Supplementary Materials with "An appraisal of the value of simulated weather data for quantifying coastal flood hazard in the Netherlands"

Cees de Valk[1] and Henk van den Brink[1]
[1]KNMI, PO Box 201, 3730 AE De Bilt, The Netherlands

Abstract. References to numbered sections, figures, tables and equations without "S" refer to the main text.

S1 Further details about the example in Section 2

S1.1 Estimates

Fig. 1 in Section 2 shows fits of a GP (Generalized Pareto) tail (Section 4, eq. (3)). Fig. S1 shows the estimates based on a different type of tail distribution, the Generalized Weibull (GW) tail (Section 4, eq. (5)). Using the GW tail, the confidence intervals are clearly less wide than if a GP tail is used (compare Fig. 1). However, the uncertainties are still of a similar order of magnitude, so the conclusions derived from Fig. 1 hold true irrespective of the choice of the tail type.

For the tail estimation and confidence intervals, the same methods were used as described in Supplement S7. In this example, the estimated return values were considered to be log-normally distributed rather than normally distributed, as this produces more plausible intervals in the case of very high uncertainty (for positive return values, both are asymptotically valid approximations).

S1.2 Differences between estimates from subrecords

The large difference between the estimates of return values for the two time periods in Figs. 1 and S1 could indicate a slow change in the storm climate, but (a) the data only provide a weak indication ("not significant") for this, as under the assumption of identical distributions, the standard deviation of this difference is already twice as large as the standard deviation of the estimate from the complete dataset, (b) similar differences are found when splitting the data of odd and even years, (c) it is not evident from the checks in Supplement S3, (d) such differences are not observed in the tide gauge data from tide gauges in the north of the Netherlands (not shown here), and (e) the difference is mainly associated with the estimates of the shape parameter, whereas the SEAS5 data indicate that climate fluctuations result in fluctuations in scale only; see Appendix D of de Valk and van den Brink (2023b). Moreover, even if the storm climates in both epochs differed, this by itself would not alter the conclusion that the uncertainty is large: we cannot (yet) distinguish between a trend in the distribution function (which could reduce the uncertainty) and a random fluctuation on a timescale of decades (which could increase the uncertainty).

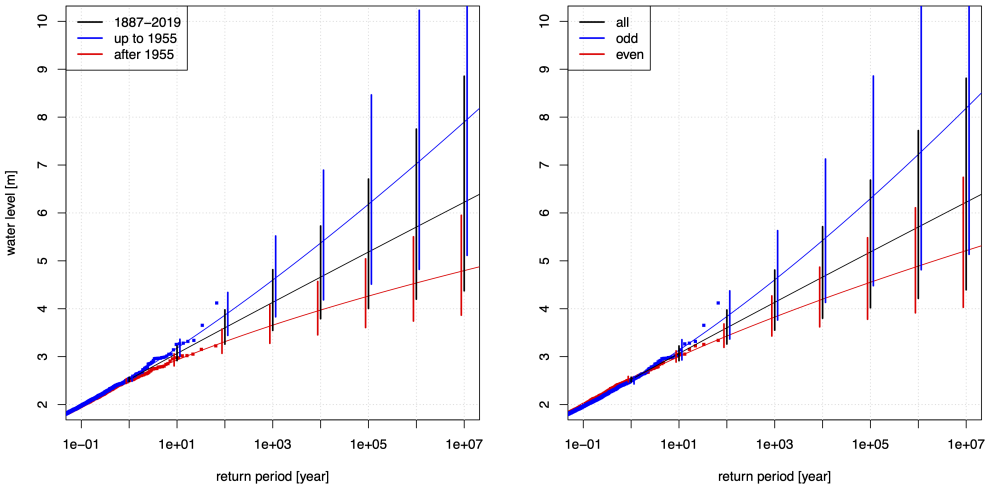

Figure S1. As Fig. 1, but estimates are based on the Generalized Weibull (GW) tail (see Section 4 and eq. (S7)).

S2 Tail dependence of stress from different ensemble members

To check for dependence between the extremes from two time-series, a sensitive metric is the coefficient of tail dependence η from (Ledford and Tawn, 1996, 1997, 1998); see Subsection 7.4. We consider estimates of $\rho := 2\eta - 1$, which can be interpreted similarly as an ordinary Pearson correlation coefficient for two Gaussian random variables (although it is estimated in a very different way!). In particular, 0 means that conditional on one of the two variables exceeding a high threshold, they are effectively independent.

To check the dependence between extremes of stress from different ensemble members of the SEAS5 forecast data for ranges in excess of one month used in de present study, we estimated ρ from all pairs of simultaneous values from different ensemble members. Estimates are made from different fractions p of the "most extreme" pairs. Because the estimates may be biased by seasonality in the data, we made estimates from all data, but also for the long storm season ONDJFM and a shorter season DJF; see Fig. S2.

For both the long and the short storm season, the tail correlation coefficient tends to practically 0 for small sample fractions, showing that the dependence between extremes of stress from different members is negligible. This supports the treatment of data from different SEAS5 ensemble members as independent in the statistical analysis.

S3 Detrending of HW measurements

Fig. S3 shows the annual mean HW and its trend line relative to the trend value in 2019 (estimated using a loess filter with span equal to $84/n$ with n the record length) for the six tidal stations. Additionally, the estimated trend line for the annual mean sea

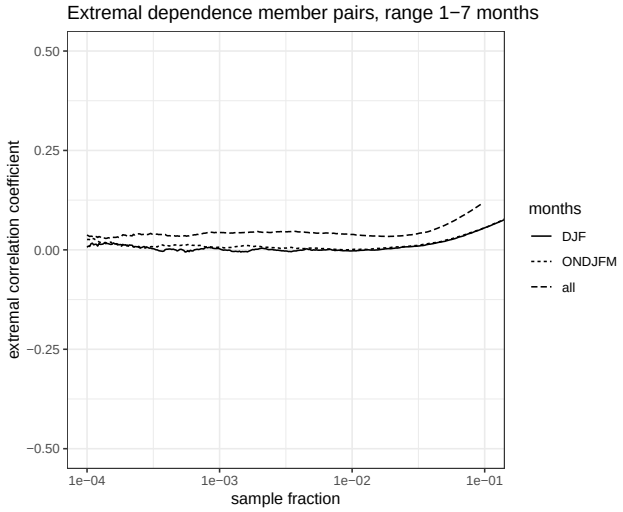

Figure S2. Estimates of tail correlation coefficient $\rho := 2\eta - 1$ from all pairs of ensemble members of SEAS5 forecasts of stress for ranges exceeding 1 month from grid point b (see Fig. 2) for months DJF (drawn), ONDJFM (dotted) and all year (dashed), derived from different sample fractions.

level (relative to its value in 2019) is plotted. The trends are increasing and, except for Den Helder, larger than the sea level rise. The plots for most stations show an accelerated rise in annual mean HW between roughly 1940 and 1980 which is not seen in the sea level rise. These deviations are largely attributed to hydraulic engineering interventions (Dillingh, 2013) but natural processes may also play a role. The records of HW were de-trended assuming that a change in the average HW has the same effect on the extreme HW values.

The considerable differences between the trends of annual mean HW and of annual mean sea level observed in Fig. S3 indicate that extrapolated trends of annual mean sea level may not be suitable for normalizing reconstructed peak values of HW or skew surge of historical flood events as in (Van Gelder, 1996; Baart, 2015). Using these reconstructions in the estimation of return levels to increase the accuracy (see Section 2) may bias the estimates.

To check the de-trending of the measured HW, several checks have been performed on the time evolution of the more extreme HWs after the correction: the annual 99% quantiles and the annual maxima; see Fig. S4. Clear residual trends cannot be discerned in these; apparent deviations are limited to short periods at the beginning or end of a record, where the trend estimation may be less accurate.

In the 99% quantiles, a very small fluctuation can be observed with a minimum around 1960 and maximum around 1920 and 1990, which resembles the pattern of the winter North Atlantic Oscillation (NAO). However, this is of little practical relevance.

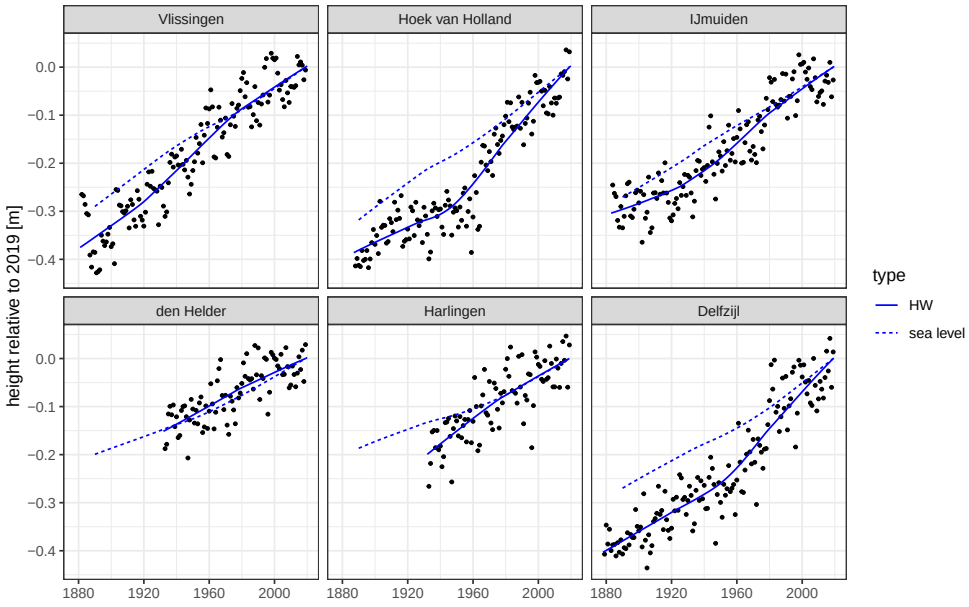

Figure S3. Time-series of annual mean HW (dots) with trend line (full) relative to the trend line value in 2019, and trend line of annual mean sea level relative to its value in 2019 (dashed).

S4 Saturation of the drag coefficient

Measurements in the laboratory and at sea (e.g. Curcic (2020); Richter et al (2021) and their references) indicate that the drag coefficient over sea saturates to a value of about 0.003 for wind speeds in the range of 25-35 m/s. Various empirical and physics-based models have been proposed, but the issue is far from settled. In fact, the ECMWS IFS model, which employs a dynamic wind-wave coupling to compute stress and near-surface wind speed, exhibits saturation even though this is not explicitly incorporated in the model (Pineau-Guillou et al, 2018). This is also seen in the SEAS5 data (generated using the IFS model); see Fig. S5 for an example (this drag relation from SEAS5 should not be trusted quantitatively, because empirical return values of wind speed from SEAS5 along the Dutch coast are considerably lower than those from wind measurements at nearby stations (not shown)). Both measurements (Fig. 3 of Curcic (2020)) and these simulated data suggest that drag saturation, if real, should be rather smooth, hence would not lead to an anomaly in the tail distribution of stress and/or near-surface wind speed. Furthermore, if drag saturation would not affect the geostrophic wind, then it would affect near-surface wind speed more than it affects stress, as can be assessed using the simple expression of Blackadar and Tennekes (1968) of wind and stress in a neutrally stable boundary layer as functions of the geostrophic wind; see also Zweers et al (2012). Finally, two-way interaction between stress and cyclone strength (Section 5.3) may further reduce the influence of drag saturation on the tail of stress.

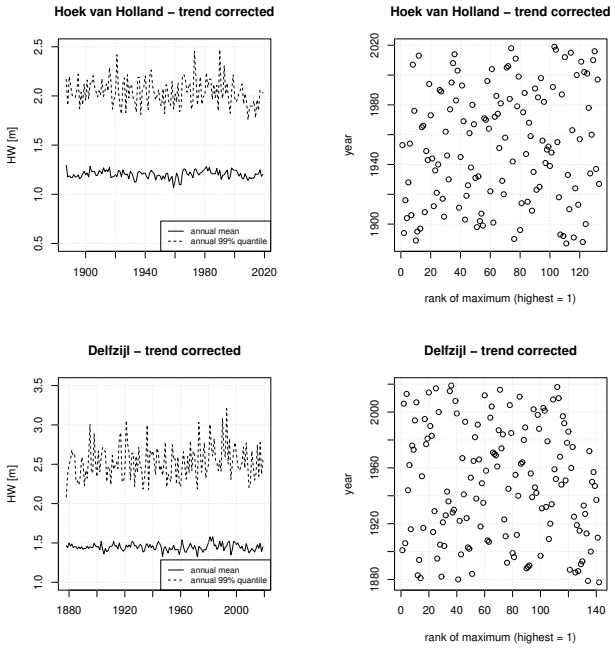

Figure S4. Annual means and 99% quantiles of trend-corrected HW (left), and year as function of the rank of the annual maximum of HW (highest = 1) (right) for Hoek van Holland (top) and Delfzijl (bottom).

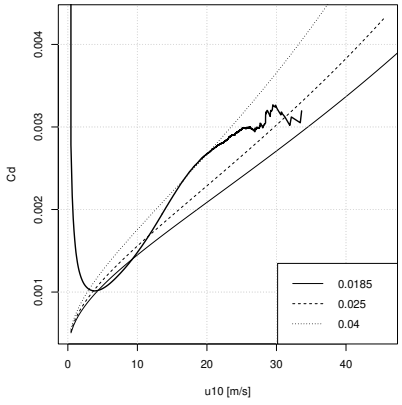

Figure S5. Empirical drag coefficient C_d from SEAS5 data for grid point b in Fig. 2 (thick). The thin lines are determined from Charnock relations with various constants (see legend)

.

S5 Tail estimates for stress

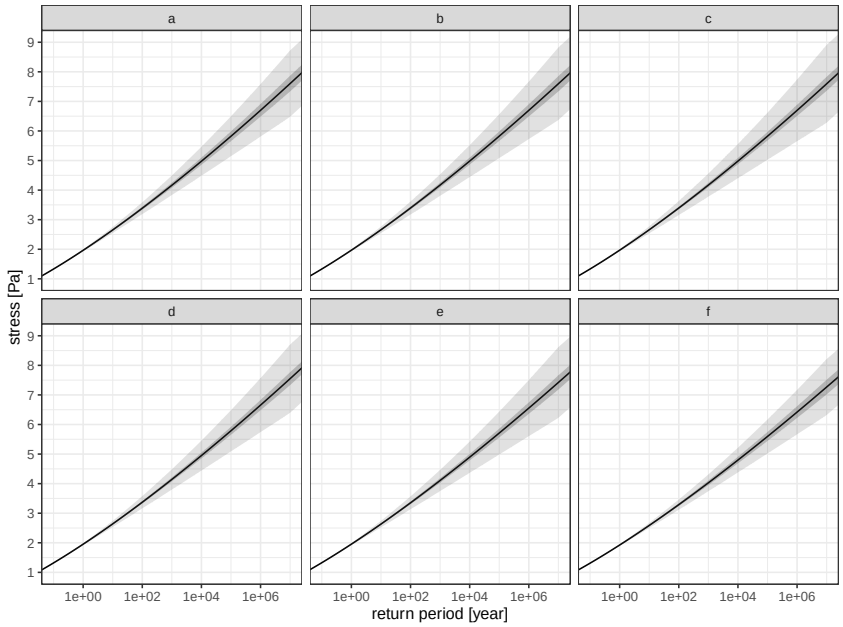

Figure S6. Return value estimates for stress derived from the corrected SEAS5 stress data with sampling error (95% confidence intervals; dark grey) and assessment of total uncertainty (idem, light grey). Grid point labels in headers refer to the triangles in Fig. 2.

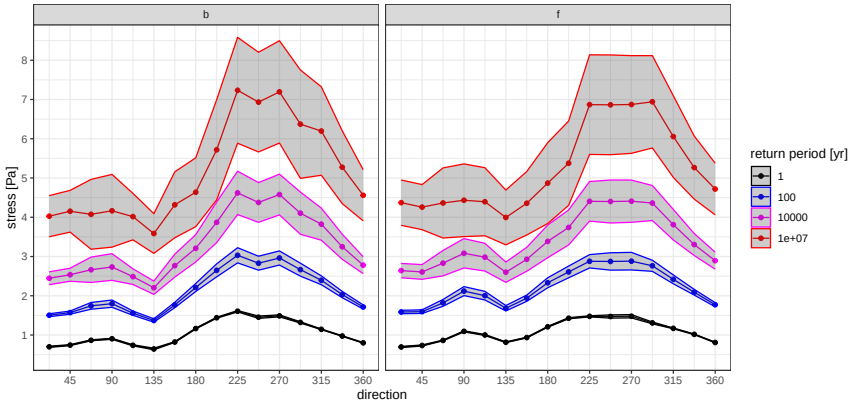

Figure S7. Return value estimates for stress restricted to 22.5° wind direction sectors derived from the corrected SEAS5 stress data with 95% confidence intervals (representing total error) for the locations labelled b and f in Fig. 2.

S6 Estimates of the extremal indices of stress/wind speed and HW

Frequencies (of exceedance) are derived from estimates of the probability, to be interpreted as fraction of time (for HW: fraction of tidal cycles). The probability of exceeding a value z is denoted as $1 - F(z)$ with F the (cumulative) distribution function. The frequency $\mu(z)$ of exceeding a high value z is determined as

$$\mu(z) = (1 - F(z))\alpha/\Delta, \tag{S1}$$

with Δ the time step (for example, 12.42/24/365.25 years for HW) and α the extremal index (EI) (Leadbetter et al, 1983): this is a factor that accounts for the clustering of extremes in time (see Section 7.1).

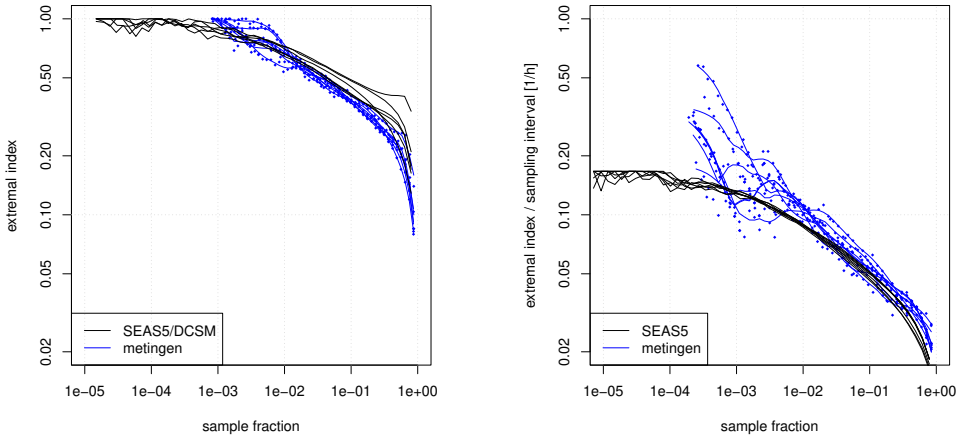

Figure S8. Left: estimates of the extremal index for HW in DJF from SEAS5/DCSMv5 data (black) and measurements (blue dots) from 6 stations; blue lines represent smoothed estimates from measurements. Right: estimates of the extremal index divided by time step for SEAS5 stress (black) and measured wind speed (blue) from nearby weather stations in DJF.

We estimate the EI using the method from Ferro and Segers (2003) from data for DJF only, to minimize bias due to seasonality. An estimate is determined from all values above a chosen threshold, indicated by the sample fraction exceeding it. Fig. S8 (left) shows the estimates for HW for the six tidal stations (see Section 3). We focus on the estimates for the lower sample fractions (high HW). For SEAS5/DCSMv5 data and measurements, the EI α for HW appears to be close to 1. For measured wind speed and SEAS5 stress, Fig. S8 (right) shows estimates of α/Δ, the extreme value index α divided by the time step Δ (6 hours for SEAS5 stress and 1 hour for measurements). For SEAS5 stress, these tend to 0.15-0.20 per hour, hence they are limited by the time step. However, for wind speed measurements, the estimates reach higher values and the increase with decreasing sample fraction does not seem to stop. Based on these results and other considerations, a value of 0.5/hour is chosen for α/Δ, corresponding to an average duration of 2 hours for a cluster of high values (see Section 7). The same values of α/Δ are used in the estimation directional frequencies of exceedance (see below).

S7 Methods for the estimation of return values of HW and stress in Section 7

S7.1 Omnidirectional tails

To approximate the tail of the distribution function F, the Generalized Weibull (GW) tail is used; see Section 4. Let $y = \log 1/p$ with p a chosen "threshold" probability. With $q(y)$ the value exceeded with probability e^{-y} as in eq. (1) and $f(y) = yq'(y)$, the GW tail approximation can be written as

$$1 - F(z) \approx e^{-y\left(1+\beta\left(\frac{z-q(y)}{f(y)}\right)\right)^{1/\beta}} \quad \text{for all} \quad z \geq q(y) \tag{S2}$$

with β some a real number (see (4)).

With this tail approximation, frequencies of exceedance are computed by (S1), using an estimate of the extremal index (Supplement S6).

The precise nature of the tail approximation (S2) is explained in de Valk (2016a). The free parameters $q(y)$ (the location parameter or threshold), $f(y)$ (the scale), and β (the shape) are estimated from the top $100p\%$ of values using an adaptation of the method from de Valk and Cai (2018); see Supplement S7.4.2 for further details. The shape parameter β and the location $q(y)$ are estimated independently of other parameters from the data. However, the estimation of the scale parameter $f(y)$ requires the estimate of β. The code is available at https://github.com/ceesfdevalk/EVTools; the tail estimates are calculated with the R function FitGW_iHilli.R.

For the estimation of the omnidirectional tails of HW and stress, $p = 0.012$ is used, based on research in de Valk and van den Brink (2023a). For stress, all parameters are estimated from SEAS5 data (and subsequently corrected for model-related bias; see Section 7.1). For HW (see Section 7.2), the shape β is estimated from SEAS5/DCSMv5 data. Subsequently, the scale $f(y)$ and location $q(y)$ are estimated from measurement data. Tab. S1 provides an overview of the data sources and settings.

S7.2 Direction-dependent tails and exceedance frequencies

First, we transform the values of HW or stress Z with distribution F to the values of a variable Y with a given distribution G, for which the exponential distribution is chosen: $Y := -\log(1 - F(Z))$ with distribution function $G(x) := P(Y \leq x) = 1 - e^{-x}$.

From the sample $Z_1, ..., Z_n$ of HW or stress Z, we can easily construct approximations of values of Y which have a deterministic distribution approximating the exponential distribution:

$$Y_i = -\log\left(\tfrac{1}{n+1}\textstyle\sum_{k=1}^{n} H(Z_k - Z_i)\right), \tag{S3}$$

with H the step function: $H(z) = 0$ if $z < 0$ and $H(z) = 1$ if $z \geq 0$. Because the empirical distribution of Y is deterministic, the dependence on wind direction is expected to be estimated more precisely from the simultaneous data of Y and wind direction θ than from the simultaneous data of Z and θ.

We work with 16 sectors $A_1, ..., A_{16}$ each $22.5°$ wide with central directions $22.5°, 45°, ...360°$. The directionality is represented by $G_j(x) := P(Y \leq x \mid \theta \in A_j)$ for $j = 1, ..., 16$. For G_j, we use again the GW approximation. Because the omnidirectional distribution of Y is exponential (and thus has a GW tail with $\beta = 1$ and $f(y) = y$), the following conditions apply:

$\beta_i \leq 1$, and if $\beta_i = 1$, then also $f_i(y) \leq y$. These conditions are imposed on the parameter estimates. To satisfy the requirement that the omnidirectional distribution of Y is exponential, we must normalise $1 - G_j(x)$. As a result, the directional frequency of exceedance $\mu_j(z)$ of z coincident with a wind direction in A_j is computed as

$$\mu_j(z) = \frac{(1 - G_j(x))P(\theta \in A_j)}{\sum_{i=1}^{16}(1 - G_i(x))P(\theta \in A_i)}\mu(z) \quad \text{where} \quad x = -\log(1 - F(z)). \tag{S4}$$

The estimates of the scale and shape of G_j are determined from the values of Y associated with directions in the $45°$ sector $\bar{A}_j$ centred at the same direction as the $22.5°$ sector A_j. This results in smoother estimates as a function of direction and higher precision (there are more data points per sector). The location of of G_j is estimated only from values with wind directions in A_j.

For stress and HW, these estimates are derived from SEAS5/DCSMv5 data with $p = 0.012$, just like the omnidirectional tail distributions. For stress, all return values are scaled afterwards by a constant (see Section 7). For HW, the analysis of directionality is partly based on measurement data: in this case, the scale and location of of G_j are estimated from measurement data, with the shape fixed at the estimate from the SEAS5/DCSMv5 data. Since the time series of simultaneous measurements of HW and wind direction are relatively short, a higher value 4.8% of p is used. The resulting estimates of direction-dependent exceedance frequencies of HW, which are still a little rough, are used to correct the previous estimates from only SEAS5/DCSMv5 data for bias. This is done by smoothing the logarithms of the ratios of the two estimates, and then applying these smoothed ratios to correct the estimates from only SEAS5/DCSMv5 data. Tab. S1 provides an overview of the data and settings used.

variable	statistic	parameter	symbol	data source	bin width for estimation	$p = e^{-y}$
stress	omni	location	$q(y)$	SEAS5	–	0.012
stress	omni	scale	$f(y)$	SEAS5	–	0.012
stress	omni	shape	β	SEAS5	–	0.012
stress	directionality	location	$q_j(y)$	SEAS5	$22.5°$	0.012
stress	directionality	scale	$f_j(y)$	SEAS5	$45°$ (overlap)	0.012
stress	directionality	shape	β_j	SEAS5	$45°$ (overlap)	0.012
HW	omni	location	$q(y)$	measurements	–	0.012
HW	omni	scale	$f(y)$	measurements	–	0.012
HW	omni	shape	β	SEAS5/DCSMv5	–	0.012
HW	directionality	location	$q_j(y)$	SEAS5/DCSMv5	$22.5°$	0.012
HW	directionality	scale	$f_j(y)$	SEAS5/DCSMv5	$45°$ (overlap)	0.012
HW	directionality	shape	β_j	SEAS5/DCSMv5	$45°$ (overlap)	0.012
HW	dir. correction	location	$q_j(y)$	measurements	$22.5°$	0.048
HW	dir. correction	scale	$f_j(y)$	measurements	$45°$ (overlap)	0.048

Table S1. Selected data and settings for the estimation of omnidirectional ("omni") tail distributions, directionality, and bias correction of directionality (dir. correction).

.

S7.3 Error estimation using the block bootstrap

In the block bootstrap (Künsch, 1989) the analysed time series are divided into blocks longer than the typical de-correlation time, and new synthetic time series of the same length as the original series are constructed by randomly drawing blocks with replacement and placing them one after the other. The entire data analysis is performed on a number of such synthetic time series, and the uncertainty in the estimates can be determined directly from the ensemble of outcomes. For the measurement data, blocks of 1 year were used to preserve the seasonality in the data. For the SEAS5/DCSMv5 model output, blocks consisting of the last 6 months of a 7-month long member of the forecast ensemble were used. Independence of simultaneous values (of which at least one is high) from different members was verified for stress; see Supplement S2. In principle, the selection of blocks of 6 month introduces a spurious inflation of the spread due to the seasonality (e.g., there may be too many or too few winter months in the resampled dataset), but this effect was shown to be negligible, due to the large size of the dataset. The number of bootstrap replicates was fixed at 100; this is sufficient for the estimation of variances.

S7.4 Tail estimators

S7.4.1 Generalized Pareto tail estimation

Generalized Pareto (GP) tails and tail shape γ (see Section 4) were estimated in two different ways: (a) from estimates of the (equivalent) GEV distribution from annual maxima (for the estimates in Section 3) and (b) from full time-series (for the tail estimates for HW at Hoek van Holland in Section 2).

In both cases, the maximum likelihood method was used to estimate the GEV (for (a), Dombry (2015)) and the GP tail (for (b), de Haan and Ferreira (2006)). For (b), the estimated tail was combined with an estimate of the extremal index (Supplement S6) to derive return values. This approach results in somewhat higher estimates of the shape γ than method (a), due to the effect of increasing dependence between successive values when these values are lower. As a result, estimates of return values for long return periods based on method (b) are somewhat higher than estimates based on method (a). Method (b) is also used for the GW tail fits (e.g. Section 5, 7 and Supplement S6).

S7.4.2 Generalized Weibull tail estimation

For fitting Generalized Weibull (GW) tails, we employ an adaptation of the estimator from de Valk and Cai (2018). The latter is introduced there as an estimator for the log-GW tail de Valk (2016a); to estimate the GW tail from a data sample, one simply skips the step of taking logarithms of the data values.

In de Valk and Cai (2018), the large-sample behaviour of this estimator was analysed in a configuration using different thresholds for shape and scale estimation. In practice, however, it proves to be more effective to use the same thresholds (see also Albert et al (2015)). In that case, the same analysis shows that errors in return value estimates are dominated by the error in the shape estimate.

Furthermore, in the notation of de Valk and Cai (2018), the estimator $\hat{\theta}_{k_n,n}$ of the shape parameter θ in their Eq. (19) is modified by replacing $\hat{\gamma}^H_{i,n}$ with the asymptotically equivalent expression

$$\left(\vartheta_{i+1,n} \frac{1}{i} \sum_{j=1}^{i} h_{\hat{\theta}_{k_n,n}}(\vartheta_{j,n}/\vartheta_{i+1,n}) \right)^{-1} \hat{\gamma}^H_{i,n}. \tag{S5}$$

This modification is motivated by replacing the first approximation of $g(\vartheta_{i+1,n})$ in Eq. (17) of de Valk and Cai (2018) with the second, more accurate, approximation. It leads to an implicit equation for the shape parameter, the solution of which is easily approximated by searching over a discrete set of values.

In practice, the outcomes of this estimator resemble those of the maximum likelihood (ML) method, but the computation takes much less time. An alternative estimator is (Albert et al, 2015).

S8 Error in skew surge predictions on the Waddenzee

The following graphs are reproduced with permission from Zijl and Laan (2021); see also Zijl et al (2022).

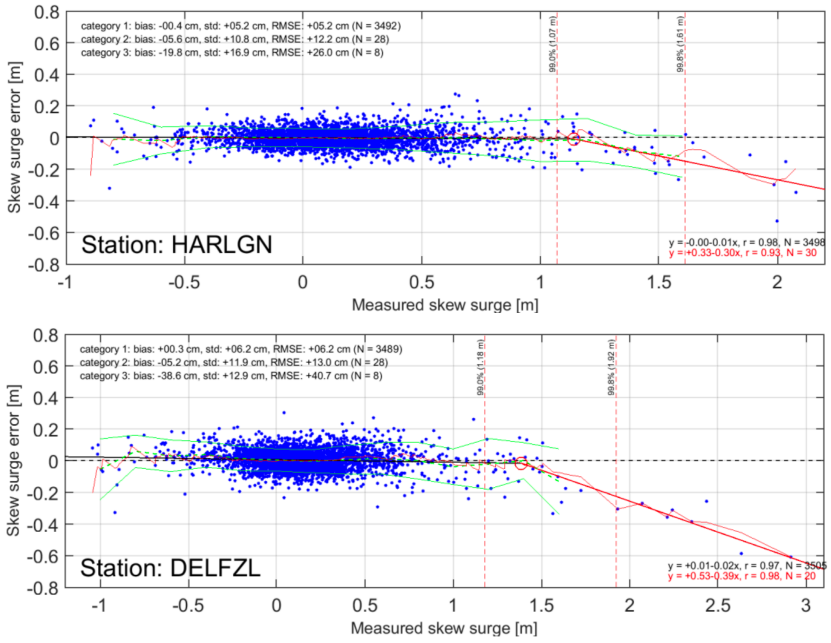

Figure S9. Error in predictions of skew surge at Harlingen (5) (top) and Delfzijl (6) (bottom) by the DCSM-FM 100m tide-surge model forced by HIRLAM output for the period 2013-2017.

References

Albert, C., Dutfoy, A., Gardes, L., Girard, S. (2020). An extreme quantile estimator for the log-generalized Weibull-tail model. *Econometrics and Statistics* **13**, 137-174.

Baart, F., Bakker, M. A. J., Van Dongeren, A., Den Heijer, C., Van Heteren, S., Smit, M. W. J., ... (2011). Using 18th century storm-surge data from the Dutch Coast to improve the confidence in flood-risk estimates. *Nat. Hazards Earth Syst. Sci* **11**(10), 2791-2801.

Blackadar, A. K. and Tennekes, H. (1968). Asymptotic similarity in neutral barotropic planetary boundary layers. *Journal of the Atmospheric Sciences* **25**(6), 1015–1020.

Curcic, M. and Haus, B.K. (2020). Revised estimates of ocean surface drag in strong winds. *Geophysical research letters* **47**(10), p.e2020GL087647.

Dillingh, D. (2013), Veranderingen in gemiddelde zeeniveaus in de Nederlandse kustwateren. *Rapport 1206182-000*, Deltares.

Dombry, C. (2015), Existence and consistency of the maximum likelihood estimators for the extreme value index within the block maxima framework. *Bernoulli* **21**(1), 420 - 436.

Ferro, C. A., Segers, J. (2003). Inference for clusters of extreme values. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **6**(2), 545–556.

Van Gelder, P. H. A. J. M. (1996). A new statistical model for extreme water levels along the Dutch coast. *Stochastic hydraulics* **96**, 243–249.

de Haan, L., Ferreira, A. (2006), *Extreme value theory - An introduction.* Springer.

Künsch, H.R. (1989), The jackknife and the bootstrap for general stationary observations. *Annals of Statistics* **17**, 1217–1241.

Leadbetter, M.R., Lindgren, G., Rootzén, H. (1983), *Extremes and Related Properties of Random Sequences and Processes*. Springer.

Ledford, A. W., Tawn, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika* **83**(1), 169-187.

Ledford, A. W., Tawn, J. A. (1997). Modelling dependence within joint tail regions. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **59**(2), 475-499.

Ledford, A. W., Tawn, J. A. (1998). Concomitant tail behaviour for extremes. *Advances in applied Probability* **30**(1), 197-215.

Pineau-Guillou, L., Ardhuin, F., Bouin, M. N., Redelsperger, J. L., Chapron, B., Bidlot, J. R., Quilfen, Y. (2018), Strong winds in a coupled wave–atmosphere model during a North Atlantic storm event: Evaluation against observations. *Quarterly Journal of the Royal Meteorological Society* **144**(711), 317–332.

Richter, D.H., Wainwright, C., Stern, D.P., Bryan, G.H. and Chavas, D. (2021). Potential low bias in high-wind drag coefficient inferred from dropsonde data in hurricanes. *Journal of the Atmospheric Sciences* **78**(7), 2339–2352.

de Valk, C. (2016), Approximation of high quantiles from intermediate quantiles. *Extremes* **19**, 661–686.

de Valk, C., Cai, J.J. (2018), A high quantile estimator based on the log-Generalised Weibull tail limit. *Econometrics and Statistics* **6**, 107–128.

de Valk, C.F., van den Brink, H.W. (2023), Comparison of tail models and data for extreme value analysis of high tide water levels along the Dutch coast. *Report WR-23-01*, KNMI, De Bilt.

de Valk, C.F., van den Brink, H.W. (2023), Update van de statistiek van extreme zeewaterstand en wind op basis van meetgegevens en modelsimulaties. *Report TR-406*, KNMI, De Bilt (in Dutch).

Zijl, F., Laan, S. (2021) Impact golfkoppeling DCSM-FM. *Memo 11206814-004-ZKS-0008*, Deltares, Delft.

Zijl, F., Zijlker, T., Laan, S., Groenenboom, J. (2022) DCSM-FM 100m: a sixth-generation model for the NW European Shelf. *Report 11208054-004-ZKS-0002*, Deltares, Delft (https://publications.deltares.nl/11208054_004_0002.pdf).

Zweers, N. C., Makin, V. K., De Vries, J. W., Burgers, G. (2012). On the influence of changes in the drag relation on surface wind speeds and storm surge forecasts. *Natural hazards*, **62**, 207-219.