Pergamon

Wat. Sci. Tech. Vol. 36, No. 5, pp. 133–140, 1997.

0273–1223/97 $17·00 + 0·00

PII: S0273-1223(97)00498-8

EXTREME VALUE STATISTICS: POTENTIAL BENEFITS IN WATER QUALITY MANAGEMENT

O. Thas, P. Vanrolleghem, B. Kops, L. Van Vooren and J. P. Ottoy

Department of Applied Mathematics, Biometrics and Process Control (BIOMATH), University of Ghent, Coupure Links 653, 9000 Ghent, Belgium

ABSTRACT

Recently extreme value statistics have proven useful in environmental applications like the assessment of sea-levels, wind speeds and ozone concentrations. In this paper, after a brief overview of the statistical theory of extreme values, modelling issues are discussed with stress on applications in water quality management. Risk analysis procedures are presented that consider the extremal behaviour of water quality in the design stage of environmental constructions. © 1997 IAWQ. Published by Elsevier Science Ltd

KEYWORDS

Extreme Value Statistics, Risk Analysis, Water Quality Management

INTRODUCTION

In many environmental processes the average behaviour of the system is much less important than the rather few extreme situations. In fish water, for instance, the changes in pollutant concentration around the average are not determinant for life in the river but in case of one extreme event sudden death may be unavoidable. An example that will be used frequently throughout this text is the simulated time series of dissolved oxygen (DO, for short) concentration (Vanrolleghem et al., 1996). Extreme low values of less than 3 $mg.l^{-1}$ lasting for longer than 1 hour will lead to considerable killings among fish. The data are presented in Figure 1. The authors believe that this extremal behaviour is to be considered in the design phase of waste water treatment plants (WWTP, for short) or sewer systems, i.e. in many cases it might be more important to reduce the risk of sudden fish death than to improve the average water quality.

Classical statistical techniques are only sensitive in the area where the data have the greatest density, i.e. around the average. Therefore these techniques cannot be applied in situations where inference about the extremes is desired. The theory of extreme value statistics is based on its own limit theorem about the asymptotic distributions of sample maxima (or minima). Of course, by definition, the main problem with extreme values is their scarceness, as well as the fact that in many applications the researcher is actually interested in estimates beyond the largest observed value.

The study of extreme values was a rather theoretical field within probability at first. The classic reference on these first methods, which are performed on annual maxima of a time series, is (Gumbel, 1958). In recent years alternative methods are presented in which the scarce data are more efficiently used by selecting the r largest observations within each year (Gomes, 1981; Smith, 1986; Tawn, 1988) or by using all observations exceeding a specified large threshold (Davison and Smith, 1990). Both approaches can be unified by adopting the point process characterization of the extreme process (Pickands, 1971). The extremal distributions have been frequently used for statistical modelling with applications in environmental processes: sea-levels, wind-speeds and rainfall. In this paper we review the main concepts of extreme value statistics which are relevant for statistical modelling of constituent concentrations (e.g. pollutants) in rivers and its applicability for water

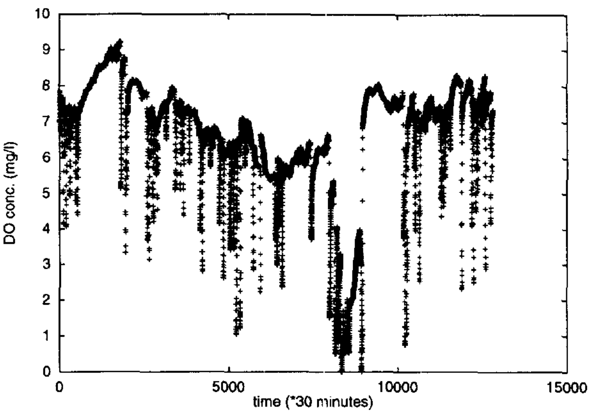

Figure 1: The simulated time series of the DO concentration ($mg.l^{-1}$) with one observation each 30 minutes.

quality management in general.

EXTREME VALUE STATISTICS

In this section a brief introduction to the theory of extreme value statistics will be given. At first some assumptions will be made. The most important is that all observations are supposed independently and identically distributed (IID, for short). In reality, however, environmental time series can show long and short range dependence. Later the theory will be generalized to the latter situation. The theory presented here is valid for univariate processes.

Extremal Distributions

Let $X_1, X_2, \ldots, X_n$ be a sequence of IID random variables with distribution function F. We will consider the distribution of the maximum order statistic $M_n = \max\{X_1, X_2, \ldots, X_n\}$. The general idea is to use a limit distribution of M_n for $n \to \infty$ as an approximation to the distribution of large but finite n. As with the central limit theorem the problem is degenerate, i.e. with probability 1 the distribution of M_n converges to the upper end-point of F, and the solution is to consider the limit distribution G after a linear rescaling. The distributions F and G are also called the parent and the extremal distribution, respectively.

Theorem: Extremal Type Theorem (Gumbel, 1958)
Suppose there exists a pair of sequences a_n and b_n with $a_n > 0$, such that, as $n \to \infty$

$$P\left\{\frac{M_n - b_n}{a_n} \leq x\right\} \to G(x)$$

for all x at which $G(x)$ is continuous and for some non-degenerate distribution G, then G is of the same type of one of the following distributions:

$$I : G(x) = \exp\{-\exp(-x)\} \quad -\infty < x < +\infty$$

$$II : G(x) = \begin{cases} 0 & x \leq 0 \\ \exp(-x^{-\alpha}) & x > 0, \alpha > 0 \end{cases}$$

$$III : G(x) = \begin{cases} \exp\{-(-x)^{\alpha}\} & x < 0, \alpha > 0 \\ 0 & x \geq 0 \end{cases}.$$

The 3 types of extremal distributions I, II and III are called Gumbel, Fréchet and Weibull, respectively. It should be noted that the theorem does not guarantee the existence of a limiting distribution G for each F,

but when it exists it must be of one of the 3 types whatever the distribution F. It is said that F lies in the domain of attraction of G.
The practical use of the limit theorem is by assuming that n is large enough for the limit to hold as an approximation. For fixed n we get for some $\mu = b_n$ and $\sigma = a_n > 0$

$$P\left\{\frac{M_n - \mu}{\sigma} \leq x\right\} \approx G(x).$$

Furthermore we have

$$P\{M_n \leq x\} \approx G\left(\frac{x-\mu}{\sigma}\right) = G^*(x)$$

which is of the same type of G. Therefore the extremal distribution can be used directly for the extreme observations of M_n. The constants μ and σ will now occur as the location and scale parameters, respectively, in the extremal distribution G. Their practical meaning and importance will be discussed in the section on modelling issues.

Since it is not straightforward to determine for a given parent distribution F to which type its extremal distribution G belongs, a more sensible approach for statistical purposes will be adopted: the generalized extreme value (GEV, for short) distribution. The GEV includes all 3 types of extremal distributions determined by the shape parameter ζ in the distribution function

$$G(x) = \exp\left\{-[1+\zeta(\frac{x-\mu}{\sigma})]^{-1/\zeta}\right\}$$

which is defined on $\{x : 1 + \zeta(x - \mu)/\sigma > 0\}$. The type II and III classes arise when $\zeta > 0$ and $\zeta < 0$, respectively, and the type I class is obtained as $\lim_{\zeta \to 0} G(x)$. For negative ζ (type III, Weibull) a finite endpoint is reached. Extremal distributions with $\zeta > 0$ (type II, Fréchet) are characterized by a very heavy tail, while the limit for $\zeta \to 0$ (type I, Gumbel) has an intermediate tail.

For statistical purposes the GEV can be used to specify the likelihood function for the data. In the simplest case the sample of M_n, which is necessary to estimate the parameters, is constructed as the sequence of annual maxima. This implies that the number of observations per year (n) must be large enough for the GEV to be a valid approximation, and in order to obtain accurate estimates the observations must be taken over many years, say p years. Then the sample can be represented as $M_{n,1}, M_{n,2}, \ldots, M_{n,p}$ which for simplicity will be denoted as $X_1, X_2, \ldots, X_p$. Furthermore, it is assumed that the distance in time between annual maxima is large enough for the maxima to be considered independently distributed.
Although other estimation techniques than maximum likelihood (ML, for short) are described in literature (graphical techniques and moment-based estimators), most recent applications are based on the ML method (Prescott and Walden, 1980; Prescott and Walden, 1983; Smith, 1985). An important advantage is its general framework for statistical modelling and inference which is well understood.
The log-likelihood is

$$l(\mu,\sigma,\zeta) = \sum_{i=1}^{p}\left\{-\log(\sigma) - (1+\frac{1}{\zeta})\log\left[1+\zeta\left(\frac{x_i-\mu}{\sigma}\right)\right] - \left[1+\zeta\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-1/\zeta}\right\}.$$

Standard errors can be obtained by the inversion of the observed information matrix. Due to the asymmetry in the likelihood function and since the validity of the interpretation of the observed information matrix depends on the quality of a quadratic approximation to the likelihood function at its maximum, confidence intervals can be calculated more accurately based on the profile likelihoods.

Alternative Formulation: Point Process Characterization

As mentioned before the application of the GEV desires a large amount of data of which only the annual maxima are used for inference. Thus a lot of information might be lost. Two alternative approaches were suggested and developed to partly overcome this limitation: (i) using the r largest observations within each year or (ii) using all observations exceeding a sufficiently large threshold u. Although both methods were derived independently (Gomes, 1981; Smith, 1986; Tawn, 1988; Davison and Smith, 1990), they can be seen as simple special cases of the point process characterization of the extreme value process (Pickands (Pickands,

1971) for a theoretical exposition, and (Smith, 1989) for the introduction into statistical applications). This will be presented first.

Let $X_1, X_2, \dots, X_n$ be an IID sample from distribution F, then $\{(i, X_i) : i = 1, \dots, n\}$ can be considered as a 2-dimensional point process. In this way the process can be specified in any region of the form $[t_1, t_2] \times (u, \infty)$, and more specifically, for high u, the extremal process can be characterized.
The asymptotic argument enters as follows. Suppose F lies in the domain of attraction of G, such that sequences $a_n > 0$ and b_n exist and a sequence of point processes on $\mathbb{R}^2$ can be defined as

$$P_n = \left\{ \left(\frac{i}{n+1}, \frac{X_i - b_n}{a_n} \right) : i = 1, \dots, n \right\}.$$

Then away from the lower boundary the process will look like a non-homogeneous Poisson process. Weak convergence of P_n to P is established on sets excluding the lower boundary. The intensity measure of the limiting Poisson process is $\Lambda(A) = E$(number of points in region A) where $E(.)$ denotes the mathematical expectation. According to the Poisson distribution for a region $A = \{(t_1, t_2) \times (x, \infty)\}$ for x large enough

$$\exp\{-\Lambda(A)\} = P\{\text{no points in region A}\} = P\{M_n \le x\} \approx \exp\left\{-[1 + \zeta x]^{-1/\zeta}\right\}.$$

As before, for fixed n, a linear rescaling is performed and in order to obtain the GEV as the distribution of the maximum over one year the constant n_y is introduced, representing the number of years of data to which the n observations correspond. In this way the intensity measure of the limiting Poisson process of $P_n = \{(\frac{i}{n+1}, X_i) : i = 1, \dots, n\}$ is

$$\Lambda\{(t_1, t_2) \times (x, \infty)\} = n_y(t_2 - t_1) \left[1 + \zeta \frac{x - \mu}{\sigma}\right]^{-1/\zeta}$$

in which the interpretation of the parameters μ, σ and ζ remains as before.

For the alternative threshold approach to find the extremal distribution G the point process method can be adapted by assuming that above a high threshold u, $P_n = \{(\frac{i}{n+1}, X_i) : i = 1, \dots, n\}$ approximates the non-homogeneous Poisson process with intensity measure Λ as before. Then within the region of the form $A_v = [0, 1] \times (v, \infty)$ for $v > u$, the likelihood is given by

$$L(A_v; \mu, \sigma, \zeta) = \exp\{-\Lambda(A_v)\} \prod_{i=1}^{N_A} d\Lambda(t_i, x_i) = \exp\left\{-n_y \left(1 + \zeta \frac{v - \mu}{\sigma}\right)^{-1/\zeta}\right\} \prod_{i=1}^{N_A} \sigma^{-1} \left(1 + \zeta \frac{x_i - \mu}{\sigma}\right)^{-\frac{1}{\zeta - 1}}$$

where $x_1, x_2, \dots, x_{N_A}$ are the N_A observed points which exceed the threshold v.
The generalized Pareto distribution, which was the original method proposed for threshold exceedance (Smith, 1984), may be derived very easily from this Poisson characterization by considering the conditional probability that $M_n > u + x$ given $M_n > u$.
This likelihood of the limiting Poisson process can be easily moderated for the r-largest order statistic approach by setting for each year the threshold u to the r-largest order statistic $M_n^{(r)}$ of that year, leading to

$$L = \prod_{j=1}^{p} \left\{ \exp\left[-\left(1 + \zeta \frac{M_{n,j}^{(r)} - \mu}{\sigma}\right)^{-1/\zeta} \right] \prod_{i=1}^{r} \sigma^{-1} \left(1 + \zeta \frac{x_{j,i} - \mu}{\sigma}\right)^{-\frac{1}{\zeta - 1}} \right\}$$

where $M_{n,j}^{(r)}$ is the r-th order statistic in year i and $x_{j,i}$ is the j-th order statistic observed in year i.

Serial Dependence

Until now independence among observations was assumed. In practice, however, time series are often characterized by serial correlation. The maxima tend to cluster. In the theory presented here (Leadbetter and Rootzén, 1988), we assume that there is no long range dependence but the rather technical formalized condition, which ensures the asymptotics of the extreme value distributions to hold, is not presented in this text. In order to quantify the effect of serial dependence the extremal index θ is defined as

$$\theta = \lim_{n \to \infty} P\{\max(X_2, \dots, X_{p_n}) \le u_n | X_1 \ge u_n\}$$

where $p_n = o(n)$ and the sequence u_n is such that $P\{M_n \leq u_n\}$ converges. The correct interpretation is that θ is the reciprocal of the mean cluster size at asymptotically extreme levels, but more loosely it is just the reciprocal of the mean cluster size.

An important theorem concerning the extremal index is based on the existance of a limiting GEV distribution of maxima of a sequence of independent variables with the same marginal properties as the original series. The theorem states that the asymptotic distribution of the maxima of a serial correlated sequence is still a GEV distribution and that the extremal index θ only affects the location and scale parameters but not the shape parameter. Hence, the type of the limiting distribution is unaffected.

For practical purposes the theorem implies that the method of annual maxima can be applied without modifications. The only effect is on the location and scale parameter which have to be estimated anyway. For procedures derived from the point process characterization, on the other hand, independence was assumed for the construction of the likelihoods, i.e. the likelihoods assume more information than there actually is in the sample of correlated data. Some modifications to tackle this problem are proposed (Coles et al., 1994).

1. The point process techniques can be applied on the cluster maxima, which are considered independent (Tawn, 1988). The detection of clusters is however arbitrary and a lot of informative data is discarded.
2. An extension to the first approach. The extremal index can be estimated as the reciprocal of the mean cluster size.
3. The likelihoods for independent data can be used, after which the standard errors are corrected for dependence.
4. The serial dependence can be modelled explicitly (e.g. Markov chain) such that the likelihood properly accounts for the dependence. The extremal index comes into the likelihood as a parameter.

The first two approaches are most frequently adopted in applications, although they do not use all informative data for estimation. The last option is model-based and its statistical behaviour is still under study.

MODELLING ISSUES

In this section some possible applications of extreme value statistics to water quality management are given. The main fields in which the authors feel that a detailed knowledge of the extremal behaviour might give an added value are (i) risk analysis (relation between probability and extreme quantiles), (ii) design and operation of WWTP and sewer systems (relation risk analysis and design parameter) and (iii) intervention analysis (statistical assessment of the effect of a sudden and known change (intervention) in the properties of the process under study). All these applications need a thorough modelling of the parameters of the GEV distributions and a good understanding of statistical prediction/estimation and inference procedures. Therefore, it is prefered to start the exposition with the modelling or parameterization of the GEV distribution and to consider the applications as examples such that the statistical techniques can be explained in a natural order.

First, however, a general remark concerning the annual maximum approach must be made clear. In the literature this method is nearly always performed on annual maxima, although from the theoretical development it is only necessary to have a sample of independent maxima which are obtained from sequences which are large enough for the asymptotic argument to hold approximately. In many examples the time series consists of daily observations, i.e. $n = 365$. Hence, in case the observations are taken more frequently (e.g. hourly), the maxima may be obtained from sequences of 1 or 2 weeks as long as the assumption of independent maxima still holds. This situation can be more realistic when data are simulated for a WWTP or a sewer system for which the extreme events might last only a very short time. This can be seen in Figure 1 for the DO example.

In the above it was always assumed that the parameters μ, σ and ζ are constant over the whole sequence of observations. In reality, however, this condition is frequently not fulfilled. Examples that can easily be seen and that are typical for constituent concentrations in rivers are seasonality and trend in the extremal behaviour which are both quantified by the location parameter μ. Note that the parameters μ and σ are not to be confused with the mean and variance of the parent distribution F. They only characterize the location and the scale of the extremal distribution and must be seen independent from F. It is for example possible to have a time series with a more or less constant mean of the parent distribution but with an increasing location parameter for G, indicating that only the extremes in the sequence become larger as time increases. Obviously, this cannot be detected rigorously with standard statistical techniques. Also the scale parameter σ can change in time independently of the variance of F. An increasing σ in time is seen by an increasing variability in the extreme values.

Models for the Parameters of the GEV Distribution

The simplest form of trend is a linear trend of μ in time. It can be parameterized by replacing μ in the GEV distribution (and hence in the likelihoods) by $\mu(t) = \alpha + \beta t$ where α and β become now the parameters to be estimated and the subject for inference from extremal data. It can be used for example to assess the effect of a slow change in water quality management policy on pollutant concentrations in rivers. A simple likelihood ratio test can be performed for testing $H_0 : \beta = 0$, i.e. testing whether there is a linear change in the extremal location parameter.

In case a new WWTP is installed on day t_0 and one is interested in the effect of its capacity on the extremal behaviour of a pollutant concentration further down the stream, the location parameter can be modelled as $\mu(t) = \alpha + \beta I$ where I is an indicator variable defined as

$$I = \begin{cases} 0 & t < t_0 \\ 1 & t \geq t_0 \end{cases} .$$

Again a likelihood ratio test can be performed for testing $H_0 : \beta = 0$, i.e. there is no effect of the WWTP on the extremal behaviour. This type of modelling is called intervention analysis.
The parameter α is interpreted as the location parameter of the extremal distribution if the WWTP would not have been installed. β is the shift in the location parameter caused by the WWTP.

Seasonality can be modelled in several ways. The most efficient approach is to model the periodicity as a superposition of sinus or cosinus functions with known frequency (harmonic frequencies can also be included). In this periodic function approach the location parameter is modelled as $\mu(t) = \alpha + \beta_0 \cos(2\pi t/365 - \beta_1)$ for a daily measured pollutant concentration with an annual cycle and no harmonics included.
Another technique is the block method. Here the parameter μ is modelled as $\mu(t) = \mu_j$ with $j = 1, \dots, 12$ representing the month at time t. It is equivalent to assume that within each month the process is stationary. Since time series of water quality variables most often have known periodicity (e.g. daily, weekly, monthly, yearly) the periodic function approach is the most sensible and economical (i.e. less parameters to be estimated) method; it essentially makes use of *a priori* knowledge.

Of course, combinations of all these parameterizations or more complex models of μ are also possible. All these models are possible for the 3 types of sampling (annual maxima, r-largest and threshold method). In case of the threshold method it might be advantageous to adopt a time-variable threshold that accounts for the temporal relative extremity of the data. This is strongly related to the general problem of the choice of threshold and the choice of r for the r-largest method. This problem is discussed in some more detail below.

Non-stationarity is not only defined on the location parameter but also on the other parameters specifying the GEV distribution (σ, ζ and θ in case the serial dependence is modelled explicitly). In contrast to the location parameter is the time-dependency of the other parameters often not visual observable and is it not advisable to model them as a continuous function of time (e.g. a linear relationship). More feasible and meaningful is to model them with an indicator variable for assessing the effect of a sudden policy change (e.g. start of a new WWTP or closing of an industry). In this case it makes indeed sense to look at the effect on the extremal behaviour quantified not only by μ and σ but also by the shape parameter ζ which gives some clear measure for the shape of the tail of the distribution. In case the model-based interdependence approach is used, it is even possible to model the extremal index θ such that the effect on the (asymptotic) mean cluster size can be assessed. In the DO example this parameter is a very important quantity and its meaning is obvious when the threshold method is applied with the threshold set to the critical DO level ($3\ mg.l^{-1}$) in which case the mean cluster size is the mean duration of the exceedance which is one of the

determinants for the fish death.
For seasonality the block method can be applied to all parameters (μ, σ, ζ) simultaneously.
In some exceptional cases it might be informative to model the shape parameter as a linear function of the time: $\zeta(t) = \alpha + \beta t$. When it turns out that the parameter β is estimated very accurately and the model diagnostics do not indicate any substantial misformulation of the model, the conclusion can indicate that the relative form of the tail of the distribution is changing. Of course, this is a very important conclusion with respect to the overall environmental change that cannot be recognized by any other technique.

For the choice of the threshold u or r for the r-largest method a robust compromise must be found between the amount of data that will be retained for the analysis and the extremity of these retained observations necessary for the approximation to the asymptotics to hold. A diagnostic check can be performed by fitting the GEV distribution for a range of u or r values and by plotting the estimates of the parameters (μ, σ, ζ) against u or r. Robust choices are located in a range where there is estimate stability. The rationale is of course that the asymptotic argument suggests that the GEV distribution with parameters (μ, σ, ζ) is valid for all observations above a certain large threshold (obtained by choosing u large or r small).

Risk Analysis and Environmental Design

All applications given above rely on estimation and statistical testing. In many classical applications of extreme value theory extrapolation or prediction is very important. For example one wants to know the probability or the frequency that a very extreme sea level might occur based on observations which do not include such an extreme level. In the example this kind of information is important for the design of constructions that must resist these extreme situations (e.g. dikes). This kind of problems essentially belongs to risk analysis. Since for water quality the critical levels often are observed, the general weakness of the extrapolation principle does not apply. Therefore the use of extreme value statistics in risk analysis is certainly a useful and reliable tool. Not only can a model based on the GEV distribution be fitted to the data and can this model and its estimated parameters be used for summarizing the extremal behaviour, but also can this information be used to estimate the extreme quantiles x_p at any probability $p = 1 - G(x_p) = P\{M_n > x_p\}$. Based on the GEV distribution this relation is analytically $x_p = \mu - \frac{\sigma}{\zeta}\left\{1 - [-\log(1-p)]^{-\zeta}\right\}$. Instead of plotting (p, x_p) a return plot is often constructed, i.e. plotting x_p against $1/p$, which are called the return level and the return period, respectively.

Closely related to risk analysis is the environmental design. E.g. the height of a dike can be specified such that the risk for accidents is controlled to some prespecified probability (Coles and Tawn, 1994). As before, in this example there is some criticism since the risk calculations are based on extrapolations in the tail of the distribution. This will not be the case in most situations concerning water quality.

The general methodology of environmental design can therefore be formalized.
Let δ be the design parameter that has to be optimized with respect to a certain risk quantity $\pi = P\{M_n \leq u_c\}$ where u_c is a user defined critical level which can be related to the design parameter (e.g. the height of the dike can also be the critical level) or just some norm specified by the authorities (e.g. a critical pollutant concentration), and suppose Δ is the set of plausible values for δ. Of course $\pi \in [0,1]$. The mapping ψ is defined as $\psi : \Delta \to [0,1] : \delta \mapsto \pi$. The problem of the environmental design is to find the solution set Ω defined as $\Omega = \{\delta : \psi(\delta) < \pi_c\}$ where π_c is the maximal acceptable risk.

Nice examples of this methodology are the optimization of design parameters for WWTP or sewer systems. First the set Δ of plausible values for the design parameter δ must be defined. Since the extremal behaviour of the process probably is not the only and also not the most important criterion, other design methods may be applied to reduce the complete set of all possible values for δ to Δ. The mapping ψ can be calculated implicitly by simulation techniques. Real input data for the WWTP or the sewer can be used to simulate the system under study. On the output data a well parameterized GEV distribution must be fitted that is subsequently used to estimate π.

CONCLUSIONS

Apart from the strong theoretical developments in extreme value theory, the GEV distribution is shown useful in applied statistics as an elegant modelling tool for extreme events through appropriate modelling of the location, scale and shape parameters and occasionally the extremal index. Each parameter has its own specific relevancy for water quality extremal behaviour. The combined information in the threshold exceedance and mean cluster size is proposed very important for the characterization of water quality w.r.t. the viability conditions of the water. Applications to several modelling approaches were given in the area of water quality management. These include descriptive aims, intervention analysis and risk analysis. Closely related to risk analysis is environmental design of which the main ideas were explained and its importance in the design of WWTP and sewer systems was stressed, as well was its general framework for a practical methodology discussed. In general extreme value statistics can give very important information about the dynamics of the water quality in the tail of the distribution where most other methods are not sensitive in this area. An interdisciplinary approach, like the one proposed for the environmental design, must be encouraged.

ACKNOWLEDGEMENTS

Financial support for the work was partly granted by the Belgian National Fund for Scientific Research (N.F.W.O.) through a research grant. The authors are also grateful to the VMM (Flemish Environmental Agency). The first author would also like to thank Dr Stuart Coles for the interesting discussion.

REFERENCES

Coles, S., Smith, R., and Tawn, J. (1994). A seasonal markov model for extremely low temperatures. *Environmetrics*, 5:221–239.

Coles, S. and Tawn, J. (1994). Statistical methods for multivariate extremes: an application to structural design (with discussion). *Applied Statistics*, 43:1–48.

Davison, A. and Smith, R. (1990). Models for exceedances over high thresholds (with discussion). *Journal of the Royal Statistical Society, Series B*, 52:393–442.

Gomes, M. (1981). An i-dimensional limiting distribution function of largest values and its relevance to the statistical theory of extremes. In Taillie, C., Patil, G., and Baldessari, B., editors, *Statistical Distributions in Scientific Work*, volume 6, pages 389–410. Reidel, Dordrecht.

Gumbel, E. (1958). *Statistics of Extremes*. Columbia University Press, New York.

Leadbetter, M. and Rootzén, H. (1988). Extremal theory for stochastic processes. *Annals of Probability*, 16:431–478.

Pickands, J. (1971). The two-dimensional poisson process and extremal process. *Journal of Applied Probability*, 8:745–756.

Prescott, P. and Walden, A. (1980). Maximum likelihood estimations of the parameters of the generalized extreme value distribution. *Biometrika*, 67(723-724).

Prescott, P. and Walden, A. (1983). Maximum likelihood estimation of the three-parameter generalized extreme value distribution from censored samples. *Journal of Statistical Computing and Simulation*, 16:241–250.

Smith, R. (1984). Threshold methods for sample extremes. In de Oliveira, J. T., editor, *Statistical Extremes and Applications*, pages 621–635. Reidel, Dordrecht.

Smith, R. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, 72:67–92.

Smith, R. (1986). Extreme value theory based on the r largest annual events. *Journal of Hydrology*, 86:27–43.

Smith, R. (1989). Extreme value analysis of environmental time series: an application to trend detection in ground-ozone. *Statistical Science*, 4(4):367–393.

Tawn, J. (1988). An extreme value theory for dependent observations. *Journal of Hydrology*, 101:227–250.

Vanrolleghem, P., Fronteau, C., and Bauwens, W. (1996). Evaluation of design and operation of the sewage transport and treatment system by an eqo/eqs based analysis of the receiving water immission characteristics. In *Proceedings WEF Speciality Conference on Urban Wet Weather Pollution: Controlling sewer overflows and stormwater runoff*, pages 14.35–14.46, Quebec (CA).