Safety and Reliability – Bedford & van Gelder (eds)

Application of various techniques to determine exceedence probabilities of water levels of the IJssel Lake

F.L.M. Diermanse, G.F. Prinsen & H.F.P. van den Boogaard
WL | Delft Hydraulics, Delft, Netherlands

F. den Heijer
RIKZ, Den Haag, Netherlands

C.P.M. Geerse
RIZA, Lelystad, Netherlands

ABSTRACT: Three different methods are applied to derive exceedance probabilities of water levels of the IJssel Lake (the Netherlands). In the first method an extreme value distribution function is identified from observed water levels. In the other two methods a numerical model is used to simulate a series of synthetic events. Subsequently, computed maximum water levels and occurrence probabilities of the synthetic events are combined in a numerical integration procedure to derive exceedance probabilities of water levels in the IJssel Lake. The main difference between the second and third method lies in the schematisation of the physical processes involved. The second method uses random variables with a clear physical meaning, whereas the third method uses a principal component representation. Comparison of the three methods shows that resulting exceedance probabilities of water levels vary strongly.

1 INTRODUCTION

Without adequate flood protection, almost two-thirds of the Netherlands would be regularly flooded. Therefore, an extensive system of dikes and dunes has been constructed in order to prevent floods of the sea, rivers and lakes. The low lying part of the Netherlands is divided into 53 so-called dike-ring areas. Each dike ring area has its desired safety level, that is stated by law (Wet op de waterkeringen, 2002). This safety level, typically a small value in the order of 10^{-3} to 10^{-4} per year, is based on both the economic value of the protected area and the extent of the threat. The same law states that every five years an evaluation is executed to verify whether the flood defence structures still offer the desired safety level. As a consequence, exceedence probabilities of extreme water levels and wave heights in the major water systems in the Netherlands are regularly updated.

This paper deals with the derivation of exceedance probabilities of extreme water levels in the IJssel Lake, one of the primary water systems in the Netherlands. Currently, exceedance probabilities of water levels in the lake are derived from a fit of an extreme value distribution function through a set of observed peak water levels. The main disadvantage of this type of method is that one looks for a regularity in measured water levels without taking into account the fact that these water levels are caused by highly non-linear and inhomogeneous processes. This observation has lead to the suggestion (e.g. Klemeš 1994, Kuchment et al. 1993) that progress in flood frequency analysis can only be expected from methods which explicitly take into account the involved physical processes.

In this study, two of such "physically based" methods are developed and applied to derive exceedance probabilities of extreme water levels of the IJssel Lake. The results of these two methods are compared with results of a purely statistical method.

2 SYSTEM AND MODEL DESCRIPTION

2.1 *The IJssel Lake*

With a total surface of 1,182 km^2, the IJssel Lake is the largest lake in the Netherlands. In the South East part of the lake the IJssel river, a distributary of the river Rhine, drains on average approximately $400\,m^3/s$ to the lake. Furthermore, a number of smaller rivers and canals (here referred to as regional contributaries) contribute an average accumulated discharge of about $165\,m^3/s$ to the IJssel Lake.

In the North, the lake is separated from the sea by a dike (the "Afsluitdijk"). Before the realisation of this dike in 1932, the area which now forms the IJssel Lake was still part of the sea. The Afsluitdijk is equipped with a number of sluices, enabling the drainage of surplus river water into the sea. The volume of water that discharges through the sluices depends on the sluice control, the water level in the lake and the sea level. The policy behind the sluice control is to keep the average water level at −0.40 m. +NAP in winter and at −0.20 m. +NAP in summer (where "m. +NAP" is the Dutch national reference level, which is approximately at mean sea water level).

2.2 *High water events*

High water levels in the IJssel Lake are the result of a period of 1 to 3 weeks during which the accumulated discharge of the IJssel river and the regional contributaries exceeds the outflow through the sluices of the Afsluitdijk into the sea. Peak water levels are generally observed a few days after a peak discharge on the IJssel river has occurred. High water events in the IJssel Lake occur almost exclusively in the winter half year since that is the season for high water events on both the IJssel river and the regional contributaries.

Besides the IJssel river and regional contributaries, wind (direction as well as velocity) and sea water level also significantly influence (maximum) lake water levels since they largely control the drainage of surplus water from the lake into the sea.

The maximum observed water level in the lake (since the start of measurements in 1932) is 0.50 m. +NAP and it occurred in October 1998. Extreme water levels like that in 1998 can only lead to floods if they occur in combination with a storm that causes wind set up in the Lake. However, the latter phenomenon will not be subject of this paper, i.e. only the dynamics of the area-average lake water level are considered. For an uncertainty analysis of water levels in the IJssel Lake which does include effects of wind set up, the reader is referred to Vrijling et al. (1999).

2.3 *The WINBOS model*

Two of the three methods presented in chapter 3 require a physically based numerical model to simulate water level dynamics in the lake during (hypothetical) events. For this purpose the model "WINBOS" is used. The model input consists of time series of processes that are relevant for the lake water dynamics, such as discharges

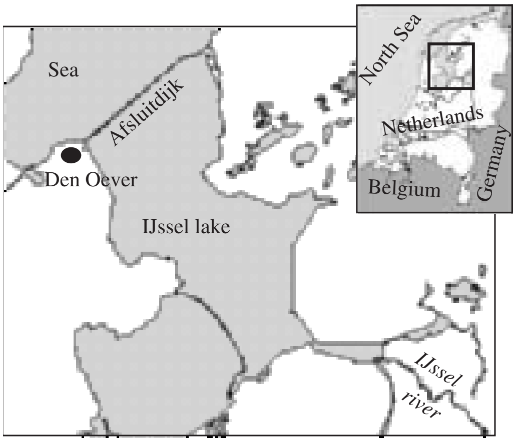

Figure 1. Location of the IJssel Lake and the IJssel river in the Netherlands.

of the IJssel river and the regional contributaries, wind direction, wind velocity and sea level. The model computes Lake-average water levels on the half-hourly time-scale. The model uses simple mass balance equations and user defined rules for sluice control. The WINBOS model has been validated and applied successfully in previous studies of the IJssel Lake dynamics (WL 1997, RIZA 1999).

3 METHODS

3.1 *Introduction*

This chapter describes the three methods applied to derive statistics of extreme water levels in the IJssel Lake. The first method consist of fitting an extreme value distribution function through the series of observed annual maxima and subsequently extrapolating this function to derive extreme value statistics.

In contrast to the first method, the other two methods explicitly take into account the physical processes that are responsible for the lake's water level dynamics. For a large number of synthetic inputs the WINBOS model is used to generate the corresponding water levels of the IJssel Lake. Combining the statistical description of the input processes and the associated water levels as computed with the model, the statistics of extreme water levels of the IJssel Lake are determined through numerical integration.

The main difference between the two numerical integration methods lies in the schematisation of the input processes. In the second method the schematisation is rather straightforward using a number of random variables with a "direct" and clear physical meaning (e.g. peak river discharge, sea water level, wind direction). In the third method the schematisation is much more complex and indirect. First, principal components of the multi-dimensional space of measured input variables are derived. Subsequently, linear combinations of the principal components are used to represent input variables of the synthetic events.

3.2 *Method 1: Extreme value distribution functions*

The first method is a purely statistically based fit-procedure of observed water levels. This type of procedure is still the most popular in extreme value analysis, mainly because it is relatively easy to apply.

The first step in the procedure consists of the selection of a subset of the available series of measurements. Since statistics of high water levels are to be derived, this subset should consist of relatively high water levels as well. Generally, either the series of annual maxima is selected or, alternatively, the series of peak water levels that exceed a user defined threshold (the so called peaks over threshold, or POT series).

Secondly, a prior estimate of the exceedance probability is derived for each observation of the selected subset. This estimate is based on the length of the period of observation and the number of times the concerning peak water level has been exceeded during this period. If the series of annual maxima has been selected, the prior estimate generally is derived as follows:

$$P = \frac{r - a}{N + b} \tag{1}$$

where P is the prior estimate of the yearly exceedance probability, r is the rank number of the peak water level (1 = highest observed peak water level, 2 = second highest peak water level etc.), N is the total number of years of observation and a and b are user defined constants. If, for instance, Gringortens method is followed, a equals 0.44 and b equals 0.12 (Shaw 1988).

The third step consists of selecting a probability distribution function, preferably from the group of extreme value distribution functions (e.g. Pareto, Gumbel, Weibull).

The fourth and final step is to determine the parameter values of the selected distribution function that provide a "good fit" for the selected subset of observations. A number of procedures are available to derive the fit, of which the method of moments and the maximum likelihood method are most commonly applied.

The derived probability distribution function provides a relation between exceedance probabilities (or recurrence intervals) on one hand and related water levels on the other hand. This means the water level with a recurrence interval of, for instance, 1000 years can be derived from the distribution function, even if the available series of observations covers less than 100 years.

In the method described above the user has to select:

- a subset of observations;
- a method to obtain the prior estimate;
- a probability distribution function; and
- a fit procedure.

The prior estimate is used to visualise the goodness of fit of the selected distribution function. However, the prior estimates have no influence on the resulting parameters of the probability distribution function. In contrast, the other three choices mentioned above can have a strong effect on the final outcome, which is demonstrated in chapter 4.

3.3 *Method 2: numerical integration*

3.3.1. *Introduction*

In method 1 it is assumed that the derived probability distribution is valid for all water levels, even those that are much higher than the maximum observed water level. In doing so, the function is extrapolated outside the range for which it is proven to be valid. One of the consequences is that this method does not take into account the fact that the nature of the physical processes during extreme (unobserved) events may change in comparison with observed events. For instance the dominant role of floodplains during extreme events may not be reflected in the observed events.

Therefore, a second method is introduced that explicitly takes into account the physical processes responsible for water level dynamics in the IJssel Lake. In this method, hypothetical events are simulated with the WINBOS model (see section 2.3). The length of simulated periods is 30 days, which is sufficient to simulate the rise and fall of the water level in the lake. Such a hypothetical event is characterised by a quantitative description of the relevant physical processes. The physical features considered are:

- the discharge of the IJssel river;
- the accumulated discharge of the regional contributaries;
- the sea water level at Den Oever;
- the wind velocity; and
- the wind direction.

The entities described above are treated as stochastic variables. The following five steps are executed in order to derive exceedance probabilities of the water level in the IJssel Lake:

1. A relatively straightforward representation and quantification is defined for the stochastic variables.
2. For each stochastic variable the range of possible realisations is divided into a number of classes.

3. For each stochastic variable the available series of observations is used to derive the probability of occurrence of the classes defined.
4. For all possible combinations of realisations of the stochastic variables (which is limited as a result of the defined classes) the resulting water level in the lake is computed with the WINBOS model.
5. The probabilities as derived in step 3 and the water levels as computed in step 4 are combined to derive the statistics of extreme water levels of the IJssel Lake through numerical integration.

Some of the steps mentioned above are described in more detail below.

3.3.2 *Representation of stochastic variables*

The stochastic variables represent processes of a dynamic nature, i.e. with a strong temporal variability. (and in some cases some spatial variability as well). For practical purposes a relatively simple representation of these is required. It is vital, though, that this representation still contains the features that are relevant in the generation of high water levels in the IJssel Lake.

The discharge of the IJssel river is characterised by the peak discharge that occurs in the simulated event. The temporal variation of the discharge is represented by a standardised hydrograph, the shape of which is based on statistics of measured hydrographs. The assumed hydrograph is such, that the peak discharge occurs on day 15 of the simulated event (the simulated period is 30 days).

The sea water level during a single tide is modelled as a sine function with an amplitude of 0.75 m. The mean of the sine function is assumed to be constant for a period of 5 days. Since a period of 30 days is simulated this means a total of 6 stochastic variables is required to represent the temporal variation of the sea water level. The choice of a period of 5 days is a trade-off between an accurate representation of temporal variability on the one hand and the required number of model simulations on the other hand.

In contrast to the discharge of the IJssel river, the combined discharge of the regional contributaries is assumed to be constant over the simulated period of 30 days. The same assumption is done for the wind direction. Finally, the wind velocity is taken as a function of both wind direction and sea water level. This means the wind velocity is assumed to be constant over periods of 5 days, similar to the sea water level.

3.3.3 *Probability distributions of the stochastic variables*

The required statistical features of the IJssel river discharge, the sea level at Den Oever and wind were taken from previous studies (HKV 2001, RIKZ 2000). These statistics have been derived by fitting extreme value distribution functions through available data in a similar manner as described in section 3.2. The available statistics of the sea level were converted from the tidal scale to the 5-daily scale, since the latter time-scale is used in the current analysis.

Statistics of the regional contributors were not readily available and consequently needed to be derived during the current analysis. Again, a data fit by an extreme value distribution was derived.

3.3.4 *Definition of classes*

In order to obtain a good coverage of the range of possible realisations, 6 classes were defined for the IJssel river discharge, 8 classes for the regional contributaries, 7 classes for the sea water level and 2 classes for the wind direction. This means the total number of realisations (and, consequently, required WINBOS runs) is equal to $8 \times 6 \times 7^6 \times 2 \approx 11{,}300{,}000$. In order to reduce the number of computations, the stochastic variables representing the sea water level of the first 5 days and the last 5 days of the simulated period of 30 days were replaced by deterministic values. This is allowed because the sea water level at the beginning and at the end of the simulated event has no influence on the maximum water level in the lake. Maximum water levels in the Lake generally occur around day 15 when the discharge of the IJssel river is at its peak. The number of required computations is now reduced to $8 \times 6 \times 7^4 \times 2 \approx 230{,}000$.

3.3.5 *Numerical integration procedure*

For a user-defined water level of the IJssel Lake the simulated events during which this water level is exceeded are identified. The probability of exceedance of the water level is equal to the accumulated probability of occurrence of these events. The available numerical integration tool assumes there is no correlation between the stochastic variables, the effect of which is discussed later on. The procedure is repeated for a range of water levels to obtain a probability distribution function.

A disadvantage of this procedure is the fact that the amount of model simulation runs required easily gets out of hand, as demonstrated in section 3.3.4. Depending on the computation time of a single model run one may be forced to reduce the amount of stochastic variables and output classes as much as possible. However, this may cause unacceptable loss of accuracy.

3.4 *Method 3: numerical integration based on principal component analysis (PCA)*

3.4.1 *Introduction*

The quantitative description of the physical processes responsible for the high lake levels plays a significant role in the method described in the previous section. Generally, a relatively simplified representation was used in order to limit the required model simulations.

The method presented in the current section offers an alternative that allows a much more dynamic representation of the processes involved. Furthermore it has the advantage that the (statistical) dependency between the processes is automatically taken into account.

The method is based on a mathematical technique called "principal component analysis" (PCA). PCA is a multivariate analysis or ordination technique that can be applied to detect common features in multiple sets of measurements. The available sets are represented by vectors of equal length, M, and as such form a multidimensional swarm of data points. The main idea of PCA is to find the directions in the M-dimensional space with the largest spread or scatter, i.e. the directions that explain the largest part of the total variance of the data set. For more detailed descriptions of PCA the reader is referred to Haan (1977) and Pielou (1984).

The PCA-analysis is used here to detect common features in observed patterns of the involved processes during high water events. This information is subsequently used to generate an ensemble of synthetic input series for the WINBOS model. As in method 2, the resulting water levels computed by the WINBOS model are used to derive exceedance probabilities of water levels through numerical integration. The method is further outlined below.

3.4.2 *PCA analysis of pre-event histories*

The first step in the analysis is to select a number (K) of high water events for which measurements of all relevant processes/stochastic variables (i.e. IJssel river discharge, sea water level, etc.) are available. For each event the measured values are "stored" in a single vector ν:

$$\nu = \left(x_{-T}^{(1)},..,x_0^{(1)};x_{-T}^{(2)},..,x_0^{(2)};....;x_{-T}^{(N)},..,x_0^{(N)}\right) \qquad (2)$$

where $x_{-t}^{(j)}$ refers to the observed value of the jth stochastic variable at t time steps before the maximum water level in the lake occurs, N is the number of stochastic variables and T is the number of time steps in the observed period.

Since K events are selected we have K vectors in a $N \times (T+1)$-dimensional space that offer a rather extensive representation of the pre-event histories. As these vectors are expected to be correlated to some extent, they are likely to contain redundant information. The main purpose of the PCA analysis is to reduce the dimension of these vectors in such a way that no vital information is lost. In doing so, the set of vectors is written as a linear combination of "base-vectors" φ and coefficients (or principal components) α:

$$\nu^{(k)} = \sum_{d=1}^{D} \alpha_d^{(k)} \varphi_d \ ; k = 1..K \qquad (3)$$

where φ_d is the dth base vector and $\alpha_d^{(k)}$ is the dth coefficient of the kth vector $\nu^{(k)}$.

The base-vectors φ_d are selected as eigenvectors of the covariance-matrix of the K vectors $\nu^{(k)}$. If the full set of K eigenvectors is selected, each vector $\nu^{(k)}$ can be exactly reproduced by the linear combination in Equation (3) above by selecting the proper α-values. Furthermore, other combinations of α-values will result in "new" ν-vectors which can be considered as mathematical representations of synthetic events. The α-values will therefore be treated as stochastic variables in order to generate a set of synthetic events. These synthetic events are subsequently applied in a numerical integration procedure, similar to the one presented in section 3.3, in order to derive exceedance probabilities of water levels in the IJssel Lake.

However, if K eigenvectors (equal to the number of selected observed events) are used, the number of stochastic variables is generally too large to handle in the numerical integration procedure. In order to reduce the number of stochastic variables, a subset of the eigenvectors is selected. First, the eigenvectors are ranked, based on the magnitude of the corresponding eigenvalues λ_k. The magnitude of λ_k is a measure of the fraction of the total variance of the set of vectors ν which can be explained by eigenvector φ_k. Subsequently, the first D eigenvectors are selected, which means a fraction f of the total variance of the set of vectors ν can be explained by equation (3), where f equals:

$$f = \sum_{d=1}^{D} \lambda_d \ / \sum_{k=1}^{K} \lambda_k \qquad (4)$$

The value of D is chosen such, that f is satisfactorily close to 1.

3.4.3 *Further reduction of the number of stochastic variables*

Application of the PCA-analysis on the case study of the IJssel Lake showed that 20 eigenvectors are required to explain 90% of the total variance of the observed pre-event histories and 27 eigenvectors to explain 95%. Since the use of 20 or 27 stochastic variables is too much to handle in the numerical integration procedure, a modification of the described approach was developed.

Each pre-event history v is written as

$$\nu = \rho(\nu) \cdot \nu^* \qquad (5)$$

where $\rho(\nu)$ represents the "norm" or "magnitude" of vector ν and ν^* is a scaled or normalised version of ν, with $\rho(\nu^*) = 1$. Several alternatives have been considered for the definition of the norm $\rho(v)$, such as the L_1, the L_2 and the L_∞ norm, where L_n is defined as

$$\sqrt[n]{\sum_{d=1}^{D} |v_d|^n} \tag{6}$$

For each of the considered alternatives the norms of observed pre-event histories were compared to the corresponding water level peaks of the lake. As a result, the L_1 norm was selected since it is a measure of the total volume of water which has drained into the IJssel Lake. This volume is a significant control on the (maximum) water level.

Subsequently, the L_1-norm of the input processes (IJssel river discharge, sea level, etc.) are treated as stochastic variables. In order to do so, a Weibull fit was derived for observed L_1 values of the input processes of 46 high water events. Then, a statistical description of the temporal variation of the input processes was derived from a PCA-analysis to the set of normalised pre-event histories ν^*. The resulting principal components, or α-values, are also treated as stochastic variables, in this case with a Gaussian distribution. In this way the pre-event histories are described by a set of random variables of which one part governs the magnitude and the other part the temporal variation.

4 RESULTS

4.1 *Introduction*

The three methods described are applied on the series of observations of the period 1951–1996. The year 1951 was selected as start date because the available measurements of the relevant processes in previous years are insufficient. The resulting exceedance probabilities and corresponding levels in the IJssel Lake are presented and differences between the methods are explained.

4.2 *Method 1: extreme value distribution functions*

Two subsets of the available data of the period 1951-1996 were selected: the 46 annual maxima and the set of annual maxima (16 in total) which exceed a threshold of 0.12 m. +NAP. Gringortens method was used to obtain prior estimates of exceedance probabilities of observed water levels, which means $a = 0.44$ and $b = 0.12$ in Equation (1). The maximum likelihood method was used to obtain an optimal fit of the data for the Gumbel-distribution function.

Figure 2 shows the resulting fits for the two selected subsets. This figure clearly shows that the selection of the subset on which the data is fitted has a strong effect on the resulting exceedance probabilities.

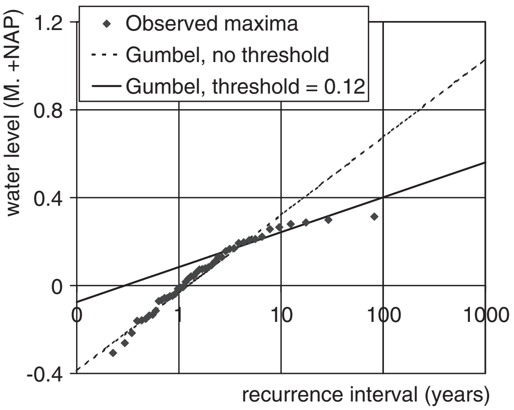

Figure 2. Gumbel fit of (1) 46 annual maxima and (2) the set of annual maxima above a threshold of 0.12 m. +NAP.

For a recurrence interval of 4,000 years (the desired safety level for dikes around the IJssel lake) the difference in derived water levels is 0.6 m. The relatively large differences are explained by the change in the pattern of the data that occurs somewhere around the water level of 0.10–0.20 m. +NAP. If the full set of annual maxima is used the resulting Gumbel function largely follows the relatively steep pattern of the lower observations, whereas for the subset of annual maxima the relatively flat pattern of the highest observed values is reflected. Consequently, the introduction of the threshold leads to a poor fit of water levels with low recurrence intervals whereas without this threshold the fit for high recurrence intervals is poor. If other extreme value distribution functions are used, the same effect occurs.

An interesting question is whether the observed change in pattern has a physical cause and, subsequently, whether the pattern of the highest observations is characteristic for the extreme recurrence intervals of interest (i.e. >1,000 years). If this is the case, the Gumbel function resulting from the set of highest observed water levels is to be preferred. However, the change in pattern may also be a pure coincidence. In that case the use of the larger set of data is to be preferred, since it holds more information. The results of the more physically based methods (described below) may provide some answers.

4.3 *Method 2: numerical integration*

4.3.1 *WINBOS results*

As explained in section 3.3 a relatively simplified and straightforward representation of stochastic variables is used in order keep the number of model runs manageable. Realisations of the stochastic variables are used as input of the WINBOS model to compute

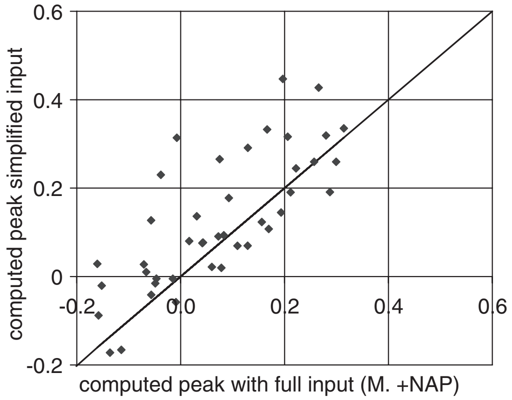

Figure 3. Derived model results, using [a] the full input information (x-axis) and [b] the simplified input representation (y-axis).

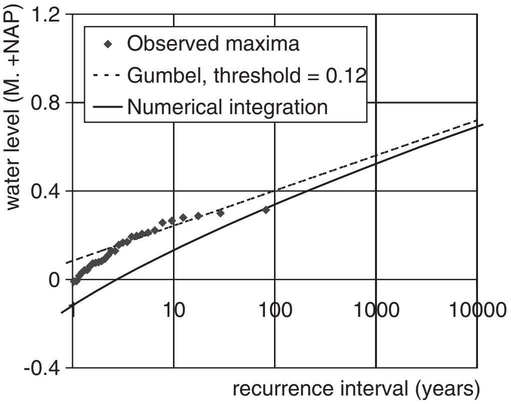

Figure 4. Results of the numerical integration (method 2) and the Gumbel fit of the set of annual maxima above a threshold of 0.12 m. +NAP (method 1).

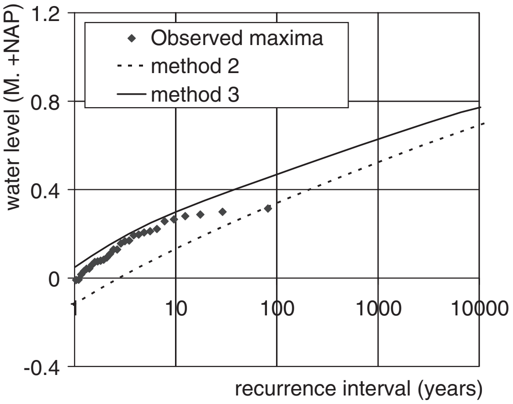

Figure 5. Resulting water levels and recurrence intervals of method 2 and method 3.

(peak) water levels of synthetic events. Due to the simplified representations of stochastic variables an error is introduced in the computed (peak) water levels. In order to quantify the error, model simulations of measured high water events have been executed with [a] measured time series of the relevant input variables and [b] the simplified representation of these input variables.

Figure 3 shows computed peak water levels for 46 events which have been simulated with both input series. The simplification of input variables causes on average an overestimation of peak water levels of 5 cm. The error appears to be independent to the magnitude of the event. Therefore, in the numerical integration procedure a correction of 5 cm. is applied to the computed water levels, to assure the model is not biased.

4.3.2 *Exceedance probabilities*

Figure 4 shows the results of the numerical integration (method 2) in combination with the results of Gumbel-fit of the set of annual maxima (16 in total) above a threshold of 0.12 m. +NAP (method 1). For low recurrence intervals (<50 years) differences between the derived water levels of the two methods are substantial. This also means the reproduction of observed water levels by the numerical integration method is rather poor, as can be seen from Figure 4.

For larger recurrence intervals (>100 years) differences between the two methods diminish. For a recurrence interval of 4,000 years the difference in derived water levels is about 3 cm. However this close correspondence should not be considered as a sound validation for both methods. Differences with observed water levels for low recurrence intervals (even for method 1, which is a fit procedure) provide a warning in this respect.

4.4 *Method 3: numerical integration based on principal component analysis (PCA)*

Figure 5 shows recurrence intervals and related water levels as derived with methods 2 and 3. The water levels derived for a given recurrence interval with method 3 are clearly higher than those derived with method 2. Differences are mainly due to the fact that method 3 accounts for the (statistical) dependency between the IJssel river discharge and the regional contributaries. As a consequence, the combined occurrence of extreme discharges on the IJssel river and the regional contributaries increases. Additional numerical integration runs show that for a recurrence interval of 100 years the total hydraulic load, i.e. the combined volume of the IJssel river and regional contributors increases with 230 million m^3.

This volume is the equivalent of an additional water level of approximately 20 cm. on the IJssel Lake.

For low recurrence intervals ($<$10 years) the derived water levels of method 3 are in close correspondence with observed water levels. For larger recurrence intervals this method appears to overestimate the water levels. However, in October 1998 a peak water level of 0.50 m. +NAP was observed (i.e. outside the period of observations as used in this studies). The recurrence interval of this event as derived with method 3 is 150 years which seems more likely than the recurrence intervals as derived with method 1 (500 years) and method 2 (750 years).

5 CONCLUSIONS AND DISCUSSION

Three different methods have been applied to derive exceedance probabilities of water levels of the IJssel Lake. Significant differences were found in the exceedance probabilities derived, using the three methods. Furthermore, the choices that have to be made within a single method (i.e. magnitudes of thresholds, number of stochastic variables) in some cases were also found to be of strong influence on the resulting exceedance probabilities.

The range of outcomes to some extent reflects the magnitude of the existing epistemic uncertainty (i.e. uncertainty resulting from a lack of knowledge of the system under extreme circumstances). Since both underestimation and overestimation of flood dangers can prove to be very costly, the range of possible outcomes preferably should be narrowed down as much as possible. In other words, some kind of validation procedure should be developed that enables the user to reject certain methods and/or ranges of outcomes.

A major problem in this respect is the length of the available series of observations, which is (much) too short to provide a sound validation of estimates of water levels with a recurrence interval of, for instance, 10^{-3} or 10^{-4} years. A verdict on the quality of predictions in any kind of (statistical) extreme value analysis is therefore partly based on good faith. A vital step in that respect is that the derived results should at least be physically realistic. Therefore, the second and third method presented here are believed to be the most promising, since they explicitly take into account the relevant physical features of the system (through application of the WINBOS model).

However, these two methods as currently applied still leave room for improvements, mainly due to the measures that were taken in order to reduce the required amount of model runs. For instance the number of classes in the second method and the number of principal components in the third method had to be reduced. Therefore, the first attempt to further improve the methods will consist of the application of alternative methods for the numerical integration procedure. An approach based on the principle of Monte Carlo simulations is considered, since it is expected to reduce the required amount of model simulation runs significantly. The amount of computation time that may be saved in this way can then be invested in a more refined representation of the processes involved.

ACKNOWLEDGEMENTS

This research was funded by the Dutch Ministry of Transport, Public Works and Water management, for which the authors are grateful. Textual comments by Micha Werner are appreciated.

REFERENCES

Klemeš, V. 1994. Statistics and probability: Wrong remedies for a confused hydrologic modeler. In: V. Barnett and K.F. Turkman, *Statistics for the environment 2: water related issues*: 345–370. John Wiley & Sons Ltd.: Chisester, England.

Kuchment, L.S., Demidov, V.N., Motovilov, Yu.G., Mazarov, N.A. and Smakhtin, V.Yu. 1993. Estimation of disastrous floods risk via physically based models of river runoff generation, In: Z.W Kundzewics, D. Rosbjerg and S.P. Simonovic (eds), *Extreme hydrological events: Precipitation, floods and droughts* (Proceedings of the Yokohama symposium, July 1993): 177–182. IAHS Publ. 213.

Haan, C.T. 1997. *Statistical methods in hydrology*. The IOWA State University Press, Ames IOWA.

HKV, 2001. *Uitbreiding afvoerstatistiek Borgharen, Lith, Lobith en Olst*. HKV report, December 2001 (in Dutch).

Pielou, E.C. 1984. *The interpretation of Ecological Data. A primer on Classification and Ordination*. Wiley & Sons, New York.

RIZA, 1999. *Meerpeilstatistiek*. RIZA report March 1999 (in Dutch).

RIKZ, 2000. *Richtingsafhankelijke extreme waarden voor HW-standen, golfhoogten en golfperioden*, RIKZ report December 2000 (in Dutch).

Shaw, E.M. 1988. *Hydrology in practice, second edition*. Chapman and Hall.

Vrijling, J.K., van Gelder, P.H.A.J.M., van Asperen, L., van de Paverd, M., Westphal, R., Berger, R. and Voortman, H.G. 1999. Uncertainty analysis of water levels on Lake IJssel in the Netherlands. *Proceedings ICASP8*1999*. Sydney, Australia, 12–15 December 1999.

Wet op de waterkeringen, 2002. Wet, houdende algemene regels ter verzekering van de beveiliging door waterkeringen tegen overstromingen door het buitenwater en regeling van enkele daarmee verband houdende aangelegenheden. *Staatsblad 304* (in Dutch).

WL | Delft Hydraulics 1997. *Onafhankelijk onderzoek Markermeer*. Report, WL | Delft Hydraulics December 1997 (in Dutch).