Appendix A: Statistical test associated with figure 2

Let m be the number of sites considered.

Let $\{Y_1, Y_2, ...Y_i, ...Y_n\}$ be the annual maximum skew surges over n years at a given site. In the context of figure 2 we are considering annual maxima simulated by our numerical model of the shelf sea, and $n = 484$.

Let λ_j be the scale parameter diagnosed by a Gumbel fit to the n annual maxima at site j, let σ_j be the standard deviation of the n annual maxima at site j, and let

$$d_j = \lambda_j - \sigma_j \frac{\sqrt{6}}{\pi}$$

This is the departure of a point in figure 2 panel (a) from the line $x = y$. Then our test metric (call it T_Y) is the root-mean-square value of $\{d_1, d_2, ...d_j, ...d_m\}$:

$$T_Y = \sqrt{\overline{d_j^2}}$$

where the overbar indicates a mean over $j = 1, 2, ...m$ sites. This test metric is a single value. To test the statistical significance of the value of T_Y that we find when $\{Y_1, Y_2, ...Y_i, ...Y_n\}$ are the annual maximum skew surges simulated by our numerical model of the shelf sea, we repeat the test, replacing $\{Y_1, Y_2, ...Y_i, ...Y_n\}$ at each site j by $\{G_1, G_2, ...G_i, ...G_n\}$ where $\{G_1, G_2, ...G_i, ...G_n\}$ is a random sample drawn from a Gumbel distibution whose scale parameter[5] is λ_j. We do this many times (say $N = 256$ times) to create a 256-element distribution of values $\{T_{G,1}, T_{G,2}, ...T_{G,k}, ...T_{G,N}\}$, each being a value of T_G that we find when $\{G_1, G_2, ...G_i, ...G_n\}$ are "easily-made pseudo-extremes" from a Gumbel distribution, instead of 'hard-won" simulated annual maxima of unknown distribution like $\{Y_1, Y_2, ...Y_i, ...Y_n\}$. The variation in $\{T_{G,1}, T_{G,2}, ...T_{G,k}, ...T_{G,N}\}$ arises due to sampling uncertainties.

When we compare T_Y with the distribution $\{T_{G,1}, T_{G,2}, ...T_{G,k}, ...T_{G,N}\}$, we find that T_Y departs from the mean of $\{T_{G,1}, T_{G,2}, ...T_{G,k}, ...T_{G,N}\}$ by more than six times the standard deviation of $\{T_{G,1}, T_{G,2}, ...T_{G,k}, ...T_{G,N}\}$, implying that this departure is not simply an artefact of sampling, but arises from the fact that the $\{Y_1, Y_2, ...Y_i, ...Y_n\}$ are *not* Gumbel-distributed. This large departure (more than six standard deviations) might seem surprising given that figure 2 (a) shows that most points are within the 95-percent (approximately 2 standard deviation) uncertainty range of the $x = y$ line. The large departure is associated with the fact that the shape parameters of the $\{Y_1, Y_2, ...Y_i, ...Y_n\}$ are predominantly negative. We can see this expressed in figure 2 panel (a), where almost all of the scatter points lie to the right of the line $x = y$, whereas in panel (b) (which is one of the 256 examples of what happens when we replace the Y_i by G_i) the scatter points lie either side of the line.

It would be interesting to apply a similar statistical test to the scatter of points in figure S1 of the supplementary material to Woodworth et al. (2021). Assuming that, as found by Wahl et al. (2017), the shape parameters are predominantly negative,

[5] We could also employ the corresponding location parameter, but there is no need because this only introduces an offset. We simply set the location parameter to zero. Incidentally, we can generate this random sample from a uniformly-distributed random sample using the probability intergral transform, among other possible approaches.

one might expect to see the test statistic T_Y similarly outside the distribution $T_{G,k}, k = 1,2,...N$, although the shortness of the tide-gauge records might reduce the statistical significance. The bias, $\overline{d_j}$, is another alternative test statistic.

Author contributions. TH performed the tests and wrote the text. Please also see the acknowledgements.

Competing interests. The author declares no competing interests.

Acknowledgements. This work was supported by the Met Office Hadley Centre Climate Programme funded by the UK Government's Department for Business, Energy and Industrial Strategy and Department for Environment, Food and Rural Affairs. Thanks to Simon Williams and Jenny Sansom for help with the CFB data.