových objemů v dílčích částech povodí Labe v období 1932–1990 (Regional study of deficit volumes in the Elbe River basin in 1932–1990). Report of a project on Participation in FRIEND international project, Prague, VÚV T.G.M.

Kašpárek, L. a Novický, O. (2000) Analýza hydrologického sucha v povodí horní Metuje a Orlice (Analysis of hydrological drought in basins of the Upper Metuje and Orlice Rivers). In Workshop 2000 – Extrémní hydrologické jevy v povodích (Workshop 2000 – Extreme hydrological events in basins), ČVTS, Prague, 81–94.

Kourková, H. (2002) Studie nedostatkových objemů v povodí horního Labe (Study of deficit volumes in the basin of the Upper Elbe River). Report of a project on Participation in FRIEND international project. Prague, VÚV T.G.M.

Kourková, H. (2003) Regionální studie nedostatkových objemů (Regional study of deficit volumes). Report of a project on Participation in FRIEND international project. Prague, VÚV T.G.M.

Řičica, J. and Novický, O. (1995) Experiments with deficit volumes (EXDEV). Description of the program. Prague, ČHMÚ.

Tallaksen, LM. and van Lanen, HAJ. (Eds) (2004) Hydrological Drought – Processes and Estimation Methods for Streamflow and Groundwater. Amsterdam, Developments in Water Sciences 48, Elsevier B.V.

# 11 Application of a flow routing simulator in water management problem: Upper Narew case study

Renata J. Romanowicz, Adam Kiczko, Jarosław J. Napiórkowski

*The management of discharges from the water reservoir Siemianówka, situated in the Upper River Narew, NE Poland, aims to reach the desirable flow conditions downstream, in the Narew National Park, required to sustain ecologically valuable wetlands belonging to the Natura 2000 region. A one-dimensional flow routing model UNET was designed for the study. However, both optimisation methods and reservoir management analysis require numerous model realizations which are in practice too computationally extensive. To speed up the computations we apply a simplified Stochastic Transfer Function (STF) model of river flow, which is calibrated on historic data and multiple 1-D model realizations. The proposed flow routing model is stochastic, enabling derivation of prediction uncertainty in a straightforward manner. The obtained optimal control policy is tested on a fully distributed model.*

## 11.1 Introduction

The Narew National Park, located in the Upper Narew catchment, NE Poland, contains valuable fluvial ecosystems that are in danger due to the decreasing spring floods and human activity in the catchment. It is the aim of this study to design a water management policy that allows maintaining the required water levels downstream using the discharges from the Siemianówka reservoir situated upstream from the Narew National Park. This task requires a solution of a control problem that incorporates the distributed inundation maps in order to fulfil the water requirements of the wetland areas. The application of a distributed flow routing model within a control problem requires too much computing time to be practically possible. We propose here an application of a much simpler flow model consisting of a system of linear transfer functions, that is able to operate within a control set up, and is numerically efficient and fit for purpose.

Among many approaches to approximating complex models very popular are statistical simulators of model output, e.g., the simulator developed by Kennedy et al. (2006) to estimate the influence of input uncertainty on the output. Another approach was presented by Young et al. (1996) who applied a system of linear dynamic models for the simulation of a complex dynamic Global Circulation System. The purpose of that paper was determining the level of complexity of model dynamics that could be explained by the data, taking into account the uncertain nature of environmental observations. The authors showed that the data allow for the identification of only very few dynamic modes of the whole global circulation process.

Distributed flow routing models are used to predict the inundation extent and the risk of flooding (e.g. Romanowicz et al., 1996; Romanowicz et al., 2004). In order to facilitate the computations, Romanowicz and Beven (1998) applied a Stochastic Transfer Function (STF) approach to update on-line flood inundation forecasts on the River Culm, UK. Maps of flood inundation probability were obtained from the distributed flood routing model, developed at Lancaster. That paper showed that few hours ahead forecasts of the probability of flood inundation can be derived using water level forecasts obtained from the STF model at the gauging station downstream. In the other example (River Severn, UK), maps of flood inundation probability obtained using the distributed flood routing model ISIS were updated on-line using forecasts from the STF based flood routing system (Romanowicz et al., 2004a). Both examples showed that large amounts of computer time are required for the derivation of flood inundation maps using a distributed model, in comparison with much more efficient lumped modelling using STF based approach, which gives the forecasts together with their uncertainty estimates. In Romanowicz et al. (2004b; 2006) a system of connected rainfall – water-level and water-level routing STF models was built for on-line flood forecasting. That model was used to estimate water levels at sites where observations were available, but was not able to interpolate the inundation extent along the river between the measurement sites, which were usually tens of km away. Interpolation of water levels at cross-sections between measurement sites can be obtained when STF models are calibrated on the distributed model simulations at these cross-sections. Work on the application of a computationally efficient, STF based model to flood inundation predictions was also carried out at Lancaster University (Beven et al., 2008). In this work we present the design of a STF based simulator of a 1D-flow routing to optimise a water management system which combines ecological and economic goals. This work describes a preliminary study fully presented in Romanowicz et al. (2010).

## 11.2 Methodology

Many popular 1D flow routing models (ISIS, HEC-RAS, UNET, MIKE 11) include river geometry described by a sequence of cross-sections, related to the river reach under study. The methodology we apply here is based on the approximation of the 1-D model simulations (flows or water levels) for each model cross-section using a discrete time STF approach. The structure of a simulator depends on the distributed model used as a target. The UNET (Barkau, 1993) model was chosen in our study. This code is a numerical implementation of 1-D Saint Venant equation. As UNET applies spatial discretisation based on cross-sections along the river and the floodplains, the simulator can also apply that type of spatial discretisation, but it can be made coarser, than in the original model, depending on the goal of the application.

There are possible two different schemes of a simulator, depending on the variable used. When water levels are used as the STF model variable, the nonlinear transformation of water levels upstream is applied in order to separate linear dynamics from nonlinear stage routing process (Romanowicz et al., 2007; Romanowicz et al., 2008). The resulting simulator takes the form of a so-called Hammerstein type model (Figure 11.1A). However, when the flow is used as the model variable, the STF model is not accompanied by a nonlinear transformation of the input, as this type of model would not be able to keep the mass balance for the steady state solution. In that case the simulator consists of a system of interconnected linear STF models. The nonlinear flow-water level transformation used to evaluate water levels at each cross-section applies a rating curve-type conversion, which does not influence the flow routing dynamics (Figure 11.1B).

Figures 11.1A and 1B illustrate both schemes for a single sub-reach between two cross-sections of a river. In Figure 11.1A $h_{1,k}$ denotes water level upstream at discrete time period k, $h_{n,k}$ denotes the water level downstream at cross-section n, $f_{1,n}(\cdot)$ denotes a nonlinear transformation between input and output (Romanowicz et al., 2008). In Figure 11.1B $Q_{1,k}$ denotes discrete flow downstream, $Q_{n,k}$ .denotes flow value downstream and $gr_n(.)$ denotes the "rating curve" transformation for the 1-D (UNET) model. At the reach scale, the discrete-time STF can be presented as (Young, 1984):

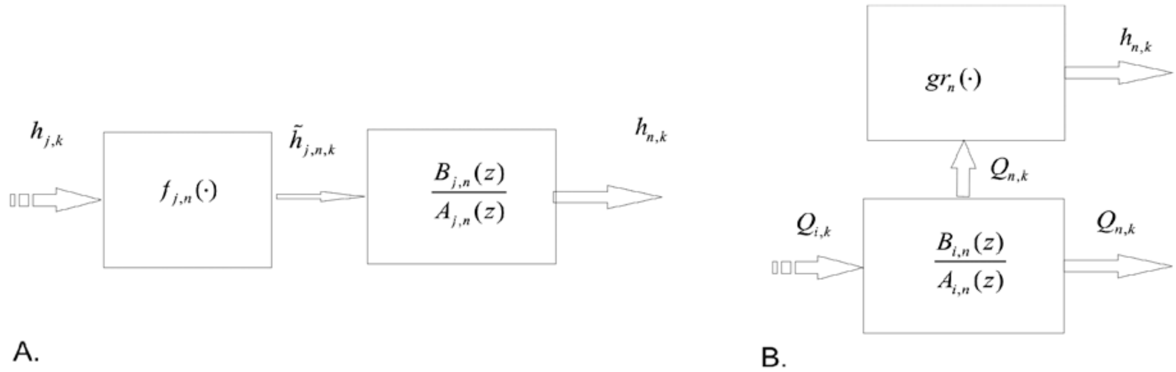



**Figure 11.1** A. The scheme of a nonlinear Hammerstein type STF simulator of one sub-reach of a 1D flood routing model; (a sub-reach without tributaries); B. The scheme of a linear STF model for flow with a nonlinear conversion for the water levels

$$\begin{aligned} x_k &= \frac{B(z^{-1})}{A(z^{-1})} u_{k-\delta} \\ y_k &= x_k + \xi_k \end{aligned} \tag{1}$$

where $u_{k-\delta}$ denotes STF model input (flow or water level), $x_k$ is the underlying 'true' flow or water level, $y_k$ is the noisy observation of this variable, $\delta$ is a pure, advective time delay of $\delta\Delta t$ time units, while $A(z^{-1})$ and $B(z^{-1})$ are polynomials of the following form:

$$A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n}; \quad \mathrm{B}(z^{-1}) = b_0 + b_1 z^{-1} + b_2 z^{-2} + \cdots + b_m z^{-m}$$

in which $z^{-r}$ is the backward shift operator, i.e. $z^{-r} y_k = y_{k-r}$, and $A(z^{-1})$ is assumed to have real roots that lie within the unit circle of the complex *z* plane. The additive noise term $\xi_k$ in (1) is usually both heteroscedastic (i.e. its variance changes over time) and autocorrelated in time. It is assumed to account for all the uncertainty at the output of the system that is associated with the inputs affecting the model, including measurement noise, unmeasured inputs, and uncertainties associated with the model structure. The orders of the polynomials *n* and *m* are identified from the data during the data-based identification process. In the following analysis, the triad [*n m δ*] is used to characterize this model structure. The model structure identification and estimation of parameters is performed using MATLAB optimisation routine together with a Simplified Refined Instrumental Variable (SRIV) routine from Captain toolbox (//lancaster.ac.uk).

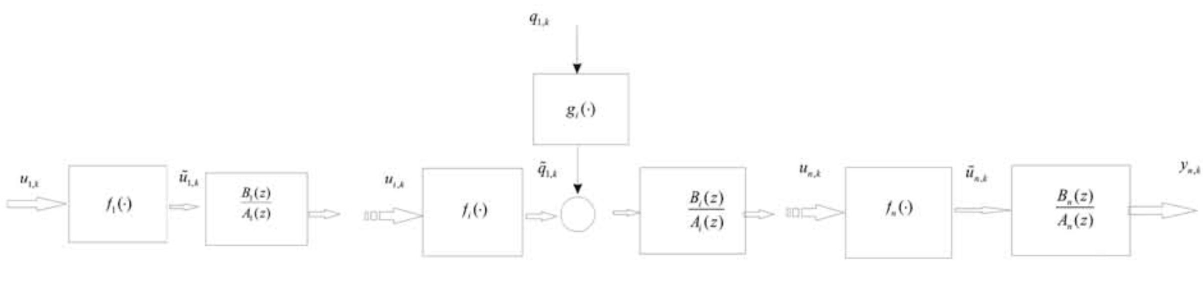


**Figure 11.2** UNET model simulator with a nonlinear sequential structure using Hammerstein type nonlinear transformation of water levels

Apart from the choice of routing variable, the choice of the STF model input determines the model structure. In the first approach, flow (water level) at the cross-sections upstream of each sub-reach of a numerical 1-D model is used as an input variable. In this case, the model of entire river reach consists of $n$ serially connected modules shown in Figure 11.1A or 1B depending on the choice of routing variables. The output from the sub-reach downstream is used here as an input to the upstream sub-reach. The other approach consists in building an independent Single Input Single Output (SISO) or Multiple Input Single Output (MISO) model for each cross-section, using the available observed inputs upstream (flows or water levels) and simulated by the 1-D model flow/water level at a cross-section as an output. In the second approach each cross-section of the 1-D model is modelled as an independent input-output system using the modules A or B from the Figure 11.1, depending on the choice of routing variable. In the case when lateral inflow is present, the multi-input model is required for the $i$-th cross-section. Therefore the whole set of 1-D model cross-sections is modelled by the set of those SISO (or MISO) transfer functions. In the following step of the procedure, the flows obtained from the 1-D simulator are transformed into water levels. That nonlinear transformation, derived using State Dependent Parameter (SDP) method (Young et al., 2001) from 1-D simulations for each cross section, is subsequently parameterised using radial basis functions (Buhmann, 2003). We shall call this approach a parallel model simulator to distinguish it from the sequential scheme. This approach should give smaller prediction errors due to the lack of propagation of the error. However when differences between the cross-sections are large, the modelling errors would occur due to smaller correlation between the sites. The diagram of a sequential structure is shown in Figure 11.2. The parallel scheme has the form shown in Figure 11.1A or Figure 11.1B.

## 11.3 Narew River case study: parallel STF simulator of UNET predictions

The case study reach (Kiczko et al., 2007) is about 90 km long. It starts at the Siemianówka reservoir and goes down over the lowland, agricultural area and Narew National Park (NPN) enclosing valuable wetland ecosystems (Figure 11.3). In recent years both a reduction in mean flows and shorter flooding periods have been observed which have resulted in a serious threat to the rich wetland ecosystems situated along the river in NPN. These undesirable changes were caused by changes in local climate manifested as recent mild winters and a reduction in annual rainfall that have resulted in reduced groundwater resources.


**Figure 11.3** Upper Narew scheme

The river reach is represented by 57 cross-sections at about 2 km intervals obtained from a terrain survey. The UNET model was calibrated by adjusting the Manning coefficients separately for the channel and left and right floodplains and a constant water surface slope was used as a downstream boundary condition. Observations of daily water levels at Bondary, Narew, Ploski and Suraż on the River Narew and Narewka and Orlanka on its two tributaries were used for the calibration and validation purpose. To keep a reasonable size of parameter dimensions, it was assumed that roughness coefficients do not change spatially between the gauges. The calibration period is 23-07-1981–28-08-1982 and the validation was performed in the period 27-08-1982–23-07-1983. Channel and floodplain (left and right) roughness coefficients for four reaches between gauging stations and downstream boundary conditions were used as model parameters. The optimization was carried out using the Differential Evolution (DE) algorithm (Storn and Price, 1997). Verification gave a good fit with a mean standard deviation less than 0.14 m for each gauging station (Figure 11.4). The uncertainty of the predictions was estimated using the Generalised Likelihood Uncertainty Estimation (GLUE) technique (Beven and Binley, 1992), as explained in Romanowicz et al. (2010).

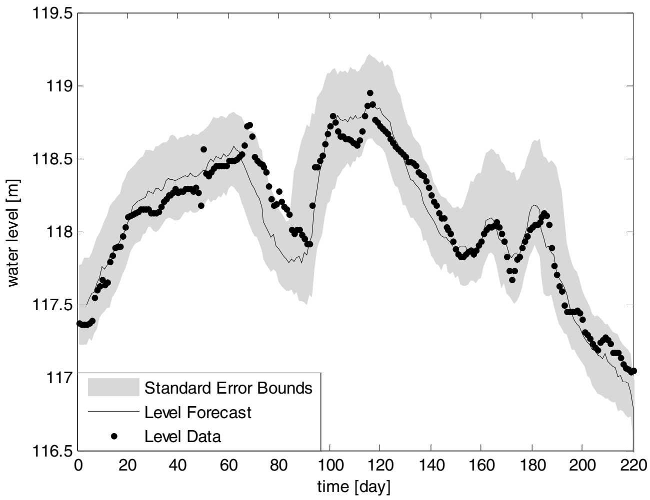



**Figure 11.4** 0.95 uncertainty bounds of UNET model predictions at Suraż obtained from MC simulations after re-sampling; shaded area denote 0.95 confidence bounds, continuous line denotes UNET predictions and dots denote observed water levels

Water level predictions are used in flood forecasting and inundation modelling. However, the analysis of UNET model results indicated that flow at each cross-section can be approximated by a linear dynamic relationship, while the level-level relationship is highly nonlinear. Simulated water levels at each cross-section can be derived from flow–water level nonlinear relationships, specific for each cross-section, for most of the cross-sections apart from one, situated near the first tributary. Therefore it was decided to use flows rather than water levels to build the flow simulator. The choice of flows as the STF model variable required the use of a parallel model structure as more suitable for the present application than the sequential scheme.

Figure 11.5 presents the obtained nonparametric relationship between water levels and flows at Suraż (upper panel) and SDP model fit (lower panel). The SDP nonparametric relation was parameterised using Radial Basis functions (Buhmann, 2003), but any other suitable parameterisation may be used. The resulting water levels at Suraż, obtained using the MISO STF [1 1 1 2 62 31 4] model with observed flows at Bondary, Narewka and

Orlanka as input variables and radial basis transformation, are shown in Figure 11.5 (lower panel) together with 0.95 confidence bounds.

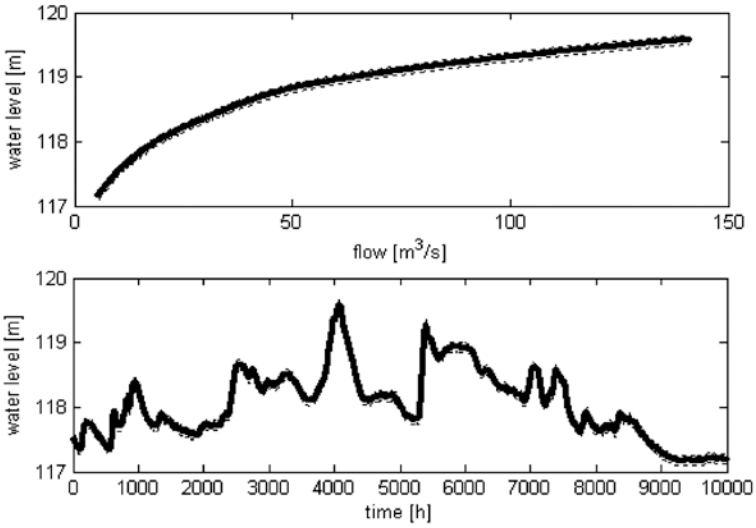



**Figure 11.5** Non-linear water level-flow relationship for Suraż (upper panel) with 0.95 confidence bounds and the SDP model fit (lower panel)

50 STF models were obtained for UNET simulated hourly flows at each cross-section along the River Narew between Bondary and Suraż, with observed flows at Bondary, Narewka and Orlanka as input variables. All the models have 1st order dynamics. The obtained goodness of fit criterion $R_T^2$ for the validation stage (1982–1983) varies between 99.33% and 99.99%.

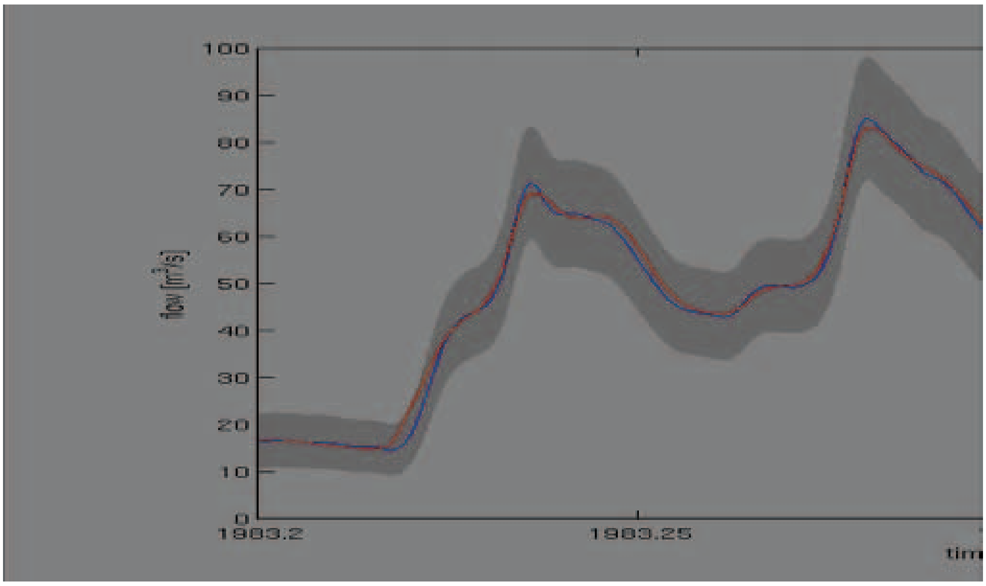



**Figure 11.6** Validation of the STF flow simulator for the cross-section at Suraż

The validation results (Figure 11.6) indicate that the UNET model has nearly linear input-output relationships for flows, i.e. it has linear dynamics.

## 11.4 Application of distributed model simulator in reservoir management

In this application we use discharges from the Siemianówka reservoir as control variables. The optimisation criteria described in Kiczko et al. (2008) are chosen to combine reservoir management and ecological requirements posed by the wetland ecosystems along the river floodplains. All considered criteria were merged into a single objective function described in Kiczko et al. (2008). Optimization of discharges from Siemianówka reservoir (control stage) was performed using historical discharges at the Bondary gauging station from the time when the reservoir was not yet built. In this way the ability of improving water conditions by introducing the discharges from the reservoir is tested. The reservoir is described using a simple discrete balance equation. Initial reservoir storage was set to the recommended value for a chosen control period by the reservoir management policy. Reservoir outflows are represented as a sum of rectangular pulses

$$Q_{out}(t) = Q_{base} + \sum_{j=1}^{NP} P_j(t, t_j, dt_j, q_j) \tag{2}$$

where: $Q_{base}$ – minimum flow (a minimum allowed discharge from the reservoir), $t_j$ – time middle point of $j$-th pulse, $dt_j$ – pulse duration time, $q_j$ – discharge and $NP$ – number of considered pulses.

The $j$-th rectangular is defined as a

$$P_j(t, t_j, dt_j, q_j) = \begin{cases} 0 \text{ for } t \in \left\langle t_0, t_j - 0.5dt_j \right) \\ q_j \text{ for } t \in \left\langle t_j - 0.5dt_j, t_j + 0.5dt_j \right\rangle \\ 0 \text{ for } t \in \left( t_j - 0.5d_j, t_k) \right\rangle \end{cases} \tag{3}$$

where $t_0$ – initial time and $t_k$ – optimization horizon.

Values of the middle time of the pulse, $t_j$, pulse duration time $dt_j$ and the pulse height $q_j$ were used as control variables describing reservoir discharges. Ten pulses were applied ($NP$ = 10), therefore there were 30 control variables.

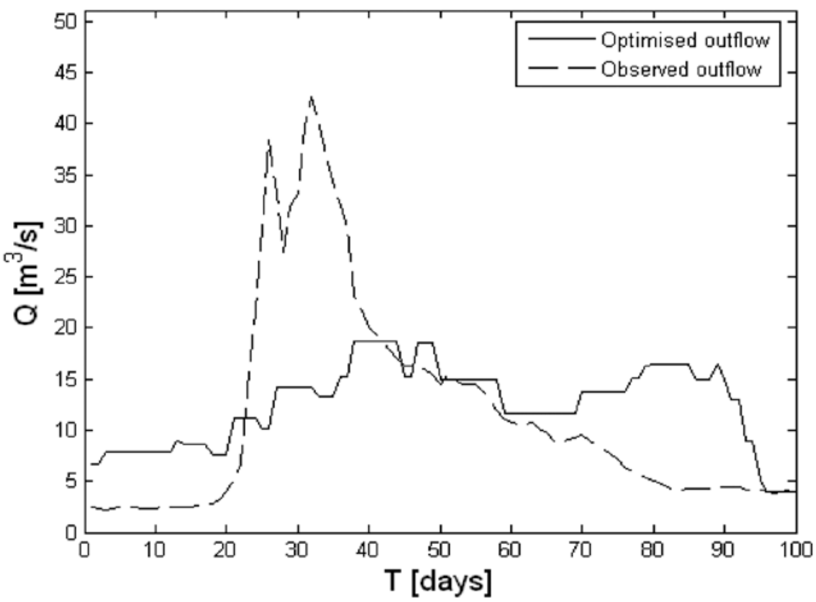



**Figure 11.7** Reservoir discharges: optimised marked by a continuous line and observed outflow marked by a dashed line

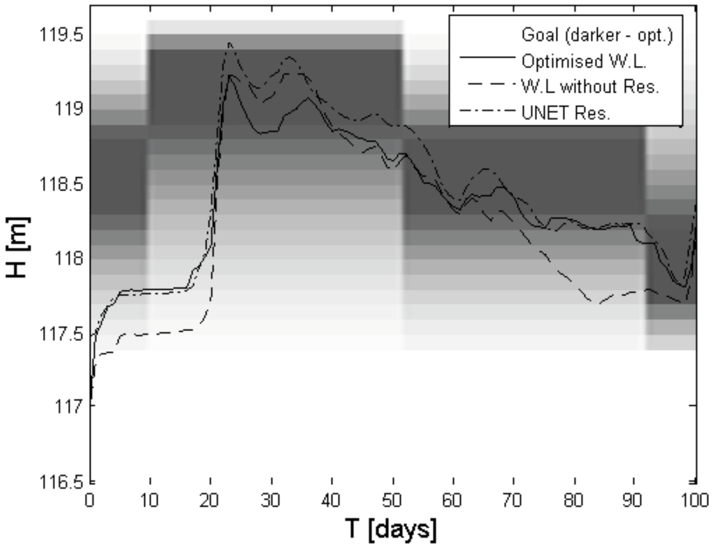



**Figure 11.8** Water levels at Suraż; historical observed marked with a dashed line, optimised marked with a continuous line; UNET simulations marked with a dash-dotted line

In the original work of Kiczko et al. (2008) three different management strategies were analyzed. In the present paper we report the results for only one scenario, from the period 0.5.10.1982–13.07.1983, in order to test the performance of the simulator. The optimisation problem was solved using DE Algorithm with flow simulator used for flow routing. Figure 11.8 presents a comparison of water levels at Suraż simulated using UNET with results obtained using the STF flow simulator when the optimised discharges from the reservoir are applied. Also shown are the observed historical water levels at Suraż. The dark shaded area presents the values of the optimum criterion applied in order to fulfil both ecological and economical goals (the darkest shade represents the best value of the criterion).

Figure 11.9 presents the inundation probability maps at high water levels and Figure 11.10 presents the inundation probability maps at low water levels, both obtained for the Siemianówka discharges optimised using the STF flow simulator within the control problem. These maps were obtained from the UNET predictions using the GLUE technique (Romanowicz et al., 2010).

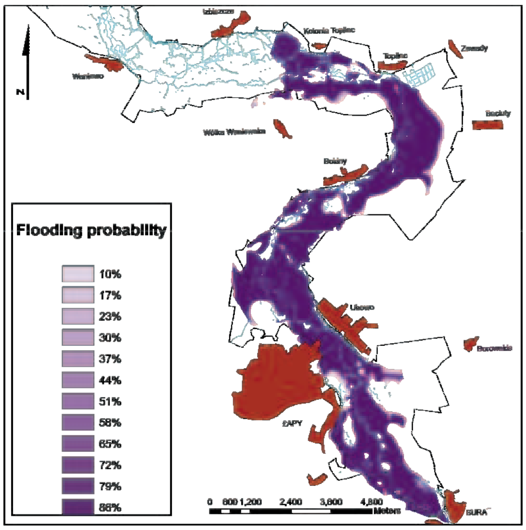



**Figure 11.9** Inundation probability maps: high water levels

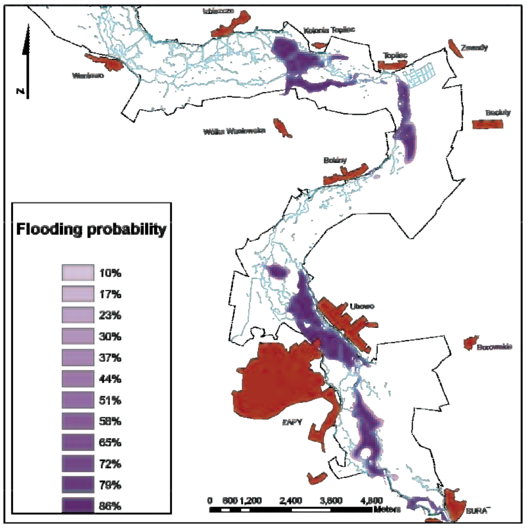



**Figure 11.10** Inundation probability maps: low water levels

## 11.5 Conclusions

We have demonstrated that the 1D flow routing UNET model can be successfully used for spatial and temporal interpolation of the observations of flow along the river, required to develop a system of lumped Stochastic Transfer Function models (so-called STF flow simulator). The simulator structure depends on the choice of the routing variable (water levels or flows) and on the choice of the input variables (sequential or parallel system). In the case of the UNET model, a choice of flows as routing variable and a parallel structure gives superior results. The flow simulator was applied within the optimisation routine to derive the best reservoir discharge scheme from the point of view of joint ecological and economic criteria. The maps of inundation probability for high and low flow conditions were derived for the optimal reservoir discharges using UNET and GLUE uncertainty estimation technique.

### References

Barkau, RL. (1993) UNET, one-dimensional flow through a full network of open channels, user's manual version 2.1. Publication CPD-66, U.S. Army Corps of Engineers, Davis, CA, Hydrologic Engineering Center.

Beven, K., Young, P., Leedal, D., and Romanowicz, R. (2009) Computationally efficient flood water level prediction (with uncertainty). In Flood Risk Management and Practice, Samuels et al. (eds), London : Taylor and Francis Group. ISBN 978-0-415-48507-4.

Buhmann, MD. (2003) Radial Basis Functions: Theory and Implementations. Cambridge : Cambridge University.

Kennedy, MC., Anderson, CW., Conti, S., and O'Hagan, A. (2006) Case studies in Gaussian process modelling of computer codes. *Reliability Engineering and System Safety*, 91, 1301–1309.

Kiczko, A., Romanowicz, RJ., and Napiórkowski, JJ. (2007) A study of flow conditions aimed at preserving valuable wetland areas in the Upper Narew Valley using GSA-GLUE methodology. Proceedings 21st International Conference on Informatics for Environmental Protection, September 12–14, 2007, Warsaw, Poland, Shaker Verlag, 175–183.

Kiczko, A., Piotrowski, A., and Napiorkowski, J., and Romanowicz, R. (2008) Combined reservoir management and flow routing modelling: Upper Narew case study, to appear in the Proceedings of River Flow Conference, Istanbul, September 2008.

Romanowicz, RJ., Young, PC., and Beven, KJ. (2004) Data assimilation in the identification of flood inundation models: derivation of on-line multi-step ahead predictions. BHS International Conference: Hydrology, Science and Practice for the 21st century. Webb, B., Arnell, N., Onfet, C. et al., 348–353.

Romanowicz, R., Beven, KJ., and Young, PC. (2004) Assessing the risk of flooding in real time. Proc. ACTIF Conference on Quantification, reduction and dissemination of uncertainty in flood forecasting, Delft, Netherlands. http://www.actif-ec.net/Workshop2/papers/ACTIF_S1_06.pdf; reviewed on-line publication.

Romanowicz, RJ., Young, PC., and Beven, KJ. (2006) Data assimilation and adaptive forecasting of water levels in the river Severn catchment, United Kingdom. *Water Resources Research,* 42: W06407. Doi:10.1029/2005WR004373.

Romanowicz, RJ., Kiczko, A., and Napiorkowski, JJ. (2010) Stochastic transfer function model applied to combined reservoir management and flow routing. Accepted for publication in *Journal of Hydrological Sciences.*

Storn, R. and Price, KV. (1997) Differential Evolution – a simple and efficient adaptive scheme for global optimization over continuous spaces. *Journal of Global Optimization*, 11, 341–359.

Young, PC. (1984) Recursive Estimation and Time Series Analysis. Berlin : Springer Verlag.

Young, PC., Parkinson, SD., and Lees, M. (1996) Simplicity out of complexity in environmental systems: Occam's Razor revisited. *J. Appl. Stat*., 23, 2 and 3, 165–210.

Young, PC., McKenna, P., and Bruun, J. (2001) Identification of non-linear stochastic systems by state dependent parameter estimation. *International Journal of Control*, 74, 1837–1857.

# 12 Fate of pollution from sediments re-suspended during catastrophic flood on quality of environment of the flood plain

Miroslav Rudiš, Petr Valenta, Jana Valentová, Ondřej Nol

*This contribution presents methods to predict the effect of a catastrophic flood on old sediments loaded with high pollution. Such sediments may re-suspend due to a flood of, say, a $Q_{500}$, they are transported downstream and settle in the flood plain. After the water decline, pollution leaches to subsoil and leaks through it to ground water. The methods presented here use the two-dimensional dynamic model FAST-2D upgraded with erosion and sedimentation terms and Visual MODFLOW model for groundwater flow and travel time modelling. The final result brings the estimate of the change of concentration of a representative heavy metal in the ground water with time as a consequence of new pollution that came with the new sediments to the flood plain.*

## 12.1 Introduction

Paper just presented deals with simulating the effect of catastrophic flood able to re-suspend old sediments loaded with high pollution on the environment of the flood plain especially on ground water layers. Research on this topic was realized with the support of European Commission in the framework of the International Project AquaTerra. This Project deals with improvement of the quality of water, suspended load and sediment transport of the Czech reach of the Elbe as a part of the total stream of the Elbe in Germany as divided to fluvial and tidal parts.

In the reach of the Czech Elbe there are three sources of highly polluted sediments that cannot be re-suspended even by $Q_{100}$ but, with higher discharges, the possibility of re-suspension and following transport and sedimentation in the flood plain increases. One of these localities is the reservoir Les Království separating the mountain and sub-mountain reach of the stream. The dam of it closed the valley in 1914. The second half of 20$^{th}$ century, its watershed was loaded by heavy machinery, paper mills and incompetent agricultural policy that ploughed the sub-mountain green fields and used high concentrated fertilizers there. Thus the enormous wash-out of soil polluted by fertilizers appeared and all this amount was settled in the reservoir. Together with paper mill wastes and heavy metals from factories, the sediment load was identified as very dangerous. In spite of this, decrease of water quality due to usual floods was not observed under the reservoir. The question was: what would be the consequences of catastrophic flood that would transport these sediments from the reservoir and transport them downstream e.g. to the flood plain over the city of Jaroměř where wells for water work supply are installed.

To this aim, total amount of sediments was determined, its chemical and geo-mechanical properties were analyzed, the discharge of $Q_{500}$ was estimated by stochastic methods and its effect on re-suspension was computed. Using hydraulic methods, the water head in the reservoir was determined and the main characteristics i.e. the mean grain diameter and settling velocity of the suspended load leaving the reservoir were estimated. With this, the transport downstream and sedimentation of it in the flood plain was modelled using the upgraded FAST 2D simulation device. As a result, we obtained the map of the new sediment layers (in kg per m$^2$) over the part of flood plain of the interest. This map was the initial basis for further estimation of the transport of pollution from new sediments to ground water horizon. The zinc (Zn) was used as a characteristic metal due to its relatively easy analyse and high concentration in overall sediments.

## 12.2 Methods

### *12.2.1 Modelling of sedimentation*

The flow simulation model FAST 2D was developed in cooperation with the University of Karlsruhe, Germany – Wenka, Valenta (1991). The first attempt of solution of re-suspension was published in Rudiš, Valenta, Valentová (2002). The practical computation of sedimentation of the load re-suspended in the reservoir of "Les Království" was published in the partial report worked out in AquaTerra Project – Rudiš and co-workers (2006).

The term expressing sedimentation of suspended load in the flood plane during the flood was defined in FAST 2D model in relatively simple form. Such expression is not applicable in alluvial flows but it may be successfully used in rivers of the type of the Czech Elbe. The limiting conditions may be seen from following assumptions:

- Concentration distribution of re-suspended sediment along the height of the flow is uniform. This assumption was verified by many all-profile measurements of velocities and concentrations during previous research on the Czech Elbe – Rudiš, Trejtnar et al. (1998). The bottom of the main channel of the Czech Elbe is formed by coarse gravel causing high turbulence and this phenomenon prevents sedimentation there. Suspended load especially under reservoirs is formed by very fine particles only, with the total lack of sand. That is the reason of the uniform concentration distribution.
- During the flood event, maximum concentrations appear in the rising limb of the hydrograph. During receding limb, on the contrary, concentration decreases intensely to minimum even at high discharges. This was verified especially in the catastrophic flood in Moravia 1997. In the case of Les Království, the supply of sediments stops abruptly in the moment of ceasing the function of spillways with the decreasing head. The bottom outlets contribute no more sediment downstream because the velocity in their erosion cones decreases with falling head and not exceeds the maximum admissible velocities for non-tenacious sediments.
- The extended FAST 2D model solves the problem of sedimentation in flood plane as a quasisteady process at constant discharges and concentrations in time intervals given by the hydrograph.
- We assume that no sedimentation appears in the main channel while, in the flood plain, sedimentation proceeds within some limits of flow velocity.
- We assume more that, in the flood plain, the transported sediment load is supplied accordingly to discharges and concentrations pertinent to the rising limb of hydrograph while, in the time of receding limb, concentration tends to zero. After the return of the flow to the main channel, the total sedimentation proceeds in remaining flood pools.

Implementation of the term representing the sedimentation process into the FAST 2D simulation tool is based on an algebraic relationship between the bed shear stress and the sediment deposition rate as proposed by Krone and Metha (1962) and Metha and Partheniades (1975). The authors investigated the deposition process of fine sediment, and proposed formulas to determine the deposition rate. Mehta and Partheniades' (1975) formula is (see also Wang and Wu, 2004) a function of shear stress limits, settling velocity and concentration. Moreover, the term depends also on a coefficient between 0 and 1 related to shear stress too. The proposed simplified methodology used here is based on an assumption that the unsteady water flow during the flood event can be approximated as a sequence of quasi-steady states and that the suspended sediment concentration specified at the model inflow boundary for every quasi-steady state prevails in the whole computational domain and thus can be assumed as a constant value at every time step of the simulation.

Together with the digitalized map of the flood plane under consideration, the numerical simulation is realized step by step on the control volume and uses the same computational grid as the FAST 2D water flow simulation tool. The result i.e. the map of new sediment layers in the flood plain over the city of Jaroměř may be seen in Figure 12.1. There, we can see not only the flood plain sediments but also the flooded pond where sedimentation is significant.

### *12.2.2 Modelling of groundwater flow with pollutants*

Leaching of heavy metals from the flood-pool sediments and from bottom sediments of Jaroměř pond can be a major risk for groundwater contamination at this site. In the framework of the study, we aimed to assess quantitatively the transport of heavy metals (solute flux) to the groundwater body and carry out a risk analysis for groundwater in the Quaternary aquifer of the Elbe River and the Úpa River via numerical modelling.

Results of chemical analyses of samples taken from the Les Království reservoir showed that zinc is one of the most important pollutants; therefore this study is focused on the zinc fate and behavior.

Zinc (Zn) was chosen as a representative among other heavy metals due to following features:

- Zn appears in sediments and, as a matter of fact, in the entire environment at relatively high concentrations enabling to perform the leaching experiments at high reliability;
- Zn is relatively uniformly spread in the top soil and sediments;
- Uniformity of concentrations originates from the high natural (geogenic) portion of Zn in east Bohemia geology at all and also from the rain wash of constructions protected by it against corrosion (as gutters, bottoms of cars and so on);
- A single drawback of zinc as a representative of heavy metals appears in high dependence of its concentration on pH (Zn is more intensively loosen with increasing pH).

Zinc can leach either from bottom sediments into the saturated zone of the Jaroměř pond or from the flood-pool sediments into the unsaturated zone and the saturated zone east of Heřmanice nad Labem, where the mathematical model of sedimentation assessed the highest volume of settled sediment in the flood plane (see Figure 12.1) Transport in the saturated zone is again dominated by advective and dispersive transport and sorption. Sorption of heavy metals can be described in each individual model cell by pH-dependent, or linear isotherms taking into account the major inorganic aqueous complexes. The three main sorbents are clay minerals, iron oxides and organic matter. The saturated sandy gravels in this site are mainly aerobic and acid. Therefore, precipitation of zinc as carbonates or oxides plays no role. Because oxide contents are probably very low, sorption on organic matter and clay minerals are key processes controlling the reactive transport. The sorption intensity is mainly dependent on the soil pH. For the purposes of this theoretical study, we neglected sorption in the Quaternary aquifer to outline the "worst" possibility of the influence on groundwater.

Based on data collected from existing monitoring wells (geometry of Quaternary aquifer, hydraulic heads, hydraulic conductivity), a 3D regional flow and transport model was

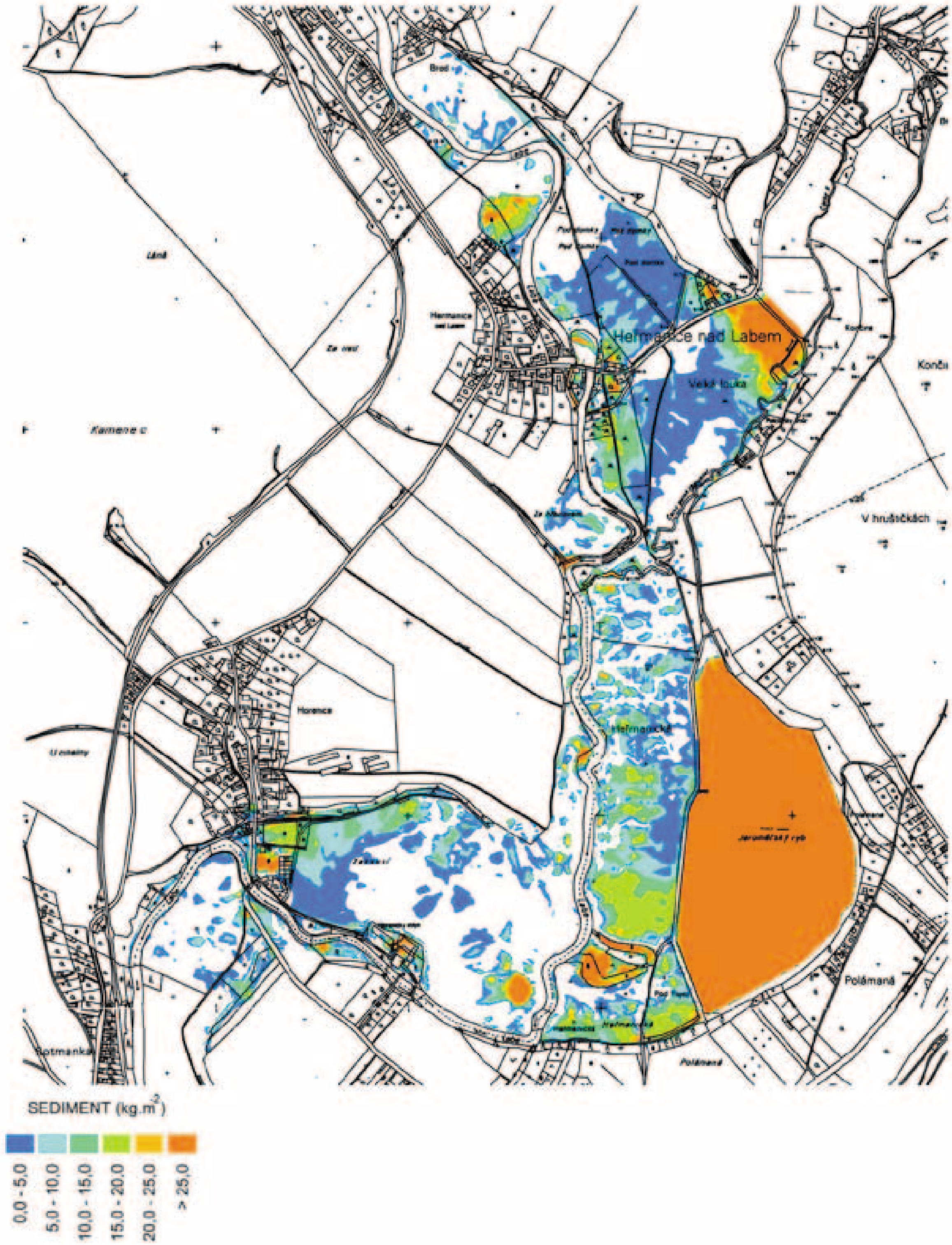



**Figure 12.1** Sedimentation over the unit of area in the flood plane of Jaroměř during $Q_{500}$

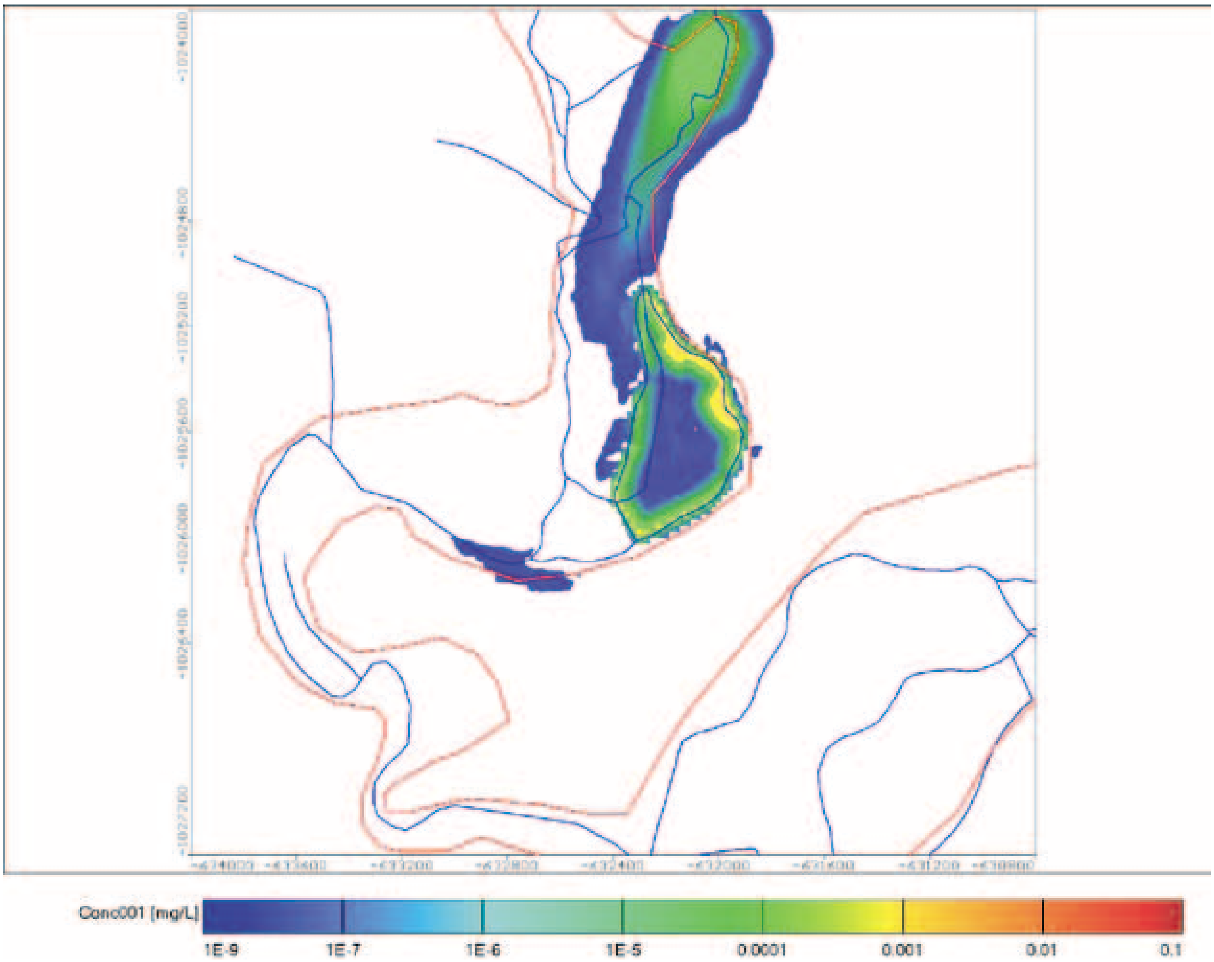



**Figure 12.2** Simulated zinc concentration distribution after 10 years – third model layer

developed that captures the main hydrological and hydro-geological characteristics of the field site. The flow model takes into account the lithological variability as well as surface water (brooks, ditches and drains).

In order to study the influence of the groundwater body, Visual Modflow written by Guiguer and Franz (2003) was used. Flow solver MODFLOW – McDonald and Harbaugh (1996) as a part of the Visual Modflow package, was used for groundwater flow and groundwater travel time modelling as was MT3DMS (Zheng and Wang, 1999) that has the capability of modelling changes in concentrations of groundwater contaminants due to advection, dispersion, diffusion, and some chemical reactions including equilibrium-controlled linear or non-linear sorption, and first-order irreversible or reversible kinetic reactions.

The vadose zone typically has a lower hydraulic conductivity because some of the pore space is filled with air, and the soil moisture in the vadose zone only travels through the wetted cross-section of the pore space. The relative proportion of air to water in the pores can vary, and consequently the hydraulic properties of the porous media can also vary.

Thus, for assessment of heavy metal flux across vadose zone vertical bench-scale numerical model transport in unsaturated soil profile was constructed. Software HYDRUS-1D (Šimůnek et al., 2005) was employed.

The HYDRUS program is a finite element model for simulating the one-dimensional movement of water, heat, and multiple solutes in variably saturated media. The program numerically solves the Richards' equation for saturated-unsaturated water flow and Fickian-based advection dispersion equations for heat and solute transport. The flow equation incorporates a sink term to account for water uptake by plant roots. The heat transport equation considers conduction as well as convection with flowing water. The solute transport equations consider advective-dispersive transport in the liquid phase, and diffusion in the gaseous phase.

## 12.3 Results

Prediction of sedimentation after the $Q_{500}$ flood event downstream the reservoir Les Království may be seen in Figure 12.1 as a map of sediment layers in $kg.m^{-2}$. One can see that the maximum load is as a flood plain sediment in the flooded valley of a small Elbe tributary and as a flood pool sediment in the flooded pond of Jaroměř. Simulating the zinc concentration after some time delay, we may see, e.g. in Figure 12.2, its distribution of concentration after 10 years in the third layer of underground.

## 12.4 Conclusions

Combining re-suspension transport and sedimentation of polluted suspended load simulated by the upgraded FAST 2D model and modelling of leaching of pollutants and their flow to underground water by visual MODFLOW model bring the useful tool for prediction of the fate of pollution represented by zinc concentration from new sediments to ground water.

**Acknowledgements.** This work was supported by the European Integrated Project AquaTerra (GOCE 505428). The project has received research funding from the European Community's Sixth Framework Programme. The second and third author were supported by the ''Revitalization of Landscape and Urban Water System Affected by Significant Anthropogenic Changes" project (MSM6840770002), funded by the Ministry of Education of the Czech Republic.

References

Guiguer, N. and Franz, T. (2003) Visual ModFlow user`s manual. Waterloo Hydrogeologic Inc., Ontario, Canada.

Krone, RB. and Mehta, AJ. (1962) Flume studies on the transport of sediment in estuarine shoaling processes. Hydraulic Engineering Laboratory, University of Berkeley, California, USA.

Harbaugh, AW. and McDonald, MG. (1996) User's documentation for MODFLOW- 96, an update to the U.S. Geological Survey modular finite-difference ground-water flow model. USGS Open-File Report 96-485.

Mehta, AJ. and Partheniades, E. (1975) An investigation of the depositional properties of flocculated fine sediment. *J. Hydr. Res.,* 13, 4, 361–381.

Rudiš, M. et al. (2006) Completion of 2D modelling of sedimentation of polluted sediments in the floodplain downstream Les Království and preparation of the modelling of percolation of pollution to groundwater for the area of Lovosice. Projekt No 505428 (GOCE) – AquaTerra, Deliverable No R4.14_BIS, T. G. Masaryk Water Research Institute, Prague, August 2006.

Rudiš, M., Trejtnar, K., et al. (1998) Transport und Belastung von Schwebstoffen und Sedimenten in der tschechischen Elbe. Projekt Nr. 523KfK9404, Abschlussbericht für 1995–97, Praha – Hradec Králové.

Rudiš, M., Valenta, P. und Valentová, J. (2002) Numerische Modelierung der Abschwemmung von Sedimenten hinter den Deichen am Stauanfang des Wehres Střekov (Schreckenstein). 10th Magdeburger Seminar, Poster No 39, Špindlerův Mlýn, CR.

Šimůnek, J., van Genuchten, MTh., and Šejna, M. (2005) The HYDRUS-1D Software Package for Simulating the One-Dimensional Movement of Water, Heat, and Multiple Solutes in Variably-Saturated Media; Version 3.0. Department of Environmental Sciences, University of California.

Wang, SY. and Wu, W. (2004) River sedimentation and morphology modelling – the state of the art and future development. Proc. Of the Ninth International Symposium on River Sedimentation, October 18–21, Yichang, China.

Wenka, T. und Valenta, P. (1991) Entwicklung und Austesten einer tiefgemittelten Version des FAST 2D Computer Programms. Universität Karlsruhe, 1991.

Zheng, C. and Wang, PP. MT3DMS (1999) A modular three-dimensional multispecies model for simulation of advection, dispersion and chemical reactions of contaminants in groundwater systems; Documentation and Users Guide, Contract Report SERDP-99-1, U.S. Army Engineer Research and Development Center, Vicksburg Miss.

# 13 Linking ensemble prediction systems of weather forecasts to hydrology within the PREVIEW project

Florian Pappenberger

*Early warning on floods is essential to initiate timely measures to prevent and reduce loss of life and economic damage. Numerical Weather Predictions and in particular Ensemble Prediction Systems form a vital part of the warning chain. This paper shows that precipitation and temperature forecasts issued by the European Centre for Medium-Range Weather Forecasts (ECMWF) are skilful up to day 5–7. Forecasts accumulated over a longer time are largely more skilful than forecasts accumulated over short time periods. Additionally, the value of using forecasts from seven global ensemble weather prediction systems is illustrated on a case study. This first assessment indicates the potential of the grand-ensembles approach to raise preparedness and thus to reduce the socio-economic impact of floods.*

## 13.1 Introduction

One major research challenge of the 21st century is to mitigate the effects of natural hazards. Of all natural disaster, flooding is the most frequent, affecting the second largest number of people after droughts causing damage in excess of several billion Euros a year (EM-DAT, 2007).

Flood forecasting based on observed precipitation or river levels, limits the lead time of the forecast to the natural response time of the catchment. However, longer lead times provide civil protection authorities with more time to prepare for the event and give an advanced warning to the public, and could reduce the socio-economic impact of the flooding. Although the incorporation of numerical weather forecasts into a flood warning system can significantly increase forecast lead time (for example Pappenberger, F., et al., 2005; Gourley, JJ. and Vieux, BE., 2005; Krzysztofowicz, R., 2002; Ahrens, B. and Jaun, S., 2007; Verbunt, M., et al., 2006; Gouweleeuw, B., et al., 2005) especially when ensemble prediction systems are used. Ensemble prediction systems (EPS) incorporate uncertainties in the initial conditions and factors of the modelling process in the numerical weather predictions and produce multiple weather forecasts (Palmer, TN., et al., 2005). It has been demonstrated that EPS have more value than a single deterministic forecast with the same resolution from the same modelling system (Zhu, YJ., et al., 2002; Roulin, E., 2007).

EPS forecasts from a single forecast centre only address some of the uncertainties inherent in numerical weather predictions and many other sources such as boundary conditions or numerical implementations exist. For example, model physics and numerics have substantial impact in generating the full spectrum of possible solutions (Fritsch, JM., et al., 2000). A multi-model approach is an effective and pragmatic approach of incorporating some of these additional sources of uncertainty (Doblas-Reyes, FJ., Hagedorn, R., and Palmer, TN., 2005).

In the first part of this paper we evaluate the suitability of ECMWF EPS forecasts for hydrological application, which has been investigated within PREVIEW projects (for more details see Cloke, HL. and Pappenberger, F., 2007; Pappenberger, F. and Buizza, R., 2007). In the second part we show the results of a hydrological multi model ensemble flood forecast for a case study of the floods in Romania in October 2007 (for more details see Pappenberger, F., et al., 2008).

## 13.2 The skill of ECMWF predictions for hydrological modelling (for more details see Cloke, HL. and Pappenberger, F., 2007; Pappenberger, F. and Buizza, R., 2007)

Pappenberger et al. (Pappenberger, F., et al., 2007) have proposed recommendations to verify forecasts using methods more relevant for hydrological applications, so that the information can be easily interpreted by hydrologists who use meteorological data to drive hydrological models. Pappenberger et al (2007) suggestions are as follows:

1. Any verification should always include variables of interest to hydrology which are specific to the catchment characteristics and antecedent conditions.
2. Comparison should be made against actual observations.
3. The benchmark model for computing skill should be persistency or a probability distribution based on the dominant hydrological processes relevant for a particular hydrological application.
4. The choice of thresholds of performance measures has to reflect the threshold behaviour of the catchments.
5. The verification area should be catchment based.
6. Spatial averaging and interpolatioin the evaluation process needs to match the use of the data in the hydrological model.
7. Smoothing performance measures or physical quantities with large time averaging windows (for example 90 days) is useful for seeing long term trends, but unacceptable when trying to understand performance on a hydrologically relevant time scale of individual events.
8. Forecasts are used continuously, meaning that a forecast for day four has to be evaluated on, for example, accumulated precipitation until day four.

This assessment follows these recommendations while evaluating the performance of precipitation and temperature forecasts.

### *13.2.1 Average of model performance of ECMWF over lead times*

*Temperature*

In Figure 13.1, the Nash-Sutcliffe score for the Danube catchment is plotted against the lead time in the form of box and whisker plots. Each box has lines at the lower quartile, median and upper quartile. The whiskers are lines extending from each end of the box to show the extent of the rest of the data. A majority of forecasts are skilful up to day 5, with a large proportion of forecasts still achieving a Nash-Sutcliffe value above 0.5 at 6 days. The figure shows large uncertainty bounds.

*Precipitation (accumulated over leads time)*

Figure 13.2 shows the Nash-Sutcliffe for precipitation accumulated over the entire lead time and the skilful predictions stretch up day four/five. The forecasts accumulated over a longer time are largely more skilful than forecasts accumulated over short time periods. The usage of accumulated predictions is important as, for example, large error at the beginning of a forecast may trigger the hydrological process of saturation excess overland flow at a lead time of 120 hrs, whereas the forecast on day four alone (ignoring antecedent conditions) may lead to precipitation to infiltrate rather than flow overland. Figure 13.2 indicates that on average a successful forecast will be possible with a lead time of four/five days. The figures show clearly that skilful forecasts can exist for the entire forecast range.

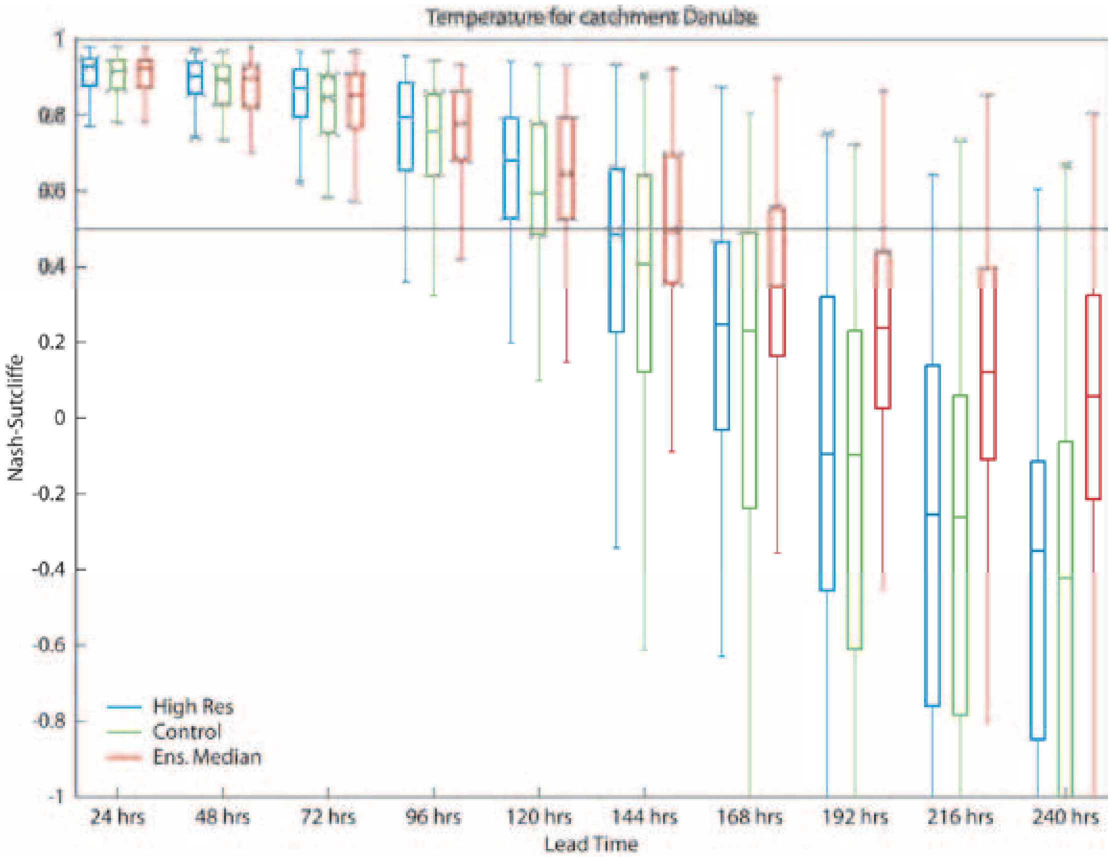



**Figure 13.1** Box and whisker plots of the Nash-Sutcliffe efficiency criteria in comparison to forecast lead time for temperature forecasts accumulated over 24 hrs. Each box has lines at the lower quartile, median and upper quartile. The whiskers are lines extending from each end of the box to show the extent of the rest of the data

## 13.3 TIGGE for hydrological applications (for more details see Pappenberger, F., et al., 2008)

TIGGE is a key component of THORPEX, a World Weather Research Programme to accelerate the improvements in the accuracy of 1-day to 2 week high-impact weather forecasts for the benefit of humanity. The usage of the multi-model TIGGE ensembles (table 13.1) in flood forecasting (i) recognizes that multiple modelling structures may be equally valid representations of the system and application in question (ii) accepts that all models have their inherent weaknesses and strengths (iii) harvests the fact that each model makes use of different information and incorporates information in different ways (for example differences in data assimilation) (Goswami, M., O'Connor, KM., and Bhattarai, KP., 2007). In this way some of the deficits of using the output of a single model EPS are overcome.

The TIGGE archive contains all variables such as precipitation, evaporation, 2 metre temperature which are necessary to drive the hydrological model of EFAS. The hydrological model of EFAS is the LISFLOOD model, a hybrid between a conceptual and a physical rainfall-runoff model with a channel routing (Thielen, J., et al., 2007; de Roo, A., et al., 2003). It is set-up for the whole of Europe on a 5 km grid. At each pixel all information is combined into early flood warning information (Ramos, MH., Bartholmes, J., and Thielen-del Pozo, J., 2007).

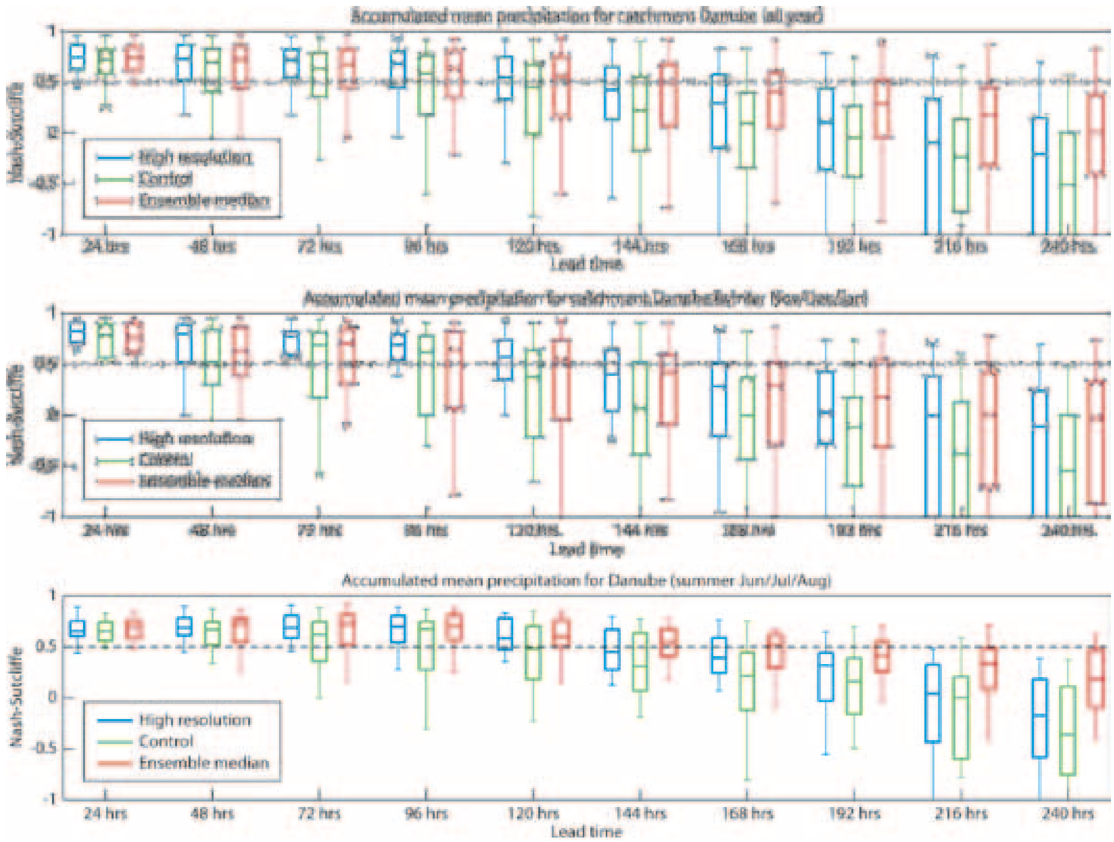



**Figure 13.2** Box and whisker plots of the Nash-Sutcliffe efficiency criteria in comparison to forecast lead time for precipitation forecasts accumulated over the entire lead time. Each box has lines at the lower quartile, median and upper quartile. The whiskers are lines extending from each end of the box to show the extent of the rest of the data. The top plot shows the performance over an entire year, the plot in the middle illustrates the performance over winter and the figure at the bottom shows the range of Nash-Sutcliffe values over summer

EFAS runs without assimilation of observed real time river discharge data (as for example by Komma, J., et al., 2007) which allows the application of this warning system to ungauged catchments (Sivapalan, M., et al., 2003). Four warning levels (*low, medium, high* and *severe*) are derived at each pixel from long-term model simulations with observed meteorological data. The highest simulated discharge of the entire simulated time series represents the highest critical threshold (*severe*) and the 99% quantile, the 2nd highest critical threshold (*high*), the *medium* warning level is defined by 98% quantile and the *low* warning level by the 97% quantile. Analysis of EFAS results through case studies (for example Kalas, M., et al., 2008) and statistical skill score analysis (Bartholmes, J., et al., 2007) has shown that medium-range probabilistic flood forecasts can provide added value information that is complementary to local and national forecasting centres.

The study area is in the country of Romania (central/eastern Europe) and concentratres on rivers Siret, Jiu, Olt and Arges which are tributaries of the River Danube. In October 2007 the hydrological regime of almost all the Romanian rivers was above the monthly mean multiannual values. In the last ten days of the month the precipitation regime exceeded the normal values of the whole month due to the high instability of the weather and torrential rainfall, especially during 20th–24th October and so the discharges in the rivers increased. Flooding was reported in parts of the country especially the south west and east.

**Table 13.1** Meteorological forecast centres and the data used in this study. For the hydrological forecasts only the first 10 days of lead time were used

| Centre | Country / Domain | Ensemble Members | Horizontal Resolution | Vertical Levels | Forecast Length |
|---|---|---|---|---|---|
| Bureau of Meteorology | Australia | 33 | TL119° | 19 | 10 |
| China Meteorological Administration | China | 15 | T213 | 31 | 10 |
| National Centre for Environmental Predictions | USA | 21 | T126 | 28 | 16 |
| UK MetOffice | United Kingdom | 24 | 1.25x0.83deg | 38 | 15 |
| Canadian Meteorological Centre | Canada | 21 | T254 (up to 3.5 days) then T170 | 64 (up to 3.5 days) then 42 | 16 |
| Japan Meteorological Agency | Japan | 51 | TL159 | 40 | 9 |
| European Centre for Medium-Range Weather Forecasts | Europe | 51 | TL399 (up to day 10) | 62 | 15 |

### *13.3.1 Initial Results*

For a location on the river Jiu where severe flooding was observed in Figure 13.3, the probability of exceedance of the high warning threshold for each forecast centre for 13 consecutive forecast dates is shown. This figure concentrates on the onset of the flood (24th of October) and therefore only shows the forecasts for the 11th of October to the 23rd of October. The exceedance levels indicate that most EPS forecasts predict flooding from the 14th of October to the 17th of October. The signal persists from forecast to forecast, which provides the necessary reassurance. This type of persistency is one method used by EFAS (Thielen, J., et al., 2007) to decide whether warnings will be issued. From the 19th October onwards, the signal is very strong, although initially the flooding is predicted one day too early. This means that there is an efficient flood warning five days in advance and a possible warning 8 days in advance.

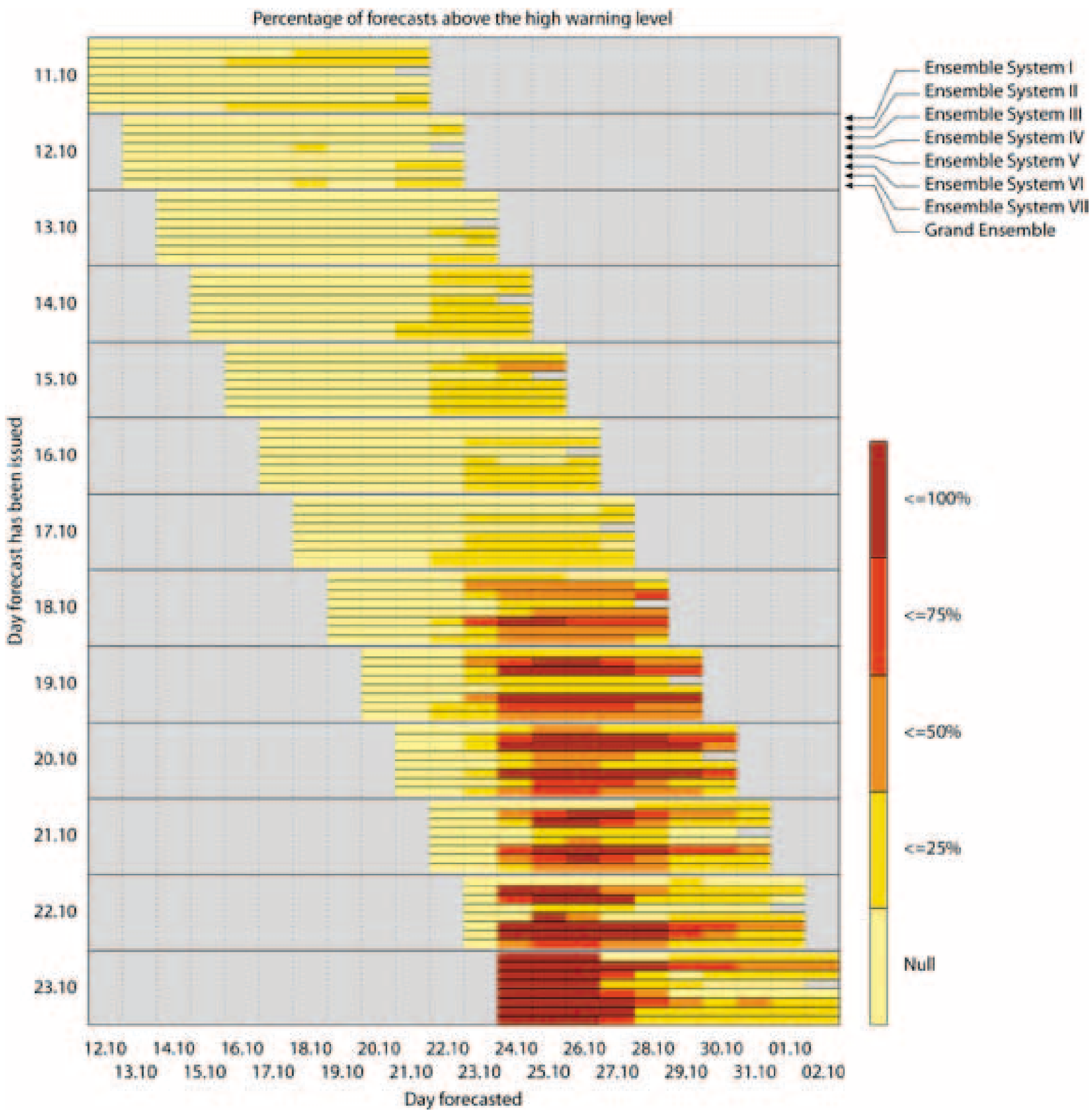



**Figure 13.3** For a location on the river Jiu where severe flooding was observed, percentage of forecasts exceeding the high thresholds from the 11th October to the 23rd October for all forecast systems

Clearly, in this case study the success of the flood warning based on an individual EPS forecast would have depended a great deal on the choice of EPS. Some EPS forecasts missed the event almost entirely and others showed inconsistent results from one forecasts to another (a lack of persistency which would lead to lower flood probabilities being assigned in EFAS). In terms of early flood warning, however, it is the missed events that have more weight: false alarms and hits are identified as the events draw nearer, while missed events lead to late preparations and can initially result in doubts of the short-term forecast results, thus reducing preparedness even more. The grand-ensemble which is composed of several EPS and thus includes a large range of possible weather forecasts, is less likely to miss an event entirely and therefore renders early flood probability estimation more reliable. We have also performed the above analysis for locations where no flooding was observed during the Romanian floods, and results indicated that in this particular case study the flood forecast based on multiple weather forecasts could reduce the false alarm rate.

## 13.4 Conclusions

In this paper the accuracy of ECMWF forecasts for hydrological applications is evaluated. The Nash-Sutcliffe criteria indicates skilful predictions up to day four/five. Forecasts accumulated over a longer time are largely more skilful than forecasts accumulated over short time periods. The analysis in this paper ignores the uncertainty in the observations, which can have a significant influence on model performance and will need to be incorporated into future analysis (see Pappenberger, F., et al., 2007, for discussion).

This paper also showcased the usage of multiple ensemble systems in hydrological forecasting. The grand-ensemble appears to produce reliable results of severe events and therefore can have significant added value for an operational flood forecasting system. Results are based on weather ensemble forecasts available thanks to the TIGGE archive, and on flood predictions generated using the EFAS system for the example of the severe flooding on the river Danube in Romania in October 2007. Although only the analysis of multiple events will allow for a statistically significant evaluation, flood events are rare, river catchment properties are difficult to compare (Bartholmes, J., et al., 2007) and flood events are highly non-linear so an analysis of medium river discharge events has only limited explanatory power for flood events. This work gives encouraging indications that a multi-system grand-ensemble can potentially provide more valuable forecasts than a single ensemble in predicting extreme flood events.

**Acknowledgements.** The work of this paper has been funded by the PREVIEW project.

References

Ahrens, B. and Jaun, S. (2007) On evaluation of ensemble precipitation forecasts with observation-based ensembles. *Advances in Geosciences,* 10, 139–144.

Bartholmes, J., et al. (2007) The European Flood Alert System EFAS – Part II. Statistical skill assessment of probabilistic and deterministic operational forecasts. Hydrological and Earth System Sciences Discussions.

Cloke, HL. and Pappenberger, F. (2007) A framework for screening performance measures to evaluate extreme events for hydrological applications. *Meterological Applications* (accepted).

De Roo, A., et al. (2003) Development of a European Flood Forecasting System. *International Journal of River Basin Management,* 1, 49–59.

Doblas-Reyes, FJ., Hagedorn, R., and Palmer, TN. (2005) The rationale behind the success of multi-model ensembles in seasonal forecasting – II. Calibration and combination. *Tellus Series A-Dynamic Meteorology and Oceanography,* 57, 3, 234–252.

EM-DAT (2007) The OFDA/CRED International Disaster Database – www.em-dat.net – Université Catholique de Louvain, Brussels, Belgium.

Fritsch, JM., et al. (2000) Model consensus. *Weather and Forecasting,* 15, 5, 571–582.

Goswami, M., O'Connor, KM., and Bhattarai, KP. (2007) Development of regionalisation procedures using a multi-model approach for flow simulation in an ungauged catchment. *Journal of Hydrology,* 333, 2–4, 517–531.

Gourley, JJ. and Vieux, BE. (2005) A method for evaluating the accuracy of quantitative precipitation estimates from a hydrologic modelling perspective. *Journal of Hydrometeorology*, 6, 2, 115–133.

Gouweleeuw, B., et al. (2005) Flood forecasting using probabilistic weather predictions. *Hydrology and Earth System Science,* 9, 4.

Kalas, M., et al. (2008) Evaluation of the medium-range European flood forecasts for the March–April 2006 flood in the Morava River. *Journal of Hydrology and Hydromechanics,* in press.

Komma, J., et al. (2007) Ensemble prediction of floods – catchment non-linearity and forecast probabilities. *Natural Hazards and Earth System Sciences,* 7, 4, 431–444.

Krzysztofowicz, R. (2002) Bayesian system for probabilistic river stage forecasting. *Journal of Hydrology,* 268, 1–4, 16–40.

Palmer, TN., et al. (2005) Probabilistic prediction of climate using multi-model ensembles: from basics to applications. *Philosophical Transactions of the Royal Society B-Biological Sciences,* 360, 1463, 1991–1998.

Pappenberger, F., et al. (2005) Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS). *Hydrology and Earth System Science,* 9, 4, 381–393.

Pappenberger, F. and Buizza, R. (2007) The skill of ECMWF predictions for hydrological modelling. Technical Memorandum, ECMWF, 558.

Pappenberger, F., et al. (2007) The skill of probabilistic forecasts under observational uncertainties within the Generalized Likelihood Uncertainty Estimation framework for hydrological applications (accepted).

Pappenberger, F., et al. (2008) TIGGE: Medium Range multi model weather forecast ensembles in Flood Forecasting (A Case Study). Technical Memorandum, European Centre For Medium-Range Weather Forecasts.

Ramos, MH., Bartholmes, J., and Thielen-del Pozo, J. (2007) Development of decision support products based on ensemble weather forecasts in the European Flood Alert System. *Atmospheric Science Letters,* 8, 113–119.

Roulin, E. (2007) Skill and relative economic value of medium-range hydrological ensemble prediction. *Hydrology and Earth System Sciences,* 11, 725–737.

Sivapalan, M., et al. (2003) IAHS decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques,* 48, 6, 857–880.

Thielen, J., et al. (2007) The European Flood Alert System. I. Concept and development. Hydrological and Earth System Sciences Discussions.

Verbunt, M., et al. (2006) Verification of a coupled hydrometeorological modelling approach for alpine tributaries in the Rhine basin. *Journal of Hydrology,* 324, 1–4, 224–238.

Zhu, YJ., et al. (2002) The economic value of ensemble-based weather forecasts. *Bulletin of the American Meteorological Society,* 83, 1, 73–83.

# 14 Drought in the Jizera Mountains, and response in the Uhlířská experimental basin

Alena Kulasová, Libuše Bubeníčková

*Hydrologically speaking, drought is understood as deficiency of water for a certain purpose. It affects various elements of an area, having various impacts on the environment, water management, farming, ecosystems, etc. The cause of drought in the vegetative period, i.e. April to October, is typically a deficit of precipitation not only in the period proper, but also in the preceding winter season, which has affected water reserves in soil during early spring. Higher temperatures during summer can be another factor. Severe frosts in winter also cause reduced flow rates.*

*The subject of our study was the occurrence of periods of minimum flows in April to October on the southern slopes of the Jizera Mountains, specifically on the Kamenice River in Josefův Důl from 1957 to 1983, and in the experimental catchment Uhlířská on the Černá Nisa River from 1982. The hydrological research was extended in the latter area by the effect of drought on waterlogging and soil moisture content.*

## 14.1 Occurrence of periods of minimum flows

### *14.1.1 Introduction*

Drought in watercourses is typically defined by the characteristics of M-day discharges $Q_{300d}$, $Q_{330d}$, $Q_{355d}$, expressing the flow rate, which, over a long period, is exceeded on average for 300, 330, and 355 days in a year, respectively. Hydrological characteristics of unmonitored small streams in the Jizera Mountains were originally determined by hydrological analogy from flow rates registered in the stations of main hydrological net in the base drainage area, while it is possible today to estimate them from the flow rates measured in experimental catchments.

The only basis to estimate the pattern of minimum flow rates before the establishment of experimental basin is the common monitoring period 1957–1983 at Bílý Potok on the Smědá River (catchment area of 26.13 km$^2$) and Josefův Důl on the Kamenice River (25.81 km$^2$). The two basins differ by their gradients as well as geographic orientation, and thence also to the prevalent trajectories of front systems, which can have a significant effect on precipitation subsidy in various meteorological situations, to the thickness and duration of snow cover, as well as other factors that may influence the runoff. A comparison of the results of monitoring of minimum flow rates in both streams, conducted in 2008, confirmed that their occurrences were not always simultaneous.

Uhlířská Basin is one of the seven experimental basins of the Czech Hydrometeorological Institute, established one by one in the Jizera Mountains after 1982. The organizations cooperating in the monitoring of various parameters affecting the environment include the T. G. Masaryk Water Research Institute, the Czech Institute of Technology, the Institute of Hydrodynamics of the Czech Academy of Sciences and others. Beside the measuring net for monitoring water levels and flow rates, rainfall and snow cover, the study includes climatic factors, quality of water and precipitation, hydraulics of soil strata, subsurface discharge, moisture content in soil and the effect of the current climatic and hydrological situation on forest growth. The Uhlířská Basin (area 1.87 $km^2$), the one best equipped with measuring instruments, is part of international ERB system (Euromediterranian Network of Experimental and Representative Basins).

As the Josefův Důl Basin is located on the southern side of the mountains just as the Uhlířská Basin, it was used for the evaluation of the minimum flows in the region.

### *14.1.2 Methodology*

Program LOWFESTIM (NIZOWKA – Low Flow Estimation), developed at the Wroclaw Agricultural College, was used to establish the characteristics of minimum discharges. The number of instances of deficiency periods of five or more days and their combined duration in April to October of each year were determined for the limit flow rates $Q_{300d}$, $Q_{330d}$, $Q_{355d}$ in the Josefův Důl and Uhlířská Basins.

The hydrological characteristics adopted for Josefův Důl were taken from the flow sequences in the period 1957–1983. The limit flow rates found were $Q_{300d}$ = 0.315 $m^3/s$, $Q_{300d}$ = 0.260 $m^3/s$, and $Q_{355d}$ = 0.160 $m^3/s$. The limit flow rates in the period 1982–2008 determined for Uhlířská were $Q_{300d}$ = 0.017 $m^3/s$, $Q_{300d}$ = 0.013 $m^3/s$, and $Q_{355d}$ = 0.009 $m^3/s$.

The dependence of minimum flow rates on the pattern of rainfall was tested using the index of antecedent precipitation, calculated for the closest climatic station at the Bedřichov Reservoir. The station, also located on the southern side of the mountains in the basin of the Černá Nisa River, is a suitable point also to derive the rainfall pattern in the basin of the Kamenice River, located further east. The index of antecedent precipitation expresses total rainfall of the previous 30 days, replenishing moisture in the adjacent basin. It is a simplified characteristics of possible saturation of the basin, not taking into account any other factors affecting the hydrological process. The index is invalid in winter, when precipitation during frost spells is deposited as snow reserve and does not replenish the basin until thaw.

The diagrams represent the season from April to October (though April is usually still affected by replenishment from the preceding winter season) by showing the sequence of daily discharges and the occurrence of flow rates equal or minor to $Q_{300d}$, $Q_{330d}$, and $Q_{355d}$ lasting 5 or more consecutive days, index of antecedent precipitation and mean daily temperature at the Bedřichov station. Each year was evaluated visually.

### *14.1.3 Results*

The driest seasons at Josefův Důl appear to be April to October 1963 and 1973, when minimum discharges lasted very long, gradually dropping to lower values, which is why even flow rates equal or minor than $Q_{330d}$ occurred for a greater number of days. In the year 1963 the drop of discharge values continued further, the duration of $Q_{355d}$ was 33 days. So little flow rates did not occur in 1973. Discharges $Q_{355d}$ were also registered in 1968 and 1982, but only for short spells. None minimal flow rates occurred in 1958, 1965–1967, 1974, 1977, and 1979 to 1981.

After the Uhlířská Experimental Basin was established, it has been possible to evaluate the occurrence of minimum discharges in the same way as at Josefův Důl. The greatest numbers of minimum flow rates were recorded in 1992, 2003, 2007 and 2008, as well as in 1982 and 1983.

For illustration, see several figures showing the most frequent minimum flow rates:

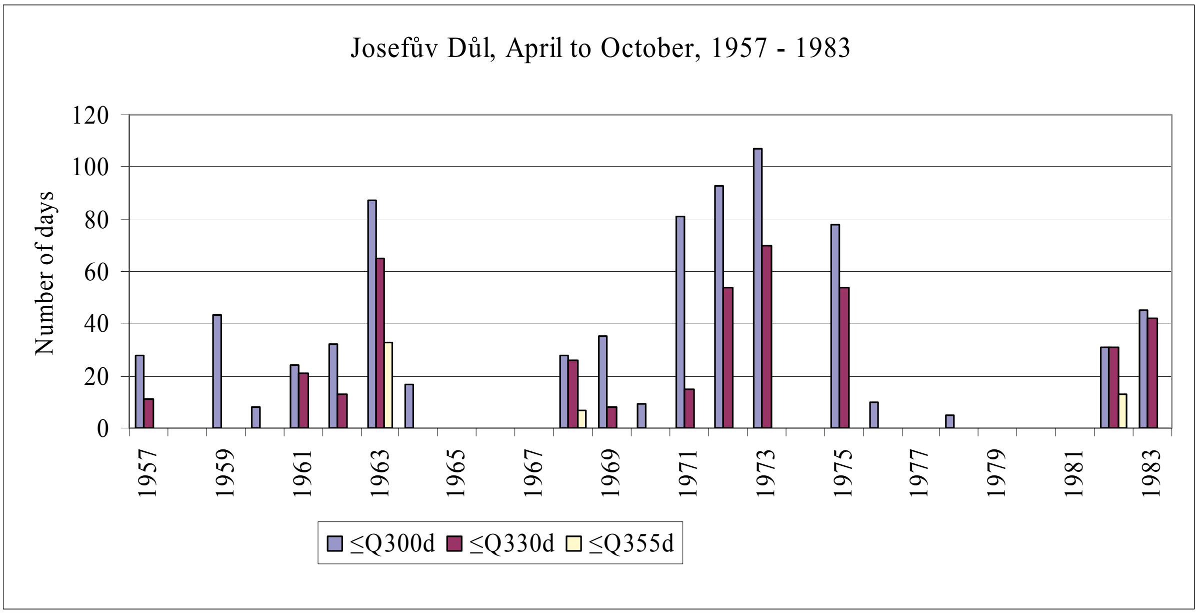



**Figure 14.1** Number of days of occurrence of discharges minor or equal to $Q_{300d}$, $Q_{330d}$, and $Q_{355d}$ at Josefův Důl in April to October 1957–1983 with minimum 5 days duration of deficiency period

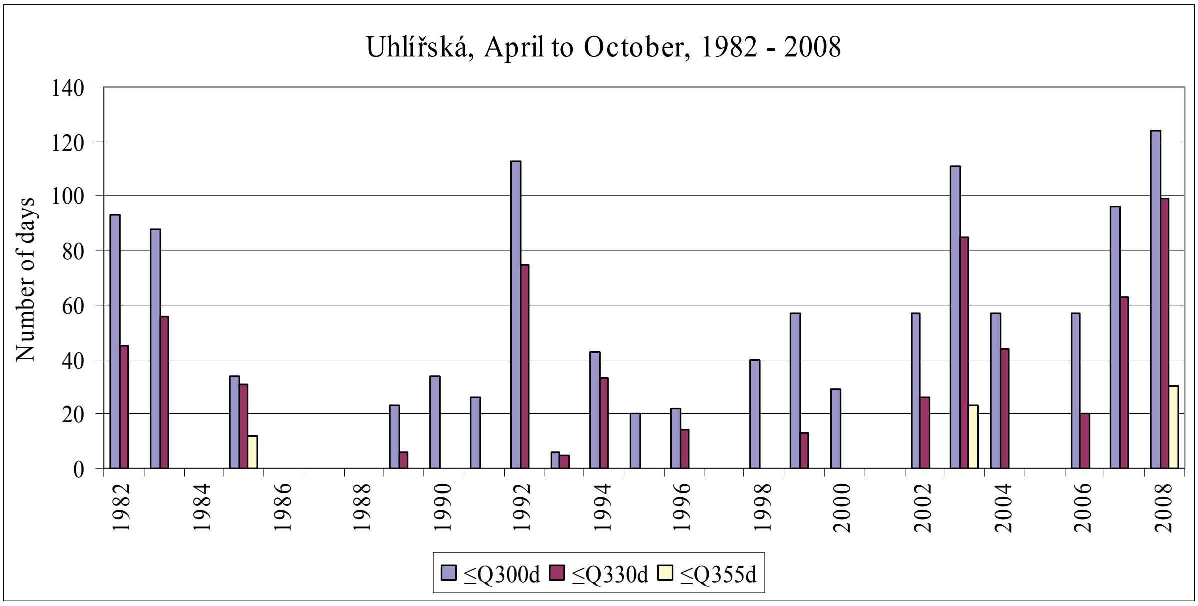
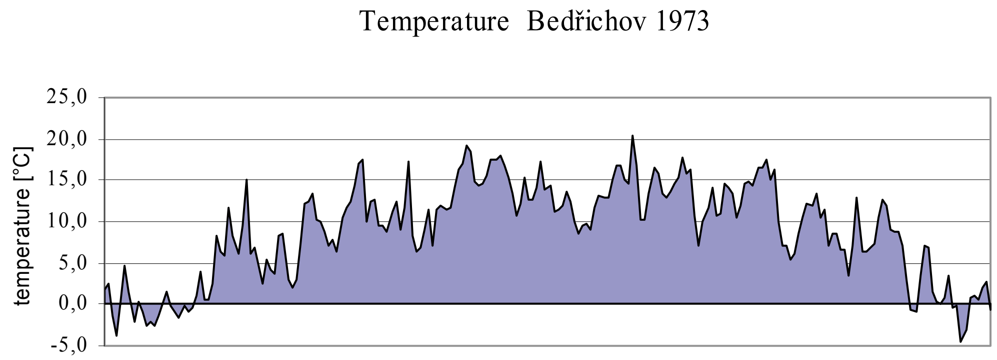



**Figure 14.2** Number of days of occurrence of discharges minor or equal to $Q_{300d}$, $Q_{330d}$, and $Q_{355d}$ at Uhlířská in April to October 1982–2008 with minimum 5 days duration of deficiency periods

The drought in 2002 was manifested as early as in mid-June, and the same applied to the following year 2003. The trend was disrupted by an exceptional rainfall on August 12–13, 2002. Small flow rates still continued in September, while temperature was declining.

Monitoring in 1982 and 1983, conducted in both basins, allowed a comparison of the occurrence of minimum flows between the larger basin of the Kamenice at Josefův Důl and the small Uhlířská basin on the Černá Nisa.



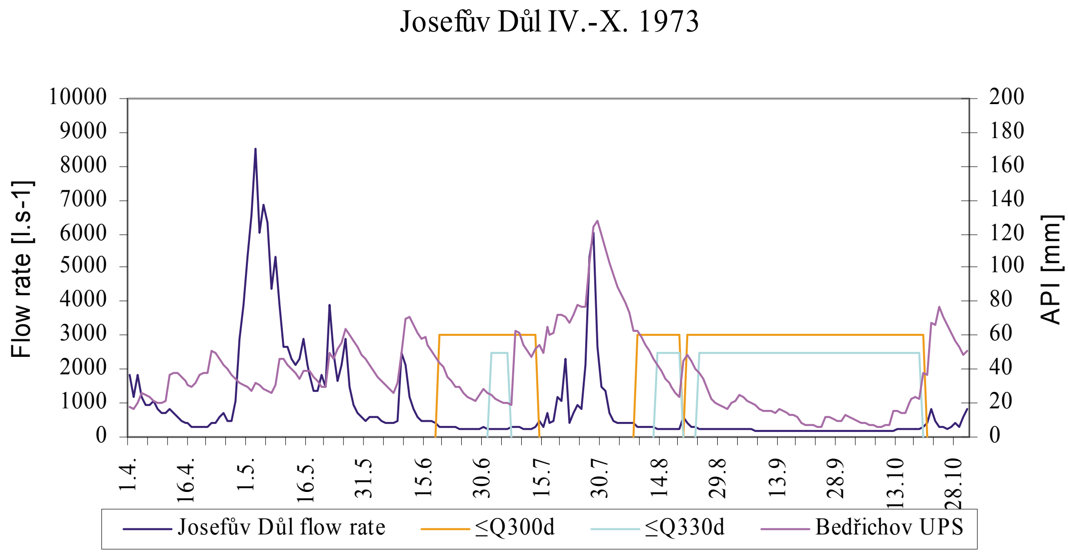




**Figure 14.3** Comparison of occurrence of minimum flow rates with the index of antecedent precipitation and mean daily air temperature at Josefův Důl, April to October 1963 and 1973

Minimum flow rates in April to October 1983 were less frequent at Josefův Důl than at Uhlířská. The pattern was similar in 1982.

In 52 years (1957 to 2008) in April to October the hydrological drought with minimal discharges in the Jizera Mountains minor or equal to $Q_{300d}$ of duration 5 or more days, lasting 40 or more days in total, occurred 17 times, specifically in 1959, 1963, 1971, 1972, 1973, 1975, 1982, 1983, 1992, 1994, 1998, 1999, 2002, 2003, 2006, 2007 and 2008.

Counting all days with flow rates minor or equal to $Q_{330d}$ (i.e. not only instances 5 or more consecutive days long) in the season from April to October, a somewhat higher total number of occurrences can be arrived at: 94 (instead of 93) in 1982, 90 (88) in 1983, 118 (113) in 1992, 75 (57) in 1999, 128 (111) in 2003, 83 (57) in 2004, 103 (96) in 2007, and 145 (123) in 2008.

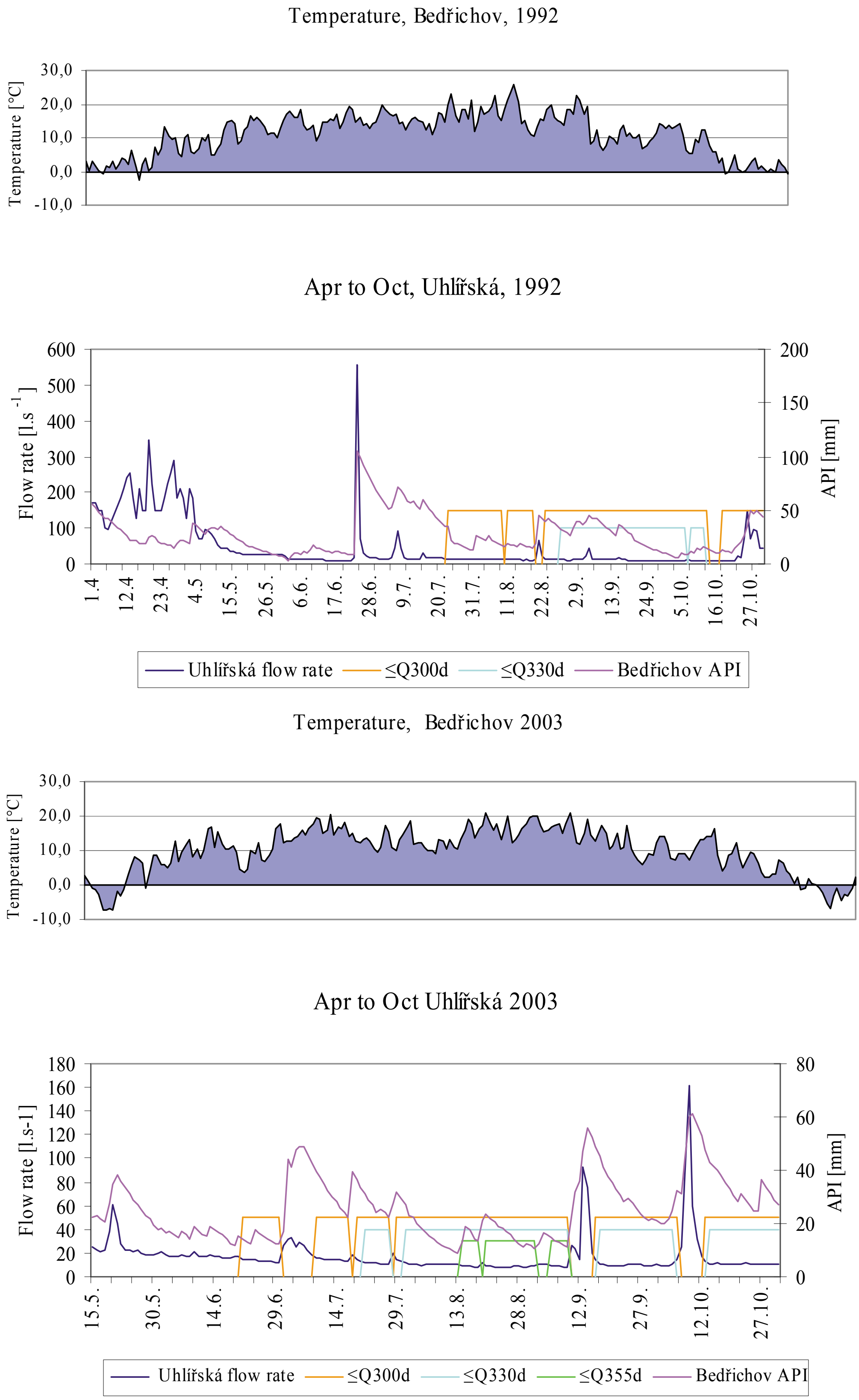



**Figure 14.4** Comparison of occurrence of minimum flow rates with the index of antecedent precipitation and mean daily air temperature at Uhlířská, April to October 1992 and 2003

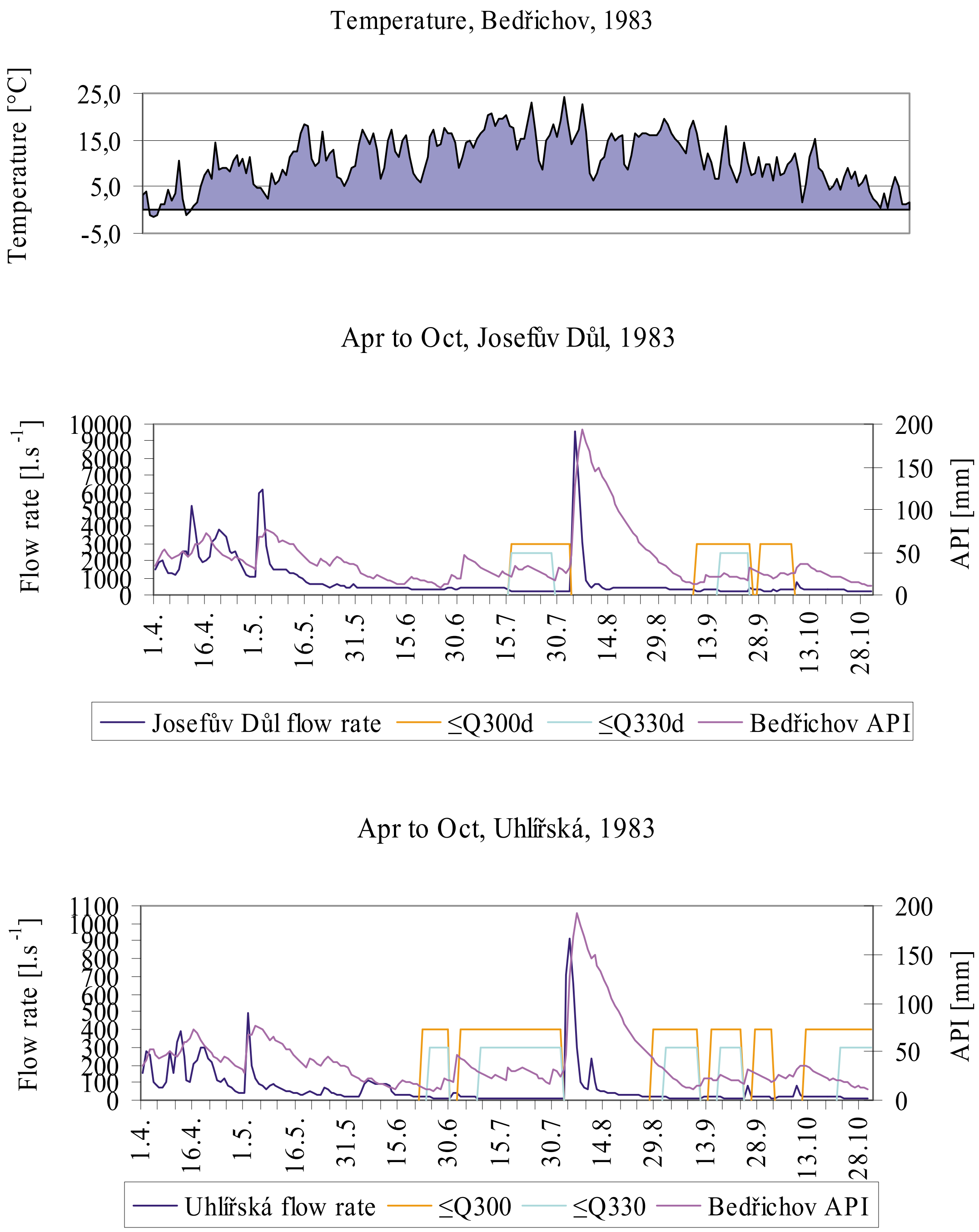



**Figure 14.6** Comparison of occurrence of minimum flow rates at Josefův Důl on the Kamenice and at Uhlířská on the Černá Nisa, April to October 1983

- Should the minima stay a longer time, they continue for some more time even after an increase of precipitation.
- Without strong rainfall the minimum flow rates continue in autumn in spite of the drop of air temperature. When the temperature drops under 0 °C, the low flow rates persist from the preceding period, precipitation is deposited as snow, and the minima often last throughout the winter season, e.g. from late November and December 1962 till January to March 1963.
- Minimum flow rates also follow strong frost spells, e.g. in December 1973.

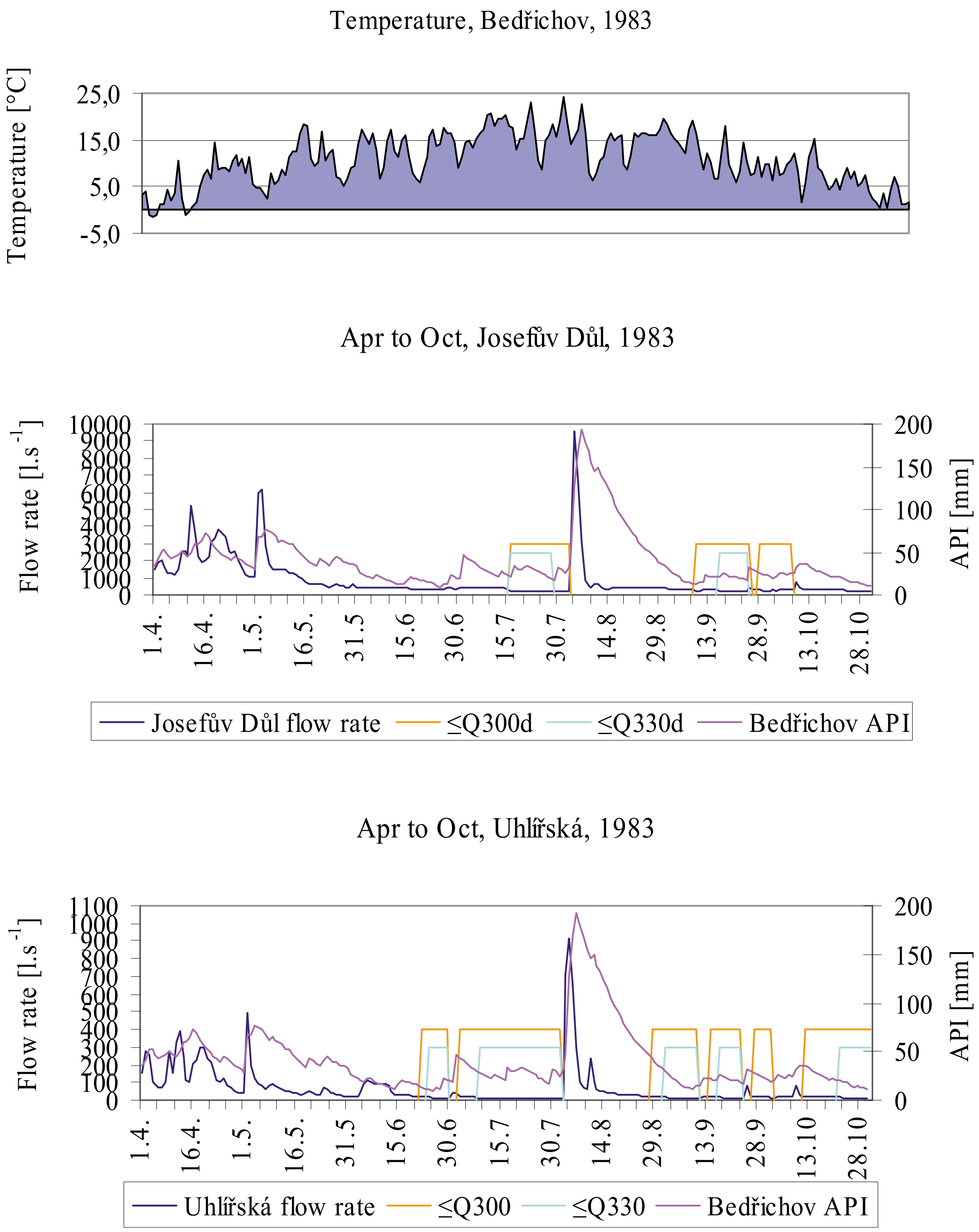



**Figure 14.6** Comparison of occurrence of minimum flow rates at Josefův Důl on the Kamenice and at Uhlířská on the Černá Nisa, April to October 1983

- Should the minima stay a longer time, they continue for some more time even after an increase of precipitation.
- Without strong rainfall the minimum flow rates continue in autumn in spite of the drop of air temperature. When the temperature drops under 0 °C, the low flow rates persist from the preceding period, precipitation is deposited as snow, and the minima often last throughout the winter season, e.g. from late November and December 1962 till January to March 1963.
- Minimum flow rates also follow strong frost spells, e.g. in December 1973.

The above findings do not allow detecting of clear general regularities, which would always apply to the basins monitored, let alone to other mountain basins. The occurrence and duration of hydrological minima are affected by a number of other factors, difficult to identify, such as the size of the basin, soil conditions, vegetation type, extent of peat moors, etc. One should also take a critical approach to the input data, precision of water readings and measurement curves, and especially the pattern of precipitation, as it is impossible to rule out absolutely the occurrence of local extremes in a monitored basin. Although precipitation measurements in the experimental basins proved that it usually rains throughout the monitored area, the rainfall to the same extent does not affect the basins, and flow response is not always identical, i.a. due to the different geographical configuration.

## 14.2 Response to drought period in the Uhlířská Experimental Basin

### *14.2.1 Introduction*

Not only water level in the stream, but the entire basin reacts to deficient precipitation. The saturation of the basin can be monitored in the field using tensiometers (measuring suction pressure of the soil) or piezometers (measuring shallow water table), or by the simple "boot method" used to check if the soil is saturated during field visits.

Soil saturation in the experimental basin of the Czech Hydrometeorological Institute at Uhlířská is monitored using tensiometers by the Faculty of Civil Engineering of the Czech Institute of Technology and the Institute of Hydrodynamics of the Academy of Sciences; the value of pressure loss in the tensiometer corresponds to the suction pressure of soil water.
If the soil is dry, it sucks water from the porous vessel of the tensiometer until equilibrium is attained. If the soil is moist, water is bound in soil by lesser forces, and a part is suctioned back into the vessel, where the pressure loss was greater than the suction pressure of soil water and the tensiometers.

The T. G. Masaryk Water Research Institute Management monitors soil saturation in the same basin using the "boot method". The criterion is whether water forms a free surface when the ground is stepped on. The object of the method is to verify the dynamics of saturated areas, especially after the melting of snow, after strong rains and in periods of drought.

Several dry periods have occurred during the monitoring period since 1995. Vegetative periods 2002 and 2003 were selected for comparison of the tensometric monitoring and the "boot method".

### *14.2.2 Climate in 2002 and 2003*

Winters in 2002 and 2003 were normal in terms of temperature (-0.6 °C, 2.1 °C respectively), but below the normal in terms of precipitation (445.4 mm, 399.5 mm). Snow cover lasted till March 30 in 2002, till April 15 in 2003. Total April precipitation was subnormal in both years (41.7 mm, 47.2 mm).

The following months from May to July in both years were also subnormal in terms of rainfall.

Bohemia suffered abundant and intense rains in August 2002. The ridge parts of the Jizera Mountains were also affected. It was mildly raining on August 11 already. Rainfall on the next day (August 12) amounted to 52.1 mm at Bedřichov and August 13 was the rainiest day with a daily total of 169.5 mm. In total 228 mm of rainfall were recorded at Bedřichov from August 11 to August 14. The monthly total was at 221.5% of the long-term normal. Precipitation was subnormal in the following month of September and again above the average in October.

The months of September and October were subnormal in 2003. Total precipitation in October was on the long-term average. Individual monthly totals from May to October in 2002 and 2003, with percentage points of the long-term mean, are shown in Table 14.1.

**Table 14.1** Climatic station Bedřichov: monthly totals of rainfall and their percentage of the long-term normal in years 2002 and 2003

| Month | Long-term normal 1961–1990 | 2002 | Percentage of long-term normal | 2003 | Percentage of long-term normal |
|---|---|---|---|---|---|
| **Climatic station Bedřichov** | | | | | |
| May | 109.2 | 42.0 | **38.5** | 57.8 | 52.9 |
| June | 122.8 | 64.0 | 52.1 | 40.3 | **32.8** |
| July | 134.8 | 113.9 | 84.5 | 91.0 | 67.5 |
| August | 135.8 | 300.7 | **221.5** | 29.4 | **21.7** |
| September | 94.6 | 65.0 | 68.7 | 77.9 | 82.4 |
| October | 86.9 | 143.4 | 165.0 | 93.7 | 107.8 |
| May to Oct | 684.1 | 729.0 | 106.6 | 390.1 | 57.0 |

### *14.2.3 Sites chosen for the monitoring of saturated areas*

*Location of tensiometers*

The Institute of Hydrodynamics (Academy of Sciences) deployed tensiometers to two sites in the Uhlířská Basin, in a small clearing amid a young spruce stand (age class 1, i.e. up to ten years), and a full-grown spruce and partly beech forest (age 40–70 years, with some beech trees up to 90 years old).

Suction pressures in soil are measured in couples at the depths of 15, 30, 45, and 60 cm. Manual measures are taken in a time span of 7–10 days.

*Mapping of saturated areas using the "boot method"*

The T. G. Masaryk Water Research Institute Management mapped saturated areas in the Uhlířská Basin since 1995. Individual points in the basin were mapped by 1997; the mapping of a transect was initiated in the same year. The first observation of the movement of water in land reclamation trenches was conducted in spring 1999. The transect was slowly grown over by vegetation and consequently the mapping of the saturated areas was distorted. For this reason chiefly selected reclamation trenches on the left bank of the Černá Nisa were mapped in 2002–2003 for comparison of saturation. The trenches had been restored by the Forest Administration in Jablonec nad Nisou in autumn 1998 to drain water from wet sites planted with a new stand.

Three reclamation trenches were selected for the monitoring, each consisting of a main trench and several lateral flumes. Water flow in the trenches was the basis for the mapping. The scale of observation is shown in Table 14.2.

**Table 14.2** Scale for the description of trench waterlogging

| |
|---|
| Degree 1: water in trench streams fast, foaming ( I – water streams – foam) |
| Degree 2: water streams through trench ( I – water streams) |
| Degree 3: level of still water ( II – level) |
| Degree 4: strong waterlogging (III – squelching) |
| Degree 5: moist only – not squelching ( IV – moist) |
| Degree 6: dry (V – dry) |

The trenches were walked one by one along the main trench as well as the flumes. The condition of the trenches was recorded according to the scale in metres of length (e.g. 16 m I – water streams, 8 m II – level). For final assessment, total length of the trenches was added up and split by the individual degrees of waterlogging.

### *14.2.4 Vegetative period 2002*

*Tensiometry*

Tensiometers were deployed in mid-May. The measurement was conducted from May 24 to October 4, 2002. Suction pressures in the clearing were quite balanced throughout the monitoring period (50–144 kPa). The highest suction pressures occurred in May and June, the lowest after the rain on August 17.

Due to the drought, water was drained in the forest by the roots of trees, which is why the suction pressures slowly grew till August 9, when the highest value was recorded. Rain in mid-August saturated the soil, causing the lowest values of suction pressure. The pressures in the forest were balanced in September and October (see Figure 14.7).

*Observation in reclamation trenches*

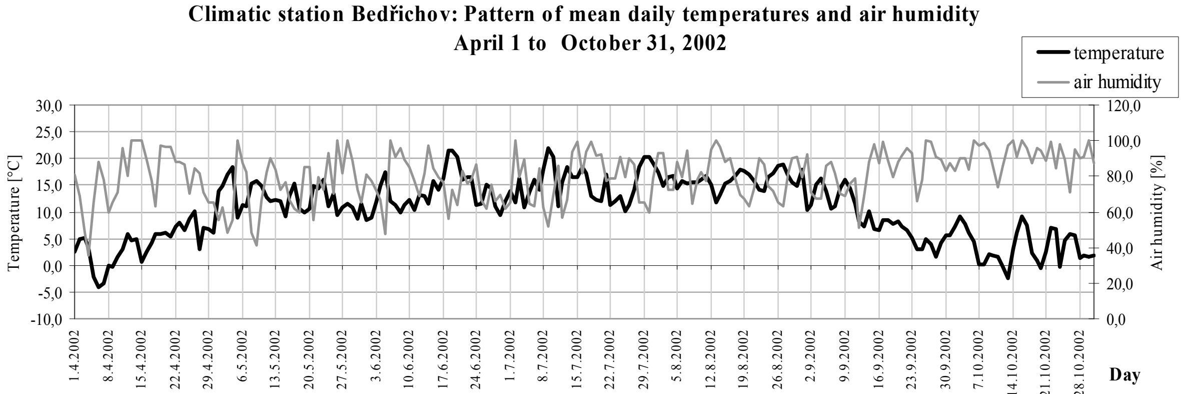
The observation after winter was conducted in a saturated basin. Water was streaming fast both in the main trench and the lateral flumes. Total rainfall in May was only 38% of the long-term mean. The mapping was conducted on May 18. Water in the trenches was flowing very slow and was rather retained at barriers. It stood still in the lateral flume, squelching under dry grasses. The ends of the flumes were turning dry, only their floors being still muddy.

The lowest degree of saturation of the basin was observed on July 1. The trenches were generally dry, with water standing only in the shaded upper sections of the main trenches. The overgrown lateral flumes were moist; peat was crumbling off the dry edges. Sphagnum moss was turning yellow. Two more observations were conducted in July after minor rains, which did not saturate the basin.

The highest degree of saturation was found after a strong rain on August 17. Water was streaming fast in the trenches and foamed in places. The ends of the trenches, overgrown with moss, squelched when stepped into. More observations took place in autumn in stable weather conditions. The basin was neither as dry as in early July nor as saturated as in mid-August. The observation of the trenches in 2002 was affected by the adjacent growing stand.

Young spruce trees grow in the neighbourhood of the upper sections of the trenches, increasingly covering the trenches with their branches. Wetland grasses (*Junus conglomeratus* L., *Deschampsia caespitosa* (L.) P. BEAUV., *Poa trivialis* L., *Carex pilulifera*) gradually overgrow the trench floors and banks. Roots and bodies of the plants form barriers, which retain water.

The following Figure 14.7 shows the pattern of mean daily temperatures, air humidity and total daily rainfall at the Bedřichov climatic station against the sequence of mean daily flow rates of the Černá Nisa at the Uhlířská outlet from April 1 to October 31, 2002. Figures 14.8 and 9 are graphic representations of tensiometry and mapping of reclamation trench A in the vegetative period of 2002.



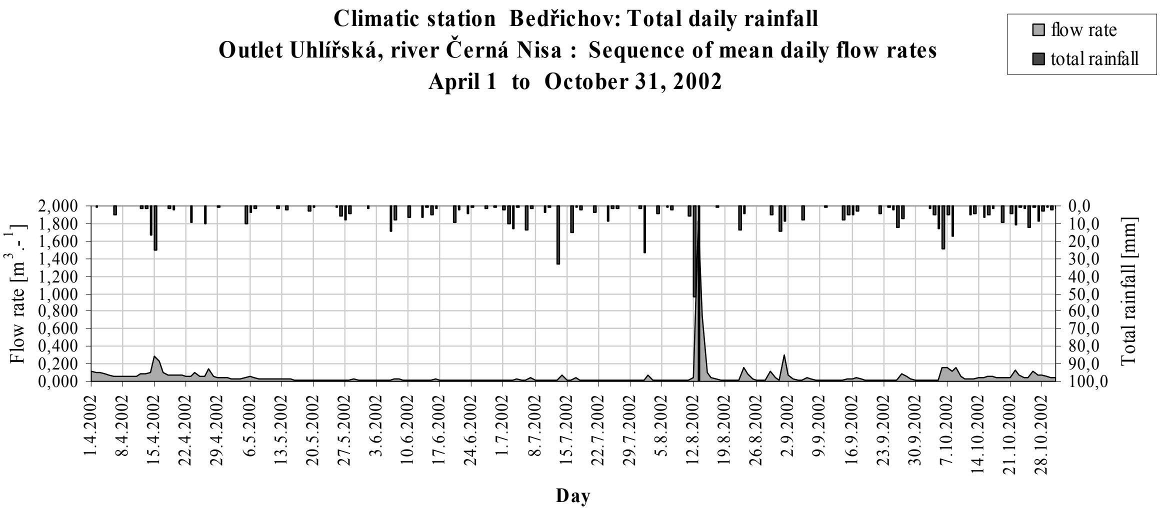




**Figure 14.7** Pattern of mean daily temperatures, air humidity and total daily rainfall at Bedřichov climatic station, and sequence of mean daily flow rates in the Černá Nisa at Uhlířská outlet from April 1 to October 31, 2002

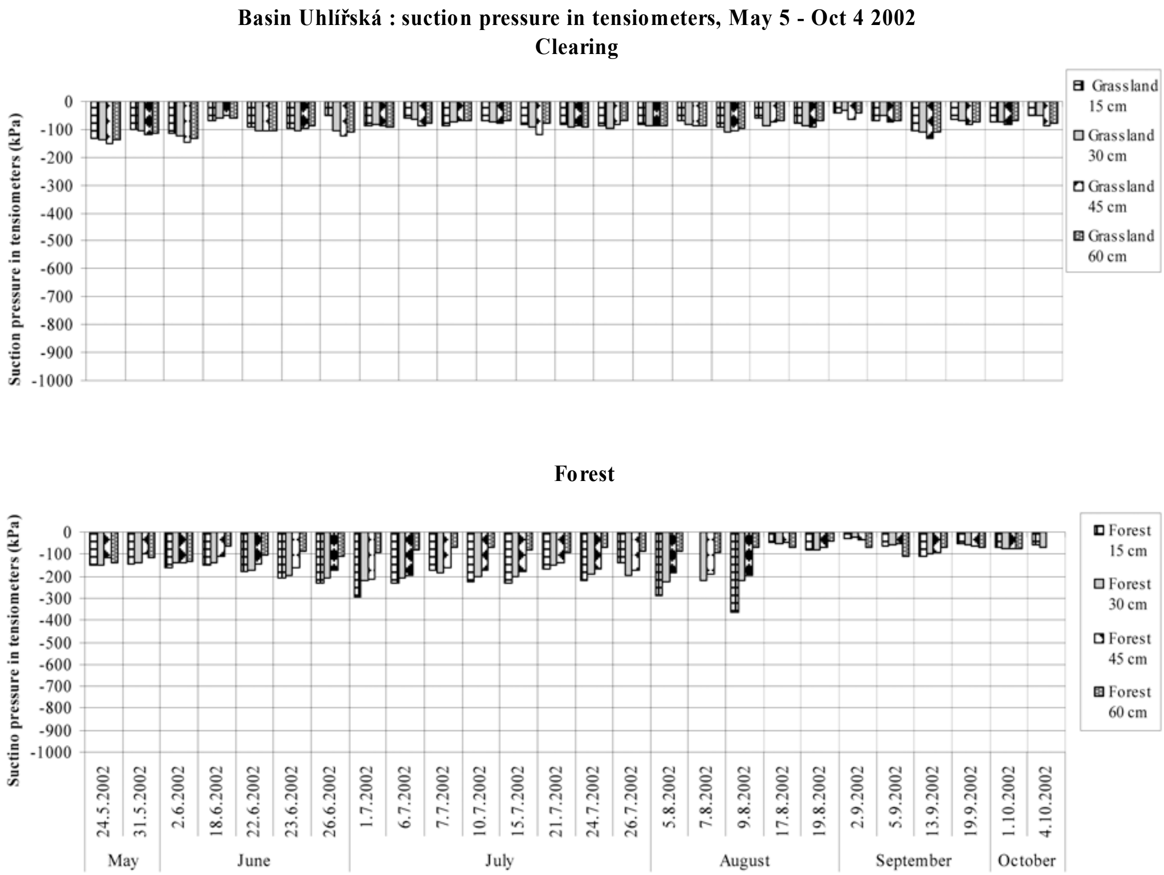



**Figure 14.8** Tensiometric results in the clearing and in forest, Uhlířská experimental basin, from May 24 to October 4, 2002

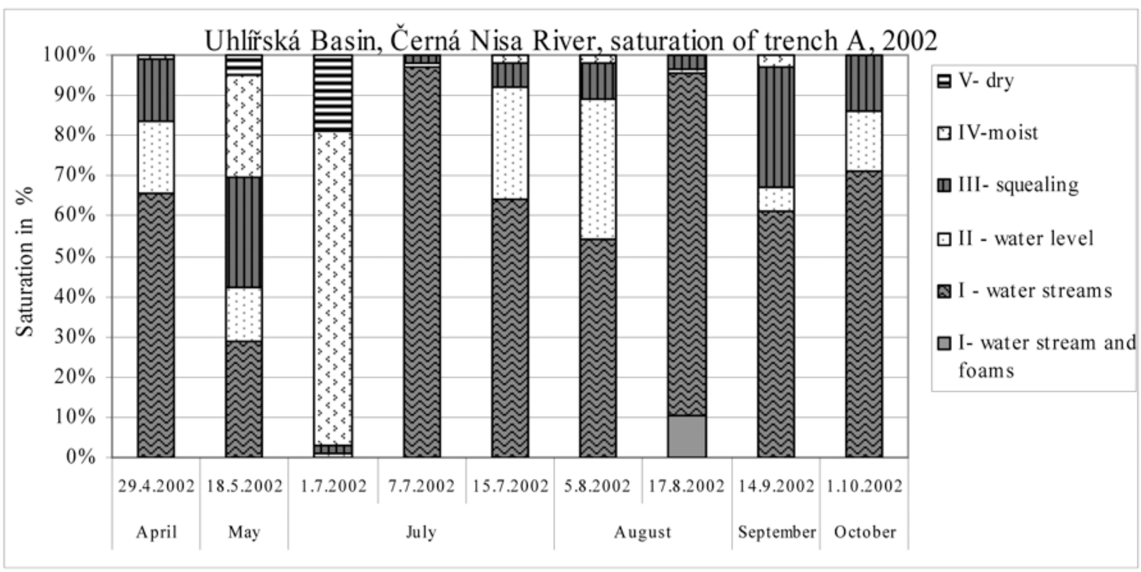



**Figure 14.9** Results of mapping of reclamation trench A in Uhlířská experimental basin from April 29 to October 1, 2002

### *14.2.5 Vegetative period 2003*

*Tensiometry*

Manual tensiometry was conducted in the clearing and in forest from May 2 to October 4, 2003. The values from the clearing were quite level till the end of May. Suction pressures gradually grew from June 11 to June 30, due to lasting dry weather. Lower suction pressures were measured on July 1, following rains (a total of 46.7 mm from June 29 to July 1). The pressures subsequently grew again due to dry weather. They reacted by a moderate decrease on August 15 and 19 to rains on August 13, 14, 18 and 19. After that they grew again due to limited rainfall, especially at the shallow depths of 15 cm and 30 cm. They did not drop again until mid-September after several days of rain (total rainfall from September 9 to 14 was 61.4 mm).

Tensiometers in the forest showed gradual growth of suction pressures, with a temporary drop after the rains of June 29 and July 1, analogously to the clearing. Tensiometers in forest did not respond to the August rains, and the suction pressures kept increasing, attaining a maximum in early September. The highest values were attained at shallow depths (15 cm, 30 cm), later also at 45 cm. The pressures dropped after the September rain, but grew again later. Suction pressures in the tensiometers in forest responded more strongly to the lasting dry weather than in the clearing.

*Observation in reclamation trenches*

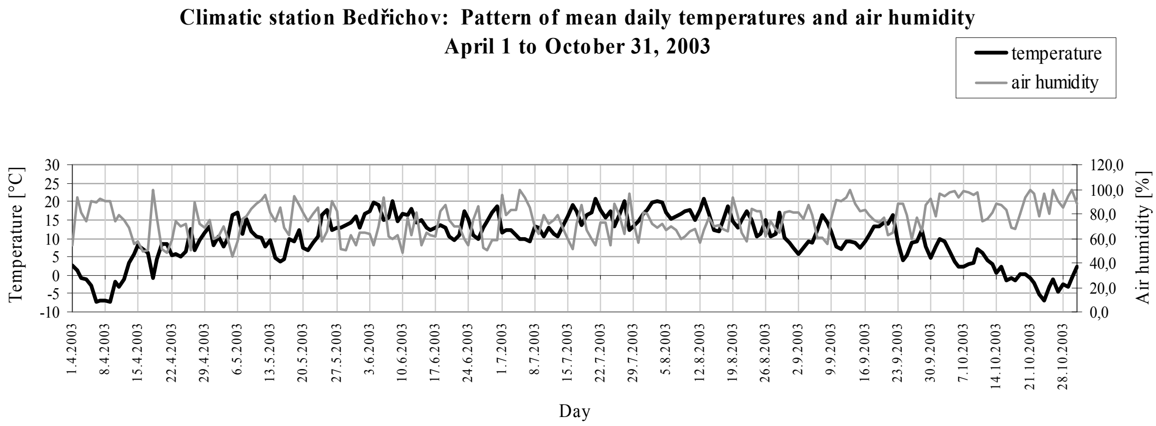
It was found in early May 2003 that the basin was still saturated. Following the rains, water was streaming, though not foaming, in the trenches on May 20. The next round of mapping took place at the end of May. Water in the trenches was slowly streaming under grass. Where clusters of grasses grew, the slowly flowing water was retained. The trenches were rather dry at the time of the June observation; shaded places were moist.

The first mapping in July was conducted on 2/7 after a rain. Water in trenches streamed slowly, standing still in places, or surfaced when moss or grass were stepped on. The next observation took place in dry weather on July 31. The banks of the trenches were crumbling off and sliding to the floor. Water in upper parts of the trenches squealed in places on stomping, but most of the trenches were only slightly moist. Two rounds of mapping were conducted during dry spell in August. The trenches were dry on August 12, and totally desiccated on August 29; sphagnum moss was yellow and edges of banks were crumbling to the floor.

Two rounds of mapping were conducted in September, the first on the 15$^{th}$, after rains (total rainfall from September 9 to 15 was 79.1 mm). Water stood in places in the trenches and squealed under grass and moss; some places were only moist (gravel). The next round of mapping took place five days later. There had been no rainfall before the mapping, so the trenches were only slightly moist. Water squealed under grass only in the upper sections of the trenches.

Two rounds of mapping were conducted in October, the first on 4/10. Water stood in some places in the trenches, squealed under grass, or the trenches were moist. The trenches were full of water on October 11 after rain; the water was flowing off in places.

The following Figure 14.10 shows the pattern of mean temperatures, air humidity and total daily rainfall at the Bedřichov climatic station during the period from April 1 to October 31, 2003, against the pattern of mean daily flow rates in the Černá Nisa at the Uhlířská outlet.



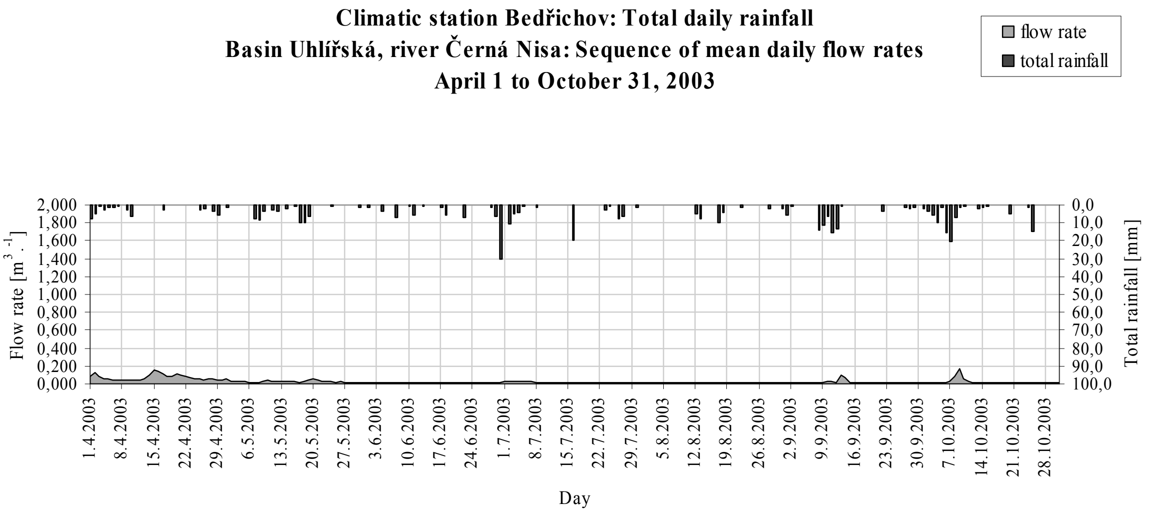




**Figure 14.10** Pattern of mean daily temperatures, air humidity and total daily rainfall at Bedřichov climatic station, and pattern of mean daily flow rates at the Uhlířská outlet from April 1 to October 31, 2003

Figures 14.11 and 14.12 show graphic representations of tensiometry and mapping of reclamation trenches in the vegetative period 2003.

The results of monitoring of the degree of saturation of the basin in the vegetative periods of 2002 and 2003 were similar at the beginning (subnormal winter precipitation, dry spring seasons).

The movement of water in reclamation trenches was similar in May and June of both years. Water and waterlogged spots occurred both in the main trench and the lateral flumes. The findings in July were similar. Water and waterlogged spots occurred only in the main trench, peat on the sides of the trenches was crumbling off, sphagnum moss was desiccating.

Tensiometry also yielded similar results. The soil was drying over time; the loss of moisture was slower in the clearing than in forest.

The results from the monitoring seasons began to differ after the intense rains in August 2002. Water was observed on August 17 to stream fast and even form foam; suction pressures in tensiometers were similar in the clearing and in forest.

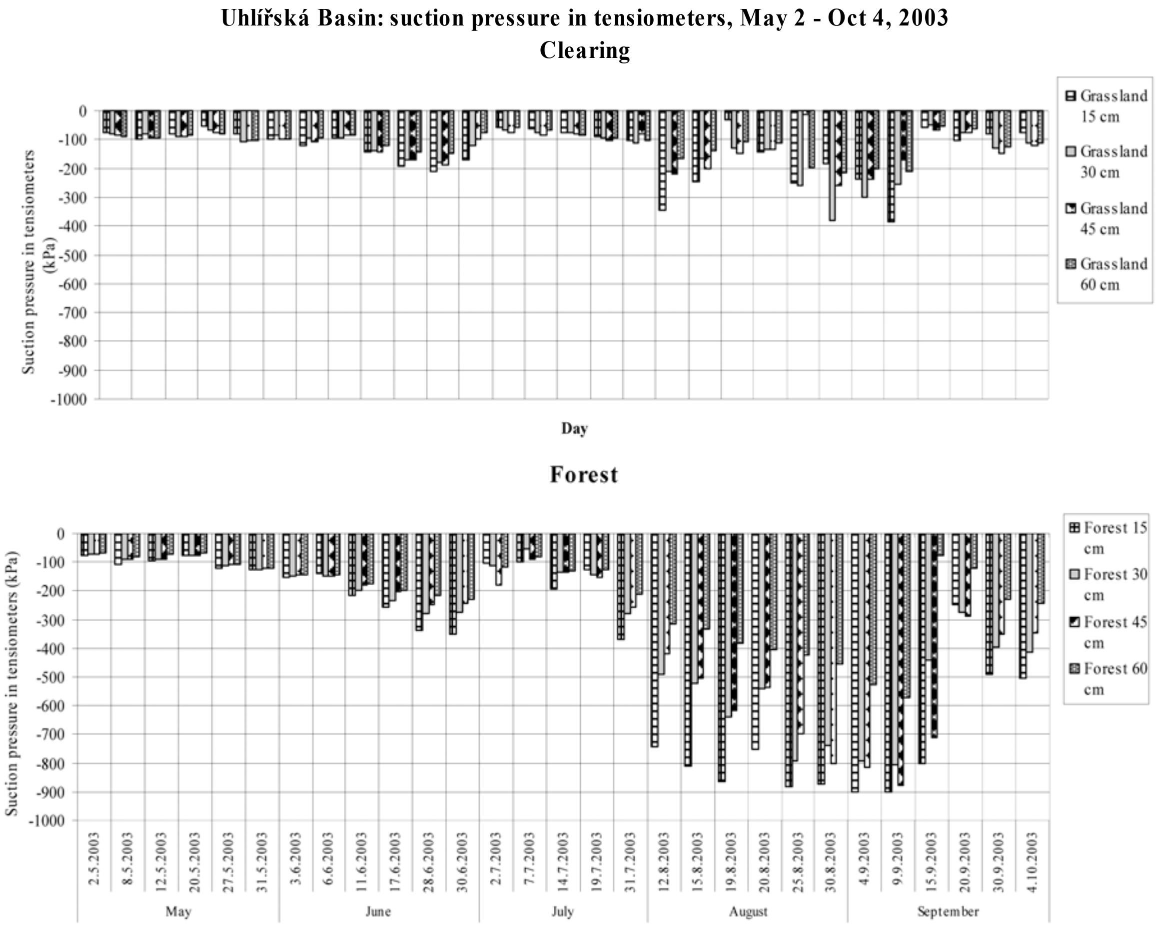



**Figure 14.11** Tensiometric results in Uhlířská experimental basin in the clearing and in forest from May 2 to October 4, 2003

On the contrary, as no major rains occurred in 2003, the ground continued to dry in the basin. The trenches were totally dry, sphagnum moss was also dry and grasses were yellow. Suction pressures attained maximum, especially in forest, at first close to the surface (15 cm, 30 cm), later also at greater depths. According to the Forestry and Game Management Research Institute lasting drought causes deficiencies in the nutrition of spruce stands, namely the intake of nitrogen, magnesium and phosphorus.

Water and wet spots were found in the reclamation trenches after rains in September and October; water was streaming again in the trenches at the time of the last round of mapping in October.

On the contrary, tensiometry suggested that the rains in September and October did not suffice to saturate the soil profile, especially in forest.

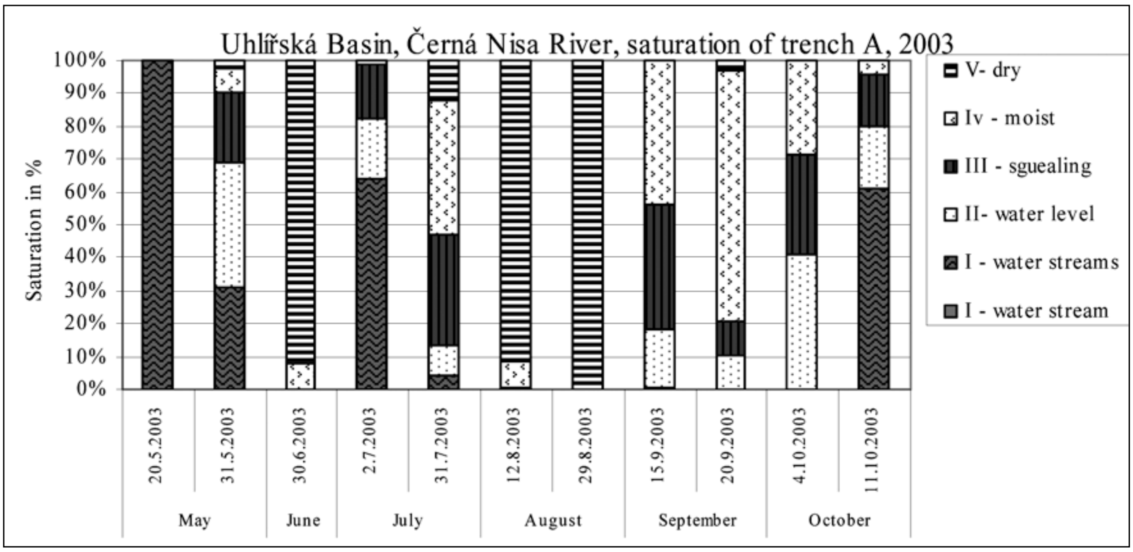



**Figure 14.12** Results of mapping of reclamation trench A in Uhlířská experimental basin from May 20 to October 11, 2003

### *14.2.6 Conclusions*

The results presented suggest that the Jizera Mountains are no exception in that dry spells recur frequently at different degrees of severity, manifested by the duration of low flow rates in the streams. The pattern of minimum flow rates corresponds to higher pressures in tensiometers and desiccation of reclamation trenches, which are important suppliers of water to forest stands and undergrowth during dry spells in small basins. Climatic observations and adequate networks of rain gauges in the basins.

**Acknowledgements.** The project was supported by the Czech Grant Agency under the number P209/11/2045.

References

Data base of climatic and hydrological measurements 1957–2008.

Data base of tensiometric results 2000–2009.

Slodičák, M. (2005) Lesnické hospodaření v Jizerských horách. Výzkumný ústav lesního hospodářství a myslivosti, Jíloviště-Strnady, Prague. ISBN 80-86461-51-3.

Romportl, D. a Bubeníčková, L. (2008) Výskyt minimálních průtoků v období 1957–1983 na horních tocích Smědé a Kamenice. Proceedings of Adolf Patera Workshop, Prague 2008, p. 117–125.

Kulasová, A. (2009) Ověřování predikcí modelů odtoku mapováním nasycených ploch v terénu. Research project MZP002071101, Výzkumný ústav vodohospodářský T. G. Masaryka, v.v.i., Prague.

## List of authors

| Author | Contanct | Chapter |
|---|---|---|
| Arnaud Patrick | Cemagref, U.R. OHAX, 3275 Route de Cézanne, 13182 Aix-en-Provence Cedex 5, France | 7 |
| Beven Keith J. | Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK<br>Geocentrum, Uppsala University, 73256 Uppsala, Sweden<br>ECHO, ISTE, EPFL, Lausanne, CH-1015, Switzerland | 0, 1, 5 |
| Blažková Šárka D. | T. G. Masaryk Water Research Institute, p.r.i., Prague, Czech Republic | 0, 5 |
| Bolgov Mikhail | Water Problems Institute of Russian Academy of Sciences, 119991, Moscow, Gubkin St., 3 | 4, 9 |
| Bubeníčková Libuše | Czech Hydrometeorological Institute, Na Šabatce 17, 143 00 Praha 4, Czech Republic, bubenickova@chmi.cz | 14 |
| Jones David A. | Centre for Ecology & Hydrology, Wallingford, UK, daj@ceh.ac.uk | 3 |
| Kašpárek Ladislav | T. G. Masaryk Water Research Institute, p.r.i., Prague, Czech Republic,_ladislav_kasparek@vuv.cz | 8, 10 |
| Kjeldsen Thomas R. | Centre for Ecology & Hydrology, Wallingford, UK, trkj@ceh.ac.uk | 3 |
| Kiczko Adam | Institute of Geophysics, Polish Academy of Sciences, ul. Ks. Janusza 64, 01 452 Warsaw, Poland | 11 |
| Kulasová Alena | T. G. Masaryk Water Research Institute, Podbabská 30, 160 00 Praha 6, Czech Republic, alena.kulas@vuv.cz | 14 |
| Lang Michel | Cemagref, U.R. Hydrologie-Hydraulique, 3 bis Quai Chauveau, 69336 Lyon cedex 09, France | 7 |
| Lavabre Jacques | Cemagref, U.R. OHAX, 3275 Route de Cézanne, 13182 Aix-en-Provence Cedex 5, France | 7 |
| Montanari Alberto | University of Bologna, Via del Risorgimento 2, I-40136 Bologna, Italy, alberto.montanari@unibo.it | 2 |
| Muller Aurélie | Cemagref, U.R. Hydrologie-Hydraulique, 3 bis Quai Chauveau, 69336 Lyon cedex 09, France | 7 |
| Napiórkowski Jarosław J. | Institute of Geophysics, Polish Academy of Sciences, ul. Ks. Janusza 64, 01 452 Warsaw, Poland | 11 |
| Nol Ondřej | Aquatest, a.s., Prague, Czech Republic | 12 |
| Novický Oldřich | T. G. Masaryk Water Research Institute, p.r.i., Prague, Czech Republic, oldrich_novicky@vuv.cz | 8, 10 |
| Osipova Nadezhda | Water Problems Institute of Russian Academy of Sciences, 119991, Moscow, Gubkin St., 3 | 4 |
| Pappenberger Florian | European Centre for Medium-Range Weather Forecasts, Reading, RG2 9AX, UK | 13 |
| Romanowicz Renata J. | Institute of Geophysics, Polish Academy of Sciences, ul. Ks. Janusza 64, 01 452 Warsaw, Poland, Environment Centre, Lancaster University, Lancaster, LA1 4YQ, U.K | 11 |
| Rudiš Miroslav | T. G. Masaryk Water Research Institute, p.r.i., Prague, Czech Republic | 12 |
| Skaugen Thomas | Hydrology Dept., NVE, P.O.Box 5091, 0301 Oslo, Norway, ths@nve.no | 6 |
| Valenta Petr | Civil Engineering Department of Czech Technical University, Prague, Czech Republic | 12 |
| Valentová Jana | Civil Engineering Department of Czech Technical University, Prague, Czech Republic | 12 |

Šárka D. Blažková et al.

# Floods, droughts and prediction uncertainties

Published by:
T. G. Masaryk Water Research Institute, public research institution
Prague, 2011



1st edition

Number of issues: 200, 9,2 AA

Coordinating editors: Josef Smrťák, Sylva Garcíová

Cover design and typeset: TAG Michal Korecký

ISBN 978-80-87402-13-9