REGIONAL FREQUENCY ANALYSIS OF EXTREME WATER LEVELS ALONG THE DUTCH COAST USING L-MOMENTS: A PRELIMINARY STUDY

Pieter van Gelder[1], Neyko M. Neykov[2]

[1] *Delft University of Technology, Faculty of Civil Engineering, P.O. Box 5048, 2600 GA Delft, Netherlands; E-mail: vangelder@usa.net WWW Home-page: http://surf.to/vangelder*
[2] *National Institute of Meteorology and Hydrology, Bulgarian Academy of Sciences, 66 Tsarigradsko shaussee, Sofia 1784, Bulgaria; E-mail: neyko.neykov@meteo.bg*

Keywords: Extreme Value Theory, Coastal Engineering, L-Moments, Regional Frequency Analysis, Detection of Outliers

ABSTRACT: The regional frequency analysis is widely used in flood analysis. The approach based on the theory of L-moments recently developed by Hosking and Wallis (1997) is a very reliable method for assessing exceedance probabilities of extreme environmental events when data is available from more than one site. The method of L-moments is analogous to the method of ordinary moments. The main advantage of L-moments is that, being a linear combination of data, they are less influenced by outliers whereas the ordinary moments requires squaring and cubing of the observed data. A well conducted regional frequency analysis involves objective and subjective techniques for defining homogeneous regions, assigning of sites to regions, identifying and fitting regional probability distribution to data, and testing hypotheses about distributions. The standard discordancy measure of Wilks for detection of multivariate outliers in terms of the sample L-moment ratios of the site's data is recommended by Hosking and Wallis (1997) as a guideline rather than a formal test during the process of initial data screening for forming the homogeneous region. It is well known, however, that this test is not robust against outliers as it is based on the sample mean and covariance matrix which are themselves affected by outliers. Existing alternatives to it based on robust estimates of the location and scatter are studied on a case study of extreme water levels at several locations along the North Sea coast of the Netherlands. The results show that the robustified analogs are superior and look promising in this case.

1 INTRODUCTION

The Netherlands is a low-lying country which has to protect itself against flooding from the sea and its rivers. Reliable flood defenses are essential for the safety of the country. The sea dikes are designed to withstand floods with a height of once every 10,000 years. This height is used to be calculated using statistics on sea levels measured along the Dutch coast since 1880. Annual maxima and peaks over threshold data are used as input data (Van Gelder, 1996). Gumbel and Exponential models are most commonly applied together with Maximum Likelihood and Least Squares parameter estimation methods.

In order to estimate the risk that a given flood will be exceeded during the design life of a structure, flood frequency analysis is needed to relate the rarity of the flood to its magnitude. The objective of flood frequency analysis is to estimate the flood magnitude corresponding to any required return period of occurrence through the use of probability distributions. However, estimating the frequencies of extreme environmental events such as floods is difficult because extreme events are by definition rare and the relevant data record is often short. Regional Frequency Analysis (RFA) developed by Hosking and Wallis (1997) can resolve this problem by trading space for time. It does so by using data from several sites, which are judged to have frequency distributions similar to the site of interest, in estimating event frequencies at that site. The main stages of the RFA procedure are: (i) screening of the data; (ii) identification of homogeneous regions; (iii) choice of a regional frequency distribution; (iv) estimation of the regional frequency distribution.

As the first three stages of the RFA procedure are subjective Hosking and Wallis (1997) recommended: the standard discordance measure of Wilks, for identifying unusual sites in a region in terms of the sample L-moments ratios as a guideline rather than a formal test; a heterogeneity measure, for assessing whether a proposed region is homogeneous; and a goodness-of-fit measure, for assessing whether a candidate distribution provides an adequate fit to the data. The RFA is an iterative procedure. However, Hosking and Wallis (1997), emphasize physical reasoning rather than formal statistical significance in data processing. In this paper the RFA procedure of Hosking and Wallis (1997) will be slightly adapted with a robust alternative for the standard discordance measure. Furthermore the robustified RFA procedure will be applied to the datasets with extreme sea levels along the Dutch coast. The calculations concerning the RFA were made by LMOMENTS package developed by Hosking (1996).

2 RFA OF EXTREME SEA LEVELS ALONG THE DUTCH COAST

Peaks over thresholds of water level data from a number of gauging stations located along the Dutch coast were collected from the RIKZ in the Hague, the Netherlands. The data set contains a total of 6818 water level observations with sample sizes at the 5 sites varying from 53 to 104 years. In figure 1 the 5 sites are depicted: from north to south they are Delfzijl, Den Helder, Harlingen, Hoek van Holland and Vlissingen.

Figure 1

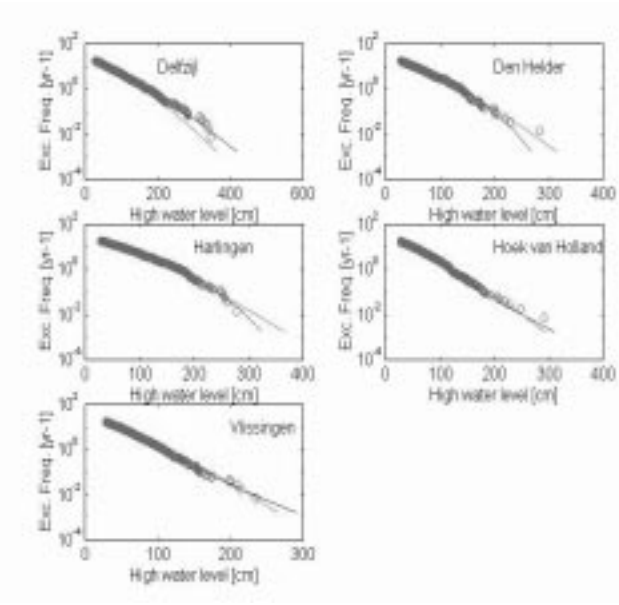


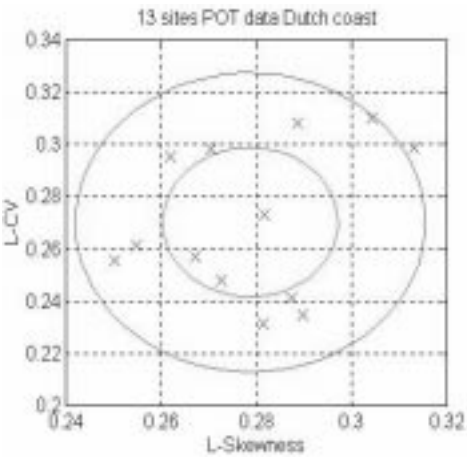
Figure 2

The number of locations were artificially increased by splitting the data sets of a location in 2 or 3 subsets. For example the Delfzijl location consists of data from 1882 until 1985. It is splitted up in data from:

1. 1882-1916
2. 1917-1951
3. 1952-1985

By doing this it is implicitly checked if the homogeneity of the data set of each location individually is valid. Up till now it was assumed that a 104-year dataset from 1882 to 1985 is homogeneous, which is a very strong and not-checked assumption. By splitting it up it automatically takes this assumption into consideration. In this way the number of stations becomes 13. The records of Den Helder and Harlingen are shorter in comparison with the others (only 53 years) and are therefore splitted into 2 subsets (from 1932 to 1951 and 1952 to 1985).

The sample L-moments ratios, i.e, the coefficient of variation, skewness and kurtoses of a distribution are calculated using the L-moments described in Hosking and Wallis (1997). They are denoted by L-Cv, L-C-skew and L-kurt respectively. The average value for L-Cv is 0.27 and for L-Cs is 0.28 based on the results given in Table 1. These values are not particularly high indicating that the frequency distributions are not necessarily highly skew. The L-Cv and L-Cs for the data are shown in figure 3. The figure also shows the unusual sites (outside the inner (1-sigma)- or outer (2-sigma) ellipse) whose data need a closer examination i.e. those sites whose L-moments are notably different from those of the other sites in the data set. For instance the site Delfzijl (1882-1916) on the upper-right side of the 2-sigma ellipse looks to be unusual.



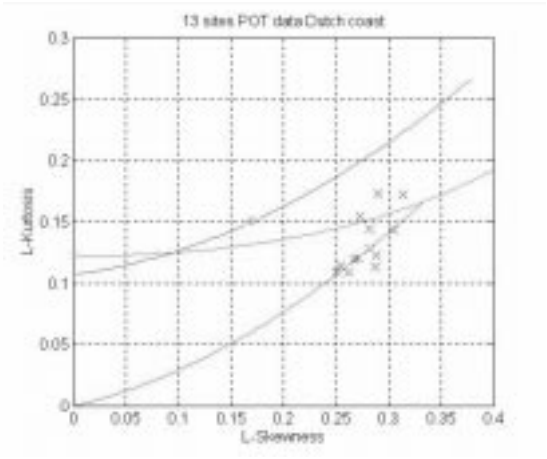


Figure 3 Figure 4

The 13 sites have the following characteristics:

SITE	N	NAME	L-CV	L-SKEW	L-KURT	Di	MDi	RDi(MCD)	RDi(REW)	RDi(HUB)	RDi(HAM)	RDi(T-BW)
1	320	Held19321951	.2611	.2548	.1143	.62	1.729	.816	1.635	1.500	1.488	1.159
2	546	Held19521985	.2730	.2817	.1280	.09	.260	.847	1.696	.434	1.221	.992
3	352	Harl19321951	.2954	.2620	.1092	.83	2.298	1.367	2.738	2.118	1.765	1.305
4	584	Harl19521985	.2982	.2705	.1206	.59	1.642	1.259	2.523	1.547	2.759	2.331
5	627	Delf18821916	.2988	.3132	.1723	1.67	4.626	30.175*	60.472*	7.426	38.448*	33.308*
6	574	Delf19171951	.3080	.2885	.1226	.78	2.150	2.051	4.109	2.258	3.526	2.793
7	566	Delf19521985	.3101	.3043	.1435	1.09	3.023	2.297	4.604	2.937	3.932	3.312
8	474	Hoek18871916	.2481	.2725	.1547	1.04	2.875	37.048*	74.243*	6.733	43.826*	37.678*
9	576	Hoek19171951	.2556	.2502	.1095	.87	2.409	.821	1.645	2.138	1.912	1.446
10	527	Hoek19521985	.2570	.2671	.1195	.23	.630	1.023	2.050	.741	1.718	1.382
11	598	Vlis18821916	.2345	.2897	.1732	1.59	4.395	53.096*	106.405*	9.321	50.999*	43.629*
12	536	Vlis19171951	.2314	.2814	.1447	.79	2.197	10.469*	20.980*	3.056	5.080	4.101
13	538	Vlis19521985	.2414	.2873	.1140	2.80	7.767	33.403*	66.930*	14.909*	70.032*	59.176*

Table 1.

The problem of identifing outliers in data is extremely difficult. It is well known that the standard measure of Wilks for detection of multivariate outliers is not robust against outliers as it is based on the sample mean and covariance matrix which are themselves affected by outliers. Alternatives to it based on robust estimates of multivariate location and scatter, developed by Rousseeuw and van Zomeren (1990) are studied. The recommended discordance measure Di of Hosking and Wallis (1997) and several of its alternatives denoted by RDi(.) based on robust estimates of the location and scatter were applied in the case study.
In table 1 the following discordancy measures have been used.
Di - is the discordancy measure of Hosking and Wallis (1997), which is based on the sample L-moments ratios (L-CV,L-SKEW, L-KURT) for each site. A site is declared as discordant if its $Di \geq 3$;
Mdi - is the classical Mahalanobis distance ($MDi=3(n-1)Di/n$);
RDi(MCD), RDi(REW), RDi(HUB), RDi(HAM), RDi(T-BW) - are the robust distances (analogs of the Mahalanobis distance) based on the Minimum Covariance Determinant (MCD) estimator of the multivariate location and scatter and some of its one-step improvements based on Huber's weights (HUB), Hampel weights (HAM), Constrained M-estimates (T-BW) (see, Rousseeuw and van Zomeren,1990; Rock and Woodroff, 1996). The reader is reminded that if n and p denote the number observations and variables respectively the MCD objective is to find $h=(n+p+1)/2$ observations out of n whose classical covariance matrix has the lowest determinant. The MCD

estimate of the location is then the average of these h points, whereas the MCD estimate of the scatter is their covariance matrix estimators (Rousseeuw and Leroy, 1987). For the robust distances $\chi^2(3,0.975) = 9.36$ is the threshold for which a site is declared as discordant. It is seen that according to the classical Di (MDi) the sites can be considered as regular whereas some are flagged as discordant by the robust distances MCD, REW, HUB, HAM and T-BW. Therefore the degree of heterogeneity is studied.

The degree of hetereogeneity within a group of sites is estimated by the heterogeneity measure H suggested by Hosking and Wallis (1993, 1997). Basically the heterogeneity measure compares the between-site variations in sample L moments for the group of sites with what would be expected for a homogeneous region. What “would be expected” is evaluated through Monte Carlo simulation from the four parameter Kappa distribution. Hosking and Wallis (1993, 1997) suggested that regions can be classified as "acceptably homogeneous" if $H<1$, "possibly heterogeneous" if $1= H<2$, and "definitely heterogeneous" if $H\geq 2$. Treating the 13 sites as a single homogeneous region, the heterogeneity measures were calculated to be 14.81, 4.77 and 1.20. These results indicate that the data set is definitely heterogeneous according to H(1) statistics of Hosking and Wallis (1997). (The H(2) and H(3) statistics based on V2 and V3 respectively lack power to discriminate between homogeneous and heterogeneous regions according Hosking and Wallis (1997)). This same conclusion could be made by the robust distances. The Di statistics was not able to identify the discordant observations. The robustified distances Rdi(.) are superior than the classical Mahalanobis distance and look promising. The result is also in agreement with our physical considerations about the locations. Sites such as Harlingen and Delfzijl are situated behind islands whereas Den Helder and Hoek van Holland lie in front of open sea with thousands of kilometers fetch lengths in north-western direction. The site Vlissingen lies also protected in an estuary. Discordant sites have to be deleted in a RFA. It was found particularly difficult to form a homogeneous region. However Hosking and Wallis (1997) state that even in regions with heterogeneity RFA is more accurate than at-site analysis.

3 GOODNESS OF FIT

Five three-parameter distributions (generalized logistic, generalized extreme value, generalized Pareto, lognormal (LN3), Pearson type III) were fitted to the region. Table 2 shows that the generalized Pareto is acceptable, according to the Hosking and Wallis (1997) goodness-of-fit measure.

```
  GEN. LOGISTIC        L-KURTOSIS=  .229   Z VALUE= 14.38
  GEN. EXTREME VALUE          L-KURTOSIS=  .199   Z VALUE= 10.32
  GEN. NORMAL                 L-KURTOSIS=  .181   Z VALUE=  7.86
  PEARSON TYPE III     L-KURTOSIS=  .150   Z VALUE=  3.56
  GEN. PARETO          L-KURTOSIS=  .123   Z VALUE=  -.15 *
PARAMETER ESTIMATES FOR DISTRIBUTIONS ACCEPTED AT THE 90% LEVEL
  GEN. PARETO      .416   .666   .142
  WAKEBY              .416   .446   .221   .225  -.029
QUANTILE ESTIMATES
  P-VALUE     .010  .020  .050  .100  .200  .500   .900  .950  .990  .999
  GEN. PARETO .423  .430  .450  .486  .563  .856  1.725 2.043 2.670 3.351
  WAKEBY      .422  .429  .450  .486  .563  .857  1.723 2.040 2.677 3.408
```

Table 2.

Figure 4 is the L-moment ratio diagram showing the L-skewness and L-kurtosis of the individual sites, including the curves of the various distributions. The L-moment ratio diagram shows that generally speaking we can use the generalized Pareto distribution for the region. Also the five-parameter Wakeby distribution can be used. This distribution is also robust to

miss-specification of the underlying distribution function of a homogeneous region and is a recommended choice for heterogeneous regions by Hosking and Wallis (1997).

Figure 2 shows quantiles of the regional frequency distribution, obtained by fitting the best-fit distribution to each site's data using the method of regional L-moments. The figure shows that the quantiles for the different sites are quite similar in the central area of the distribution (i.e. non-exceedance probability equal to 0.5). However, for non-exceedance probabilities greater than 0.95, marked differences become clear. The fitted distribution is the generalized Pareto at site (in bold) and regionalised (in light).

4 CONCLUSIONS

The 'preliminary' conclusions of this study are:

The regional frequency analysis which is usually applied to river flow data has been used in this paper to examine the extreme sea levels along the Dutch North Sea coast. The sites along the Dutch coast do not form a homogeneous region. This was already clear from physical considerations, and the statistical measures have confirmed this expectation. It is recommended to include sites along the North Sea coast of other countries such as Belgium, England, Germany, France, Sweden, Norway, and Denmark. The collection of POT-data of extreme water levels from these countries have already started. In the case study with only Dutch sites, only the sites with time periods:

Den Helder 1932-1951
Den Helder 1952-1985
Delfzijl 1952-1985
Hoek van Holland 1952-1985

appeared to be 'regular' according to the robust distances. It was seen, however, that according to the measure H these sites do not form a totally homogeneous region (H(1) = 6.72). Hosking and Wallis stated however that even in regions with heterogeneity RFA is more accurate than at-site analysis. The RFA resulted in a best distribution for the extreme sea levels along the Dutch coast: the Generalized Pareto.

Apart from the collection of foreign sites, the process of generation of new sites can also be continued by splitting the data.

The inference based on robust distances and H measures complement and agreed each other for these particular data sets. It is expected that in general the robust distances can better be used than discordancy measure Di of Hosking and Wallis (1997). This should be proved by means of an extended Monte Carlo investigation.

Acknowledgements

The authors are grateful to Dr.’s P. Neytchev and V. Todorov for valuable assistances in programming of the robust distanses and some extensions based on LMOMENTS package.

4 REFERENCES

Hosking,J.R.M., and Wallis,J.R.,(1993), Some Statistics Useful in Regional Frequency Analysis, Water Resources Research, 229, pp. 22271-281.

Hosking,J.R.M.(1996). Fortran routines for use with the method of L-moments, Version 3. Research Report RC205225, IBM Research Division, Yorktown Heights, N.Y.

Hosking,J.R.M., and Wallis,J.R.,(1997), Regional Frequency Analysis, An Approach Based on L-moments, Cambridge University Press.

Van Gelder, P.H.A.J.M., (1996), A new statistical model for extreme water levels along the Dutch coast. Stochastic Hydraulics, pp. 243-250, A.A. Balkema, Mackay.

RIKZ, Rijksinstituut voor Kust en Zee (Ministry of Coast and Sea), The Hague, The Netherlands.

Rousseeuw,P. and A. Leroy,(1987), Robust Regression and Outliers Detection, Wiley, New York

Rousseeuw,P. and B. van Zomeren,(1990) Unmasking Multivariate Outliers and Leverage Point (with discussion), Journal of the American Statistical Association, 85, pp.633 - 651).

Rocke,D.M. and Woodruff,D.L., (1996). Identification of Outliers in Multivariate Data, Journal of the American Statistical Association, 91, pp. 1047-1061.