AN ADVANCED STATISTICAL EXTREME VALUE MODEL FOR EVALUATING STORM SURGE HEIGHTS CONSIDERING SYSTEMATIC RECORDS AND SEA LEVEL RISE SCENARIOS

Christoph Mudersbach[1] and Jürgen Jensen[2]

In this paper, a non-stationary extreme value analysis approach is introduced in order to determine coastal design water levels for future time horizons. The non-stationary statistical approach is based on the Generalized Extreme Value (GEV) distribution and a L-Moment parameter estimation as well as a Maximum-Likelihood-estimation. An additional approach considers sea level rise scenarios in the non-stationary extreme value analysis. All the methods are applied to the annual maximum water levels from 1849-2007 at the German North Sea gauge at Cuxhaven. The results show, that the non-stationary GEV approach is suitable for determining coastal design water levels.

Keywords: extreme value statistics; non-stationary time series; design levels; sea level rise scenarios

INTRODUCTION

The design of coastal defence structures in Germany is based either on deterministic or statistical approaches (e.g. MLR 2001). For the statistical methods water levels with assigned return periods are needed. Considering an integrated coastal risk management approach, there is a growing demand for calculating exceedance probabilities of water levels, where the extremes are of particular importance. The calculation of exceedance probabilities or return periods is widely based on an extreme value analysis of observed sea level data.

The theory of extreme value analysis of hydrological data (e.g. sea level data) has been studied very intensively in the past. (e.g. Gumbel 1958; Jensen 1985; Coles and Tawn 1990; Kotz and Nadarajah 2000; Katz et al. 2002; Hawkes et al. 2008). The extreme value analyses can be based both on block maxima (e.g. annual maxima) and peaks over threshold. In recent years, the Generalized Extreme Value distribution (GEV) has become the most widely used method for block maxima. In order to use peak over threshold data the Generalized Pareto distribution (GPD) is recommended (e.g. Hawkes et al. 2008).

In the present paper, an approach is used based on annual maximum sea level data of the North Sea gauge at Cuxhaven, where high quality data records are available since 1849 (Jensen and Mudersbach 2007). The basic assumption of an extreme value approach is that the data are independent and identically distributed (iid, e.g. Coles 2001). While the demand for independency by using annual maximum sea level data may be justified due to the temporal resolution, the assumption of stationarity is unlikely in most cases due to existing trends in the mean and/or variability of the data (Khaliq et al. 2006). Before starting an extreme value analysis the time series needs to be tested against trends or changes/shifts in the mean and the variability in order to assess the stationarity or non-stationarity of the data (Hawkes et al. 2008). If there are significant non-stationarities in the data, these effects are commonly removed to obtain a stationary time series (Salas 1993). This approach is practical, but contains a decisive weakness: the results of the extreme value analysis are only valid for the present state. From a coastal engineer's point of view, not only is the present state is important, but also extreme value results for future time horizons. This is due to the design of coastal structures with a certain lifetime, e.g. 50-100 years. It has to be ensured that the design water level is valid up to the projected end of the lifetime. Assuming a coastal dike has to be designed with a 100-yr return level and has a projected lifetime up to 2100, then the 100-year return level for the year 2100 needs to be estimated (Fig. 1). For the determination of future design levels, the expected development of the water levels up to a future time horizon has to be considered. In many cases this may be achieved by adding an expected sea level rise (SLR) of, for example, 0.3 cm/yr. The latter does not provide the calculation of valid return periods for future design levels. Thus, an advanced statistical model is needed which eliminates this weakness.

[1] Research Center for Water and Environment, University of Siegen, Paul-Bonatz-Str. 9-11, 57076 Siegen, Germany, E-Mail: christoph.mudersbach@uni-siegen.de

[2] Research Center for Water and Environment, University of Siegen, Paul-Bonatz-Str. 9-11, 57076 Siegen, Germany, E-Mail: juergen.jensen@uni-siegen.de
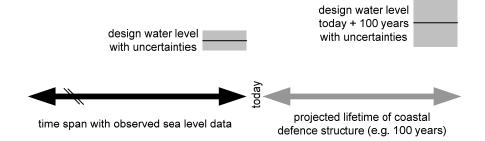
Figure 1. Estimation of future design levels for coastal defence structures.

However, in the light of climate change, it is likely that the assumption of stationarity in hydrologic time series becomes doubtful and, therefore, advanced methods in extreme value analysis have to be developed and applied (Khaliq et al. 2006). In recent years, the concept of non-stationary extreme value analysis has been improved and is nowadays used more often (e.g. Coles 2001; Strupczewski et al. 2001; Katz et al. 2002; Cunderlik and Burn 2003; Khaliq et al. 2006; Butler et al. 2007; Nogaj et al. 2007; Mendez et al. 2007; El Adlouni et al. 2007; Riberau et al. 2008; Hundecha et al. 2008). In the non-stationary approach, the parameters of the distribution functions are replaced by time-dependent parameters, so that the results of the extreme value analysis also vary with time. The benefit of this approach is that the original data no longer have to be detrended and can be used directly. The non-stationary extreme value analysis enables also the extrapolation of the results up to future time horizons.

The causes of a non-stationary behaviour of coastal water level data have been studied intensively (e.g. Coles 2001, Pugh 2004, Mendez et al. 2007). One of the main contributions to extreme sea level data is the mean sea level rise (e.g. Pugh 2004). Furthermore, not only sea level rise, but also changes in the intensity and frequency need to be considered. For example, Grossmann et al. (2006) found that, on basis of the IPCC scenario A2 for the time horizon 2085, an increase of the annual maximum water level at the Cuxhaven gauge of approximately 20 cm may be expected without any consideration of the mean sea level rise. By adding an expected sea level rise of 30 cm to the results, the total increase amounts to 50 cm.

Time-dependent models of the Generalized Extreme Value distribution (GEV) for determining return periods and applications to hydro-meteorological data have been studied recently. Coles (2001) investigated annual maximum sea level data at Fremantle, where only the location parameter was set as time-dependent with a linear model. In addition, the location parameter was linearly related to the Southern Oscillation Index (SOI). Katz et al. (2002) introduced a non-stationary GEV model, and suggested a linear model for the location and a log-transformed model for the scale parameter, whereas the shape parameter was kept constant. Mendez et al. (2007) published a wide application of a time-dependent GEV model to monthly extreme sea levels. They used nonlinear time-dependent models for all three parameters containing seasonal and long-term effects. Hundecha et al. (2008) analysed changes in extreme annual wind speeds in Canada with a non-stationary GEV, where location and scale parameters were time-dependent. Models have not yet been developed which would employ a non-stationary GEV distribution specifically to estimate future design levels on the German North Sea coastline.

The aim of this paper is to give a brief introduction to non-stationary extreme value analysis methods. A pragmatic GEV approach is described in order to estimate future design water levels for coastal engineering tasks. The approach described provides an advanced method to extrapolate return levels up to future time horizons. With this, an assessment of common design levels in comparison to time-dependent design levels is feasible. The methods are applied to the annual maximum water level data at the German North Sea gauge Cuxhaven.

THEORY OF NON-STATIONARY EXTREME VALUE ANALYSIS

In the present paper the non-stationary form of the GEV is used, which is well described by Coles (2001) and Cunderlik and Burn (2003) and is generally given in the form:

$$GEV(x,t) = \exp\left[-\left(1 + k(t) \cdot \frac{x - a(t)}{b(t)}\right)^{-\frac{1}{k(t)}}\right]$$

where x is the independent value (e.g. water level), a(t) the time-dependent location parameter, b(t) the time-dependent scale parameter, and k(t) the time-dependent shape parameter.

In general, non-parametric and parametric methods can be used to describe the time-dependent behaviour of the GEV parameters. In order to obtain return levels for future time horizons, the non-parametric approach is not applicable, due to the fact that it can not be extrapolated (Strupczewski et al. 2001). Therefore, a parametric approach is needed. For the description of the time-dependent behaviour of the parameters, several functions can be used. It is better to use rather simple relations, e.g. linear or exponential functions, in order to avoid difficulties in extrapolating. Furthermore, it is also possible to consider physically based covariates in the functions (Coles 2001). Mudersbach and Jensen (2010) used two different time models for the location and scale parameters are used. For the location and scale parameters, linear and exponential models may be used for the sake of simplicity in extrapolation. Both the linear and exponential models are physically based. Many sea level data sets change linearly over the observation period, whereby it seems plausible that this trend will continue to exist in the near future. The exponential model is also physically based since all climate scenarios of the IPCC-Report 2007 state a slight change in sea level rise by the middle of the century, while the changes by the end of the century are expected to be greater (e.g. Meehl et al. 2007). In this paper, we use a linear model for the first (location) and second (scale) parameters (α and β) respectively (GEV11). Thus, the following model result for the location parameter $\alpha(t)$:

$$a(t) = \alpha_0 + \alpha_1 t \tag{1}$$

Accordingly, the time model for the scale parameter $\beta(t)$ is given as:

$$b(t) = \beta_0 + \beta_1 t \tag{2}$$

As mentioned above, the shape parameter k is kept constant, so that

$$k(t) = k \tag{3}$$

In these time models the parameters α_0, α_1, β_0, β_1 and k need to be estimated from the observed data. For the calculation of the parameters different methods are generally available, which are initially known from the stationary extreme value analysis approach. The most widely used estimation methods are the method of moments (MM, e.g. Rao and Hamed 2000), the method of probability weighted moments (PWM, Greenwood et al. 1979), the method of L-moments (LM, Hosking and Wallis 1997), and the maximum-likelihood estimation (MLE, e.g. Coles 2001). Over the past few years, especially the parameter estimation methods PWM, L-moments and MLE have been used and developed further. El Adlouni et al. (2007) developed a so-called generalized maximum-likelihood estimation for the time-dependent parameters of the GEV. Ribereau et al. (2008) published a new approach based on PWM, the so-called generalized probability weighted moment approach. Cunderlik and Burn (2003) calculated the parameters of the GEV by using time windows, where for each time window the parameters are estimated using L-moments (see also Mudersbach and Jensen 2010).

DATA

The hydrodynamic system of the German Bight is dominated by semi-diurnal tides with a tidal range of about 3 to 3.5 m. Storm surges at the German North Sea coast result mainly from a build-up of water masses along the coasts, which are superimposed on the astronomical tide. The most distinctive flooding events have been the storm surge on 16/17 February 1962, with over 300 fatalities in Hamburg and the extreme flood on 3 January 1976, where the highest water level occurred up to now (Müller-Navarra et al. 2006).

The gauge at Cuxhaven is located at latitude 53° 52' N and longitude 08° 43' E in the central German Bight (Southern North Sea) at the mouth of the river Elbe estuary (Fig. 2). It is one of the most important gauges at the German North Sea coastline due to the long term records of water levels. The Cuxhaven gauge has an important role for designing flood defence structures both on the German coastline and in Hamburg.

The water level data used in this study are the annual maximum water levels (HThw) from 1849 to 2007 (Fig. 3). Alongside the HThw data, the time series of annual mean tidal high water (MThw) is also shown in Fig. 3, which is important for normalising purposes. The HThw data show an increasing linear trend s, which amounts to $s_{HThw} = 0.40$ cm/yr. The MThw data show also an increasing trend with $s_{MThw} = 0.26$ cm/yr. The water level data used in this analysis have been converted to the German reference datum NN (normal zero).
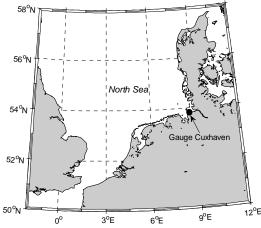

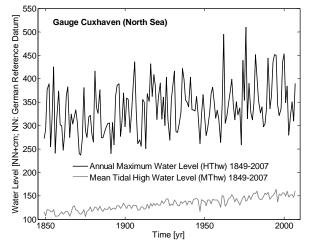**Figure 2. Map of the North Sea and location of the gauge Cuxhaven.**

Figure 3. Time series of annual maximum and mean tidal high water level at gauge Cuxhaven from 1849-2007.

RESULTS

The dataset of the tide gauge Cuxhaven was analysed by use of a non-stationary extreme value approach. In this paper, two parameter estimations methods were conducted and compared. The two models are one approach based on L-Moments with sliding windows (GEV11-LM) and one based on Maximum-Likelihood-estimation (GEV11-MLE). The results of the adopted models can be modified to a standardized version, thus, common model diagnostic tools can be used (Coles 2001). Therefore, the goodness of the estimation method is assessed by probability-probability plots (PP-plots), where the empirical probability is plotted against the theoretical probability. Figure 4 shows PP-plots for the parameter estimation method based on L-Moments (sliding window approach) and for a Maximum-Likelihood approach, respectively. It can be stated, that both models lead to satisfying results, i.e., both models are appropriate to calculate the non-stationary parameters of the GEV11-model.
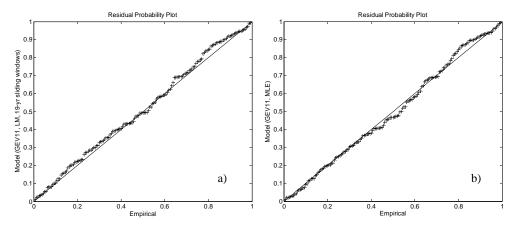
Figure 4. a) PP-plot of L-Moment approach. b) PP-plot of Maximum-Likelihood approach.

In Figure 5 some results of the non-stationary GEV11-models can be seen. The solid lines represent the 99%-, 98%-, and 95%-quantiles, respectively. The parameter estimation is based on the observed dataset of annual maximum water levels at the gauge Cuxhaven from 1849 to 2007. In Figure 5 additionally an extrapolation of the calculated trends was performed by the year 2100. As expected, the results from both parameter estimation methods differ. E.g., the 99%-quantile in 2007 (GEV11-LM) is about NN+508 cm and will be NN+540 cm by 2100, which is an increase of about 32 cm (Figure 5a). Considering the GEV11-MLE model, the same value in 2007 is NN+515 cm and will be NN+557 cm by 2100, which is an increase of about 42 cm (Figure 5b). However, these uncertainties are common in statistical extreme value analyses an do not lead to a rejection of one of the methods.
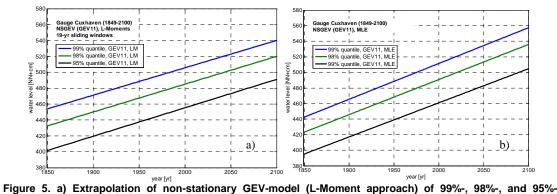

Figure 5. a) Extrapolation of non-stationary GEV-model (L-Moment approach) of 99%-, 98%-, and 95%-quantiles by 2100. b) Extrapolation of non-stationary GEV-model (Maximum-Likelihood approach) of 99%-, 98%-, and 95%-quantiles by 2100.

Next to the possibility to extrapolate quantiles by use of the trends based on the observed data, sea level rise scenarios can be applied to calculate quantiles for future time horizons. Doing so, it is supposed, that the available sea level rise scenarios represent the development of the trend of the location parameter. As common sea level rise scenarios give estimates for the development of the mean sea level and not for higher quantiles, the replacement of the trend of the location parameter by the trend of a sea level rise scenario is not fully justified. However, the results will give a possible range of future changes in the quantiles. If available, more appropriate scenarios (e.g. scenarios for 99%-quantiles) should be used.

In recent years numerous sea level rise scenarios were published. In Table 1 some well known scenarios are listed. An assessment or discussion of sea level rise scenarios is not part of this paper. The given scenarios will rather used to incorporate them into the non-stationary extreme value approach and to compare these results with those mentioned above. For the purpose of simplification

the sea level rise scenarios listed in Table 1 are combined to a mean sea level rise scenario, which is from 0.61 to 1.47 m by 2100.

Table 1. Sea Level Rise Scenarios (in Jevrejeva et al. 2010)	
Author	**sea level rise by 2100**
Meehl et al. (2007)	0.18 – 0.59 m
Rahmstorf (2007)	0.50 – 1.40 m
Pfeffer et al. (2008)	0.80 – 2.00 m
Vermeer and Rahmstorf (2010)	0.75 – 1.90 m
Grinsted et al. (2010)	0.80 – 1.30 m
Jevrejeva et al. (2010)	0.60 – 1.60 m
Mean	**0.61 – 1.47 m**

Figure 6 shows the results by incorporate the mean sea level rise scenario for the gauge Cuxhaven (GEV11-LM). The solid lines represent the extrapolation of the quantiles by 2100 based on the trends calculated by use of the observed data (same as in Fig. 5a). The dashed lines show the development of the different quantiles considering the upper and lower boundary of the mean sea level rise scenario, respectively. By incorporating the sea level rise scenarios only the location parameter of extreme value distribution is replaced. The scale and shape parameters are still the same as not using scenarios. The results show, e.g. that using the upper boundary of the mean sea level rise scenario (1.47 m) the 99%-quantile in 2100 will increase of about 115 cm in comparison to using not the sea level rise scenario. Looking at the same value, but using the lower boundary of the mean sea level rise scenario (0.61 m) there will be an increase of about 29 cm.
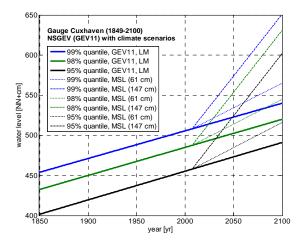

Figure 6. Extrapolation of non-stationary GEV-model (L-Moment approach) of 99%-, 98%-, and 95%-quantiles by 2100 considering different sea level rise scenarios.

CONCLUSIONS

Coastal engineers often are faced to non-stationary water level time series and are requested to determine design values for costal protection purposes. In principle, non-stationary extreme value analysis methods provide appropriate tools to achieve these tasks. There are some different methods available, where the most important differences exist in the time-models for the parameters of the extreme value distribution. Here, linear as well as higher order functions can be used to describe the future development of the parameters. The choice of the model should be based on physical findings. Since design values for future time horizons will be needed, sea level rise scenarios can be incorporated into the non-stationary extreme value analysis. Even if sea level rise scenarios have significant uncertainties, the use will give a possible range of future changes in the quantiles of extreme water levels. It should be mentioned, that all analyses should be evaluated critically, since the chosen time model and the range of extrapolation should lead to physical reasonable results.

REFERENCES

Butler, A., Heffernan, J. E., Tawn, J. A., and Flather, R. A. (2007) Trend estimation in extremes of synthetic North Sea surges. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 56 (4), 395-414.

Coles, S. (2001) An Introduction to Statistical Modeling of Extreme Values. Springer, London.

Coles, S.G. and Tawn, J. (1990) Statistics of coastal flood prevention. *Phil. Trans. R. Soc. Lond. A*, 332, 457-476.

Cunderlik, J.M. and Burn, D.H. (2003) Non-stationary pooled flood frequency analysis. *Journal of Hydrology*, 276, 210-223.

El Adlouni, S., Ouarda, T.B.M.J., Zhang, X., Roy, R., and Bobée, B. (2007) Gerneralized maximum likelihood estimators for the nonstationary generalized extreme value model. *Water Resources Research*, 43, W03410.

Grossmann, I., Woth, K. and von Storch, H. (2007) Localization of global climate change: Storm surge scenarios for Hamburg in 2030 and 2085. *Die Küste*, 71.

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

Hawkes, P.J., Gonzales-Marco, D., Sanchez-Arcilla, A., and Prinos, P. (2008) Best practice for the estimation of extremes: A review. *Journal of Hydraulic Research*, 46, Extra Issue 2, 324-332.

Hosking, J.R.M. and Wallis, J.R. (1997) Regional frequency analysis. Cambridge University Press, Cambridge.

Hundecha, Y., St-Hilaire, A., Ouarda, T.B.M.J., El Adlouni, S., and Gachon, P. (2008) A nonstationary Extreme Value Analysis for the Assessment of Changes in Extreme Annual Wind Speeds over the Gulf of St. Lawrence, Canada. *Journal of Applied Meteorology and Climatology*, 47, 2745-2759.

Jensen, J. (1985) Über instationäre Entwicklungen der Wasserstände an der Deutschen Nordseeküste, Dissertation, Technische Universität Braunschweig (in German)

Jensen, J. and Mudersbach, Ch. (2007) Zeitliche Änderungen in den Wasserstandszeitreihen an den Deutschen Küsten. *Berichte zur deutschen Landeskunde*, 81 (2), 99-112 (in German).

Jevrejeva, S., Moore, J.C., and Grinsted, A. (2010) How will sea level respond to changes in natural and anthropogenic forcings by 2100? *Geophysical Research Letters,* Vol. 37, L07703

Katz, R.W., Parlange, M.B., and Naveau, P. (2002) Statistics of extremes in hydrology. *Advances in Water Resources*, 25, 1287-1304.

Khaliq, M.N., Ouarda, T.B.M.J., Ondo, J.-C., Gachon, P. and Bobee, B. (2006) Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review. *Journal of Hydrology*, 329, 534-552.

Kotz, S. and Nadarajah, S. (2000) Extreme Value Distribution - Theory and Applications. Imperial College Press, London.

Meehl, G.A., T.F. Stocker, W.D. Collins, P. Friedlingstein, A.T. Gaye, J.M. Gregory, A. Kitoh, R. Knutti, J.M. Murphy, A. Noda, S.C.B. Raper, I.G. Watterson, A.J. Weaver, and Z.-C. Zhao, 2007: Global Climate Projections. In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Solomon, S.,D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Mendez, F. J., Menendez, M., Luceno, A., and Losada, I. J. (2007) Analyzing monthly extreme sea levels with a time-dependent GEV Model. *Journal of Atmospheric and Oceanic Technology*, 24, 894-911.

MLR (2001) Generalplan Küstenschutz - Integriertes Küstenschutzmanagement in Schleswig-Holstein, Ministerium für ländliche Räume, Landesplanung, Landwirtschaft und Tourismus des Landes Schleswig-Holstein (in German).

Mudersbach, C. and Jensen, J. (2010) Nonstationary extreme value analysis of annual maximum water levels for designing coastal structures on the German North Sea Coastline, *Journal of Flood Risk Management* 3 (2010), 52-62

Müller-Navarra, S., Bork, I., Jensen, J., Koziar, Ch., Mudersbach, Ch., Müller, A., and Rudolph, E. (2006) Modellstudien zur Sturmflut und zum Hamburg-Orkan 1962, *Hansa,* 12, Hamburg, 72-88 (in German)

Nogaj, M., Parey, S., and Dacunha-Castelle, D. (2007) Non-stationary extreme models and a climatic application. *Nonlinear Processes in Geophysics,* 14, 305-316.

Pugh, D. (2004) Changing Sea Levels - Effects of Tides, Weather and Climate. Cambridge University Press, Cambridge.

Rao, A.R. and Hamed, K.H. (2000) Flood frequency analysis. CRC Press, New York.

Ribereau, P., Guillou, A., and Naveau, P. (2008) Estimating return levels from maxima of non-stationary random sequences using the Generalized PWM method. *Nonlinear Processes in Geophysics*, 15, 1033-1038.

Salas, J.D. (1993) Analysis and Modelling of Hydrologic Time Series. In Maidment, D.R. (Ed.): *Handbook of Hydrology*, McGraw-Hill Inc., New York.

Strupczewski, W.G., Singh, V. P., and Feluch, W. (2001) Non stationary approach to at-site flood frequency modelling 1: Maximum likelihood estimation. *Journal of Hydrology*, 248, 123-142