Improving 10^4-year surge level estimates using data of the ECMWF seasonal prediction system

H. W. van den Brink, G. P. Können, J. D. Opsteegh, G. J. van Oldenborgh, and G. Burgers
Royal Netherlands Meteorological Institute, De Bilt, Netherlands

Received 6 May 2004; accepted 16 August 2004; published 9 September 2004.

[1] The vulnerability of society on extreme weather has resulted in extensive research on the statistics of extremes. Although the theoretical framework of extreme value statistics is well developed, meteorological applications are often limited by the relative shortness of the available datasets. In order to overcome this problem, we use archived data from all past seasonal forecast ensemble runs of the European Centre for Medium-Range Weather Forecasts (ECMWF). For regions where the forecasts have very little seasonal skill the archived seasonal forecast ensembles provide independent sets that cumulate to over 1500 years. We illustrate this approach by estimating 10^4-year sea-surge levels at high-tide along the Dutch coast. No physical mechanisms occur in the ECMWF model that make the distribution of very extreme surges different from what is inferred from a direct analysis of the observations. In comparison with the observational sets, the ECMWF set shows a decrease in the statistical uncertainty of the estimated 10^4-year return value by a factor four. *INDEX TERMS:* 3309 Meteorology and Atmospheric Dynamics: Climatology (1620); 3364 Meteorology and Atmospheric Dynamics: Synoptic-scale meteorology; 3367 Meteorology and Atmospheric Dynamics: Theoretical modeling; 4564 Oceanography: Physical: Tsunamis and storm surges; 1821 Hydrology: Floods. **Citation:** van den Brink, H. W., G. P. Können, J. D. Opsteegh, G. J. van Oldenborgh, and G. Burgers (2004), Improving 10^4-year surge level estimates using data of the ECMWF seasonal prediction system, *Geophys. Res. Lett.*, *31*, L17210, doi:10.1029/2004GL020610.

1. Introduction

[2] Meteorological extremes have large impacts on society. Typical examples are flooding of rivers caused by extreme precipitation, extended droughts, extreme temperatures, and flooding from the sea caused by extreme wind speeds. The higher the extremes, the more difficult it is to obtain their statistics from the observational datasets. However, these very extreme meteorological situations cause the most disastrous events. For many types of extremes, the meteorological situations causing these extreme events are of synoptic scale ($O(10^3$ km)), and last for longer times (>12 hours). These properties make them appropriate to be explored with the set of ensembles of simulations generated by the ECMWF seasonal prediction system [*Anderson et al.*, 2003]. The resolution in space and time of the dataset is high enough to resolve extremes on synoptic scales.

[3] The ECMWF seasonal forecast dataset has two properties advantageous for examining current-climate extremes. First, it combines high resolution in space (1.875°, 40 levels) and time (6-hourly output) with large record length (1569 years in total by May 2004). This length exceeds that of most high-resolution climate model runs [*Kharin and Zwiers*, 2000; *Kysely*, 2002; *Kiktev et al.*, 2003]. Second, the ECMWF model does not drift far from the observed climatology, as the individual forecast ensemble members are only 6 months in length. Here, we illustrate the power of the dataset by estimating extreme surge levels along the Dutch coast.

2. GEV Analysis of Observed Surges

[4] Approximately 40% of the Netherlands is below sea level. This part, with millions of inhabitants, is protected against flooding from the sea by dikes. Dutch official policy is that a flooding event is 'allowed' to happen with a probability of at most 10^{-4} per year, hence with a mean return period of 10^4 years. However, the heights of the dikes that correspond to this probability are hard to determine from Dutch observational sea level records, which cover order hundred years. So, an extrapolation over two orders of magnitude in probability is required, resulting in an 95%-confidence interval of several meters, which is considerably larger than the value of the expected sea level rise in the coming century [*Church et al.*, 2001].

[5] We follow the common choice in empirical studies to fit the annual maxima to the Generalized Extreme Value (GEV) distribution, and plot the results on a Gumbel plot, i.e., a plot with the ordered values on the ordinate and on the abscissa the Gumbel variate $x = -\ln(-\ln(F(x)))$, with $F(x)$ the cumulative distribution function of the variable x.[1] Figure 1 shows the annual maxima of the 117-years observational record for the Dutch coastal station Hoek van Holland, the fitted GEV distribution, its extrapolation to a return period of 10^4 years, and the corresponding 95%-confidence interval at this return period.

[6] There are no known physical processes that limit the surge height to values below the estimated upper 95%-confidence level (6.44 m for Hoek van Holland). The high 95%-confidence level has large consequences for dike design. Another uncertainty in the extrapolation from the ~100 years of observations is the question if all surge extremes up to the 10^4-year return period can be described by one GEV distribution with fixed parameters, a condition which is not always satisfied [*van den Brink et al.*, 2004].

[1] Auxiliary material is available at ftp://ftp.agu.org/apend/gl/2004GL020610.

0094-8276/04/2004GL020610$05.00

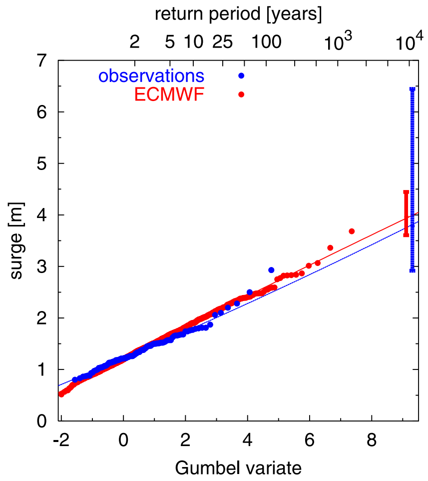

Figure 1. Gumbel plot for the 117 annual surge maxima of the Hoek van Holland observational set for 1887–2004 (blue), and for the 1569 annual maxima according to the archived data generated by the ECMWF seasonal forecast ensembles for 1987–2004 (red). Also shown are the GEV fits up to a return period of 10^4 years and the 95%-confidence interval of the 10^4-year return level.

[7] The ECMWF dataset offers the possibility to check this condition, as well as to decrease the statistical uncertainty in the 10^4-year estimate, due to the thirteen times larger amount of data compared to the observations.

3. ECMWF Model

[8] Since August 2001 the European Centre for Medium-range Weather Forecasts (ECMWF) produces every month an ensemble of 40 global seasonal forecasts up to six months ahead, i.e., amply surpassing the 2-weeks horizon of weather predictability from the atmospheric initial state. Over the period 1987–2001, hindcasts, that is forecast runs on historical data, have been performed with smaller ensembles for calibrating the forecast system. The system consists of a coupled atmosphere-ocean model [*Anderson et al.*, 2003]. The atmospheric component has a horizontal resolution of T95 (1.875°) and 40 levels in the vertical [*Ritchie et al.*, 1995; *Gregory et al.*, 2000; *Anderson et al.*, 2003]. The ocean component has a resolution of 1.4° and 29 vertical levels [*Wolff et al.*, 1997]. The ECMWF dataset provides, among other fields, global fields of 6-hourly winds and 2 m-temperatures, 12-hourly sea level pressures and temperatures, and 24-hourly precipitation amounts.

[9] We constructed 1569 calendar years by combining pairs of ensemble members with six months difference in starting date (see supplementary information[1] for details), all of them generated by the so-called System-2 [*Anderson et al.*, 2003].

[10] Since the ECMWF model has very limited skill in predicting the NAO index [see also *Palmer et al.*, 2004], effectively the simulations sample all different NAO situations.

[11] The GEV location parameter μ for the annual maximum of 6-hourly wind speed (averaged between 30°N–60°N and 90°W–30°E) is constant within 1% for different forecast times. This indicates that the wind climatology of the system shows no detectable deterioration with forecast time.

[12] The dependence between the ensemble members in the first weeks of the forecasts has negligible influence on the estimates of the GEV parameters of the surge.

[13] We compared the daily-mean annual minima of the sea level pressure (SLP) at the Dutch coastal station Den Helder over 1906–2004 with the 1569 annual minima of the ECMWF SLP at the nearest sea grid point to Den Helder. The Gumbel plot is shown in Figure 2. There is a good agreement between the annual minima of the ECMWF data and the observations.

[14] We conclude that the ECMWF seasonal forecast system generates (deep) depressions with the same frequency and intensity as observed.

4. Surge Equation

[15] We use the following equation [*van den Brink et al.*, 2003] to calculate from the meteorological data the surge at high-tide (i.e., the difference between the observed high tide and the calculated height of the astronomical high tide) at the coastal station Hoek van Holland:

$$\text{Surge} = A\, C_d\, u_{10}^2\, \sin(\phi - \beta) + \frac{1015 - \text{SLP}}{100.5} \quad [\text{m}] \qquad (1)$$

with C_d the drag coefficient, u_{10} the wind speed at 10 m averaged over 12 hours at a central grid box over the North Sea (depicted in Figure 3a), ϕ the wind direction and A and β empirically determined constants by *Timmerman* [1977]. The second term on the right hand side of equation (1) represents the barometric pressure effect, with SLP the instantaneous sea level pressure in Hoek van Holland in hPa. Here, we describe the dependence of C_d on the wind speed u_{10} as

$$10^3 C_d = 0.738 + 0.068\, u_{10} \qquad (2)$$

where u_{10} is expressed in ms^{-1}, and in which the constants were obtained from a linear fit between the instantaneous,

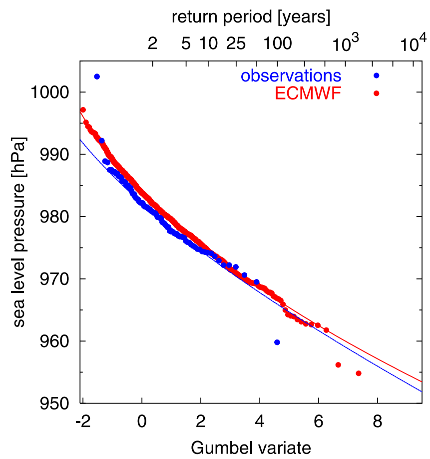

Figure 2. Gumbel plot for the 98 annual SLP minima of the Den Helder observational set for 1996–2004 (blue), and for the 1569 annual SLP minima of the ECMWF dataset (red). The lines are the GEV fits up to a return period of 10^4 years.

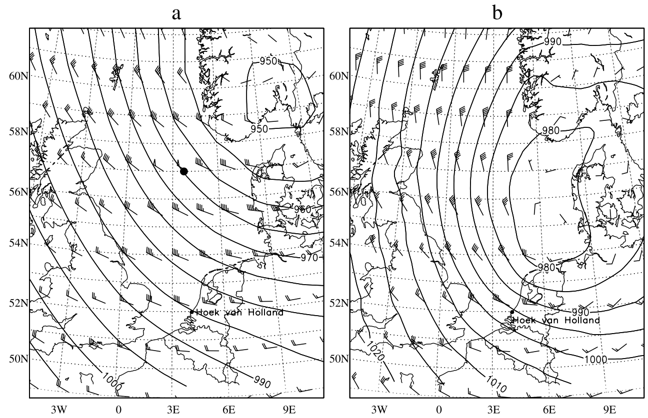

Figure 3. Wind and pressure field for the situations of highest surge at high-tide in Hoek van Holland. **a**: in the ECMWF dataset for day 57 in ensemble member 15, starting from 1 November 1987 (calculated surge in Hoek van Holland: 3.68 m). The dot is the location used for calculating the surge at high-tide via equations (1) and (2). **b**: ECWMF Reanalysis, for the highest surge in the observations for Hoek van Holland (2.93 m, on 1 February 1953).

once-a-day available drag coefficient of the ECMWF dataset and u_{10}. To assure that equation (2) describes optimally the specific case of strong north-westerly winds, we used in the determination of the constants in equation (2) only those situations that resulted in the annual maximum surges in Hoek van Holland. In the 12–27 m/s range, which covers our range of interest, equation (2) fits closely to a Charnock relation [*Charnock*, 1955] with parameter 0.016. Our estimate compares well with other estimates of high-speed drag over sea [e.g., *Smith et al.*, 1992; *Bonekamp et al.*, 2002].

[16] The surge equation was validated by comparing the 1957–2002 observed annual extreme surges in Hoek van Holland with the annual extreme surges calculated from equations (1) and (2) using the wind and pressure of the ERA40-Reanalysis data [*Simmons and Gibson*, 2000]. Supplementary Table 3 and supplementary Figure 1 show good agreement between the GEV distributions fitted to the observed and calculated surges. About 1/2 of the annual extremes according to the ERA40 dataset correspond to the same storm as the annual extremes in the observations.

5. Results

[17] We calculated the surge at high-tide for coastal station Hoek van Holland with equations (1) and (2), using SLP and u_{10} of the ECMWF dataset. The 1569 annual extremes are shown on a Gumbel plot in Figure 1, together with the 117 annual extremes of the 1887–2004 observational set. The following four features are apparent from Figure 1. First, the ECMWF-based data indicate that for extreme surges, a single GEV distribution is appropriate up to return periods of at least 10^3 years. So, the ECMWF data gives no indication that physical processes limit the strength of extreme storms, nor that the 10^3-year winds are caused by another type of storms than 10-year winds, as in the less comprehensive model discussed in *van den Brink et al.* [2004]. Second, the GEV location parameter μ (representing the surge level with an exceedance probability of once a year) estimated from the ECMWF dataset equals that of the observational record within one cm (see also supplementary Table 4). This implies that systematic differences between the observed data and the results from the ECMWF system with the surge equation are small compared to the statistical uncertainties. Third, 10^4-year surge level estimates from the ECMWF dataset (3.96 m) and from the observational record (3.78 m) are nearly equal. Fourth, the 95%-confidence interval of the 10^4-year estimate reduces from 3.52 m for the observational set to 0.84 m for the ECMWF set, i.e., a reduction by a factor four.

[18] The meteorological situation in the ECMWF data that leads to the highest surge at Hoek van Holland (3.68 m) is depicted in Figure 3a. For comparison, Figure 3b shows the meteorological situation according to the ECMWF-Reanalysis of the largest real event in the observations (2.93 m, on 1 February 1953). Both situations show a large-scale depression, generating a strong north-westerly flow over the entire North Sea. The 25 hPa deeper depression and the more north-easterly position of the depression in Figure 3a with respect to the situation in Figure 3b leads to a 0.75 m higher surge level at Hoek van Holland. Figure 3 shows that the largest surge from the ECMWF dataset is caused by a realistic meteorological situation.

6. Conclusions

[19] The ECMWF seasonal forecast dataset can serve as a powerful tool for estimating 10^3–10^4-year return values for meteorological extremes that are caused by synoptical weather systems. The statistics of extreme storm surge levels in the Netherlands in this dataset can be described well by a single GEV distribution for return periods ranging from 1 to 10^3 years. The statistical uncertainty in the 10^4-year return surge level is reduced by a factor four with respect to observations. The dataset offers potentials to estimate 10^3- to 10^4-year return values for wind, temperature, precipitation and related variables as surge and river discharges, with unprecedented accuracy.

[20] **Acknowledgment.** We thank the ECMWF for providing the seasonal forecast data.

References

Anderson, D. L. T., et al. (2003), Comparison of the ECMWF seasonal forecast systems 1 and 2, including the relative performance for the 1997/8 El Niño, *Tech. Memo. 404*, Eur. Cent. for Medium-Range Weather Forecasts, Reading, U. K.

Bonekamp, H., G. J. Komen, A. Sterl, P. A. E. M. Janssen, P. K. Taylor, and M. J. Yelland (2002), Statistical comparisons of observed and ECMWF modeled open ocean surface drag, *J. Phys. Oceanogr.*, *32*, 1010–1027.

Charnock, H. (1955), Wind stress on a water surface, *Q. J. R. Meteorol. Soc.*, *81*, 639.

Church, J. A., J. M. Gregory, P. Huybrechts, M. Kuhn, K. Lambeck, M. T. Nhuan, D. Qin, and P. L. Woodworth (2001), Changes in sea level, in *Climate Change 2001: The Scientific Basis: Contribution of Working Group 1 to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, edited by J. Houghton et al., pp. 639–694, Cambridge Univ. Press, New York.

Gregory, D., J. J. Morcrette, C. Jakob, A. C. M. Beljaars, and T. Stockdale (2000), Revision of convection, radiation and cloud schemes in the ECMWF integrated forecasting system, *Q. J. R. Meteorol. Soc.*, *126*, 1685–1710.

Kharin, V. V., and F. W. Zwiers (2000), Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere-ocean GCM, *J. Clim.*, *13*, 3760–3788.

Kiktev, D., D. M. H. Sexton, L. Alexander, and C. K. Folland (2003), Comparison of modeled and observed trends in indices of daily climate extremes, *J. Clim.*, *16*, 3560–3571.

Kysely, J. (2002), Comparison of extremes in GCM-simulated, downscaled and observed central-European temperature series, *Clim. Res.*, *20*, 211–222.

Palmer, T. N., et al. (2004), Development of a European multi-model ensemble system for seasonal to inter-annual prediction (DEMETER), *Bull. Am. Meteorol. Soc.*, in press.

Ritchie, H., C. Temperton, A. J. Simmons, M. Hortal, T. Davies, D. Dent, and M. Hamrud (1995), Implementation of the semi-Lagrangian method in a high resolution version of the ECMWF forecast model, *Mon. Weather Rev.*, *123*, 489–514.

Simmons, A. J., and J. K. Gibson (2000), The ERA-40 project plan, *Tech. Rep. ERA-40*, Eur. Cent. for Medium-Range Weather Forecasts, Reading, U. K.

Smith, S. D., et al. (1992), Sea surface wind stress and drag coefficients: The HEXOS results, *Boundary Layer Meteorol.*, *60*, 109–142.

Timmerman, H. (1977), *Meteorological Effects on Tidal Heights in the North Sea*, Staatsdrukkerij, The Hague, Netherlands.

van den Brink, H. W., G. P. Können, and J. D. Opsteegh (2003), The reliability of extreme surge levels, estimated from observational records of order hundred years, *J. Coastal Res.*, *19*, 376–388.

van den Brink, H. W., G. P. Können, and J. D. Opsteegh (2004), Statistics of extreme synoptic-scale wind speeds in ensemble simulations of current and future climate, *J. Clim.*, in press.

Wolff, J.-O., E. Maier-Reimer, and S. Legutke (1997), The Hamburg Ocean Primitive Equation model HOPE, *Tech. Rep. 13*, Dtsch. Klimarechenzent., Hamburg, Germany.

G. Burgers, G. P. Können, J. D. Opsteegh, H. W. van den Brink, and G. J. van Oldenborgh, Royal Netherlands Meteorological Institute, Wilhelminalaan 10, NL-3730 AE De Bilt, Netherlands. (brinkvdh@knmi.nl)