# Author response to reviewer 2 comments

Marvin Lorenz and Ulf Gräwe

September 8, 2023

We thank both referees, Tarmo Soomere and one anonymous reviewer, for their detailed comments and thoughts. Our responses to reviewer 2 are given below.

> This paper presents extreme sea level analysis for the Baltic Sea, based on a GETM simulation of the water elevation dynamics for the years 1979-2018. The paper is well written and addresses an important topic. However, certain aspects of the research should be clarified, particularly related to the ocean model configuration, before the manuscript can be recommended for publication.

Thank you. We answer your comments in the following.

> Major comments:
> The model configuration, presented in Section 2.1 should be elaborated. The authors should provide sufficient information about the model configuration such that the simulations can be reproduced by independent researchers. In particular, as atmospheric pressure and wind stress are key drivers for extreme water levels, used forcing methods should be presented in more detail. Which wind stress formulation was used? Why that particular one was chosen? Can it be argued that the existing and widely-used formulations are suitable for the Baltic Sea (to the best of my knowledge, many wind stress formulations are designed for global ocean simulations). These details should be added and their impact also discussed in the manuscript.
> In this study, the wind speeds have been artificially increased by 3 to 11 percent to take into account the fact that atmospheric models have a tendency to underestimate extreme wind speeds. Altering the wind stress parametrization would be another way of achieving the same, potential benefit being that one can only alter the wind stress for high wind speeds.

In GETM, the wind stress is computed with:

$$\vec{\tau} = C_D \rho_{\mathrm{air}} |\vec{u}_{10}| \vec{u}_{10},$$

with $\rho_{\mathrm{air}} = 1.25$ kg m$^{-3}$ and the 10m wind speed vector $\vec{u}_{10}$. The drag coefficient $C_D$ is computed by the formulation of Kara et al. (2000). In our barotropic simulations we do not consider temperature differences between the ocean and atmosphere. The drag coefficient reads as

$$C_D = 10^{-3}\left(0.862 + 0.088w - 0.00089w^2\right),$$

where this parameterization limits the wind speed $w = \max(2.5, \min(32.5, |\vec{u}_{10}|)$. You are right that this information is really important for the reader, thus we included this into the model description:
"The wind stress is calculated from the 10 m wind fields with

$$\vec{\tau} = C_D \rho_{\text{air}} |\vec{u}_{10}| \vec{u}_{10}, \tag{1}$$

with $\rho_{\text{air}} = 1.25$ kg m$^{-3}$ and the 10 m wind speed vector $\vec{u}_{10}$. The drag coefficient $C_D$ is computed by the formulation of Kara et al. (2000). In our barotropic simulations we do not consider temperature differences between the ocean and atmosphere. Therefore, the drag coefficient reads as

$$C_D = 10^{-3}\left(0.862 + 0.088w - 0.00089w^2\right),$$

where the wind speed is limited: $w = \max(2.5, \min(32.5, |\vec{u}_{10}|))$."
Since this formulation considers higher drag coefficients for higher wind speeds, we believe the main variability and the main reason for the necessary winds speed adjustments lie in the atmospheric datasets. Concerning your point on the validity of such formulation based on a parameterization for global ocean models for the Baltic Sea, there are parameterizations based on near shore measurements (Smith et al., 1992). However, if these different formulations perform better or worse have to be tested in the future.
We added a new paragraph to the discussion where we briefly elaborate your two raised points: "There exist many formulations, i.e. polynomial fits, of the drag coefficient which translate wind speeds into winds stresses (e.g. Kara et al., 2000; Large and Yeager, 2009). Here we have used the formulation of Kara et al. (2000). Using a different formulation, e.g. one that yields a larger drag coefficient for high wind speeds, could have improved the representation of ESLs and potentially decreased the wind factors we applied. Since there exist many formulations, each formulation would lead to different ESLs, thus introducing a source for variability. The different formulations depend on the geographic locations of the underlying data which is fitted, e.g. offshore versus onshore (Smith et al., 1992). The formulation we have used (Kara et al., 2000) is based on observations made in the Arabian Sea. One could argue that this formulation may not suitable for the Baltic Sea. However, we expect for this study that the variability of the atmospheric datasets, Figs. 9 and 10, to be much larger than the differences in ESLs due to the drag coefficient formulations. Still, the exploration of the effect of different drag coefficient formulations on ESLs in the Baltic Sea should be carried out in a future study. A calibrated drag coefficient formulation using an adjoint model could be a possible solution (Peng et al., 2013). However, depending on the wind input fields, different best fits have to be expected. Therefore, this inverse method cannot be expected to provide a one-size-fits-all solution."

> In order to obtain reliable extreme value estimates, the model should be sufficiently accurate in representing the tides, emptying and filling of the Baltic Sea basin, seiche waves, and atmospherically-driven effects - extreme SSH values are formed as a superposition of all of these. (While the tides might not play a major role in the Baltic Sea itself, they may affect the volume flux to/from the North Sea.) The manuscript focuses mostly only on the atmospheric effects. It would be good to discuss the other effects and their impact as well.

Our model does include all of the above processes except tides, which are not important in the Baltic Sea (Gräwe and Burchard, 2012). The emptying and filling of the Baltic is also atmosphere driven, thus included in the results. However, the disentanglement of these processes is not trivial as they interact with each other, e.g. seiche-surge interaction can increase ESLs by up to 10cm in the Baltic Sea (Arns et al., 2020). Also, their statistical distributions are different. Thus, the GEV- and GPD approach may not work well, as mentioned by the other reviewer. Since this point goes into the direction of a point also raised by reviewer 1, we included a paragraph on these points into the discussion: "ESLs are the superposition of many sea level processes: preconditioning, storm surges, and standing waves (seiches). For example, since preconditioning can increase the mean sea level in the Baltic Sea on weekly time scales, the total sea level during a storm surge can be increased. These two (and more) compounding mechanisms follow different statistical distributions (e.g. Suursaar and Sooäär, 2007). Our analysis neglects this point as the GEV and GPD methods assume a single statistical distribution which still led to good fits. Furthermore, the mentioned mechanisms are interacting non-linearly (Arns et al., 2020). Disentangling these two (and more) compounding processes of this ensemble remains a matter for a future study."

> Section 3 shows a comparison of extreme sea levels. To better assess the skill of the model, I recommend including a comprehensive statistical analysis of the reference model SSH in the Baltic Sea (and perhaps the Danish waters). It would then be clearer whether we are only testing the effects of different atmospheric forcings using a poor SSH model, or whether the model reproduces SSH dynamics well in general (giving more confidence that combined SSH effects can be reproduced).

We added the Root Mean Square Errors and correlation coefficients between the modelled and observed time series for each tide gauge to the Supplement (see Fig. 1 which is the new Figure S1) and referred to it in Section 3.1.:
"The time series comparison for each tide gauge station shows a good agreement with low Root Mean Square Errors (RMSE $\leq 0.1\,\mathrm{m}$) and high Pearson correlation coefficients $R \approx 0.9$, see Fig. S1."
And in the Supplement:
"In addition to the ESL comparison in the main text, we compare here the full length time series of the tide gauge stations with the different model runs. We compare the Root

Mean Square Error (RMSE),

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\eta_i^{\text{obs}} - \eta_i^{\text{mod}}\right)^2}, \tag{2}$$

and the Pearson correlation coefficient $R$,

$$R = \frac{1}{N}\sum_{i=1}^{N}\frac{(\eta_i^{\text{mod}} - \bar{\eta}^{\text{mod}})(\eta_i^{\text{obs}} - \bar{\eta}^{\text{obs}})}{\sigma^{\text{mod}}\sigma^{\text{obs}}}, \tag{3}$$

where $\eta_i$ denotes the discrete time series of the observed sea level and the modelled sea level, respectively, $\bar{\eta}$ denotes the temporal mean of the respective time series, and $\sigma$ denotes the respective standard deviation. For all simulations the correlation coefficients are all around 0.9 and most RMSEs are smaller than 0.1,m (Fig. 1). For stations in the Kattegat, the $R$-values are smaller since our simulations excluded tides which are still present in this area. Also for the tide gauges in coastal lagoons, e.g. Althagen, the correlation is much smaller since the sea level dynamics cannot be captured correctly due to the coarse resolution."

> The model uses a constant bottom roughness length value (z0 = 1 mm) for all the simulations. Is this a realistic value? How was it chosen? Was the same value used of all nesting levels? Tuning the bottom friction for North Sea-Baltic Sea simulations is not a trivial task (e.g. Kärnä et al. (2023) and Kärnä et al. (2021)) if one wishes to represent SSH dynamics well across the domain - it also affects the attenuation of seiche oscillations in the Baltic.

We have used $z_0 = 5$ mm for the North Atlantic nest and indeed $z_0 = 1$ mm for the North Sea / Baltic Sea nest. You are right that a constant bottom roughness is not the most realistic. However, past studies have intensively calibrated and validated the model setup and led to the value of 1 mm, e.g. for sea level dynamics and ESLs (Gräwe and Burchard, 2012), but also for Major Baltic Inflows (MBI, e.g. Mohrholz et al., 2015) (in a baroclinic version of this setup). Correctly reproducing the timings and strengths of MBIs is very sensitive to the choice of the bottom roughness parameter. The spatially constant $z_0$ is capable to do that. Therefore, we believe the constant value is suitable for our application. Kärnä et al. (2021) also use a constant bottom roughness of 1mm for the Baltic Sea. Nevertheless, future studies, especially going to higher and higher resolution, have to revisit this topic.

> I highly encourage the authors to share the GETM source code, input files, and also post-processing scripts for reproducibility. Quite often some data (e.g. forcings) cannot be shared due to licence restrictions, but even under such circumstances, the authors should provide links to where the data can be accessed.

We included links to the code and data now in the "code and data availability"-statement.

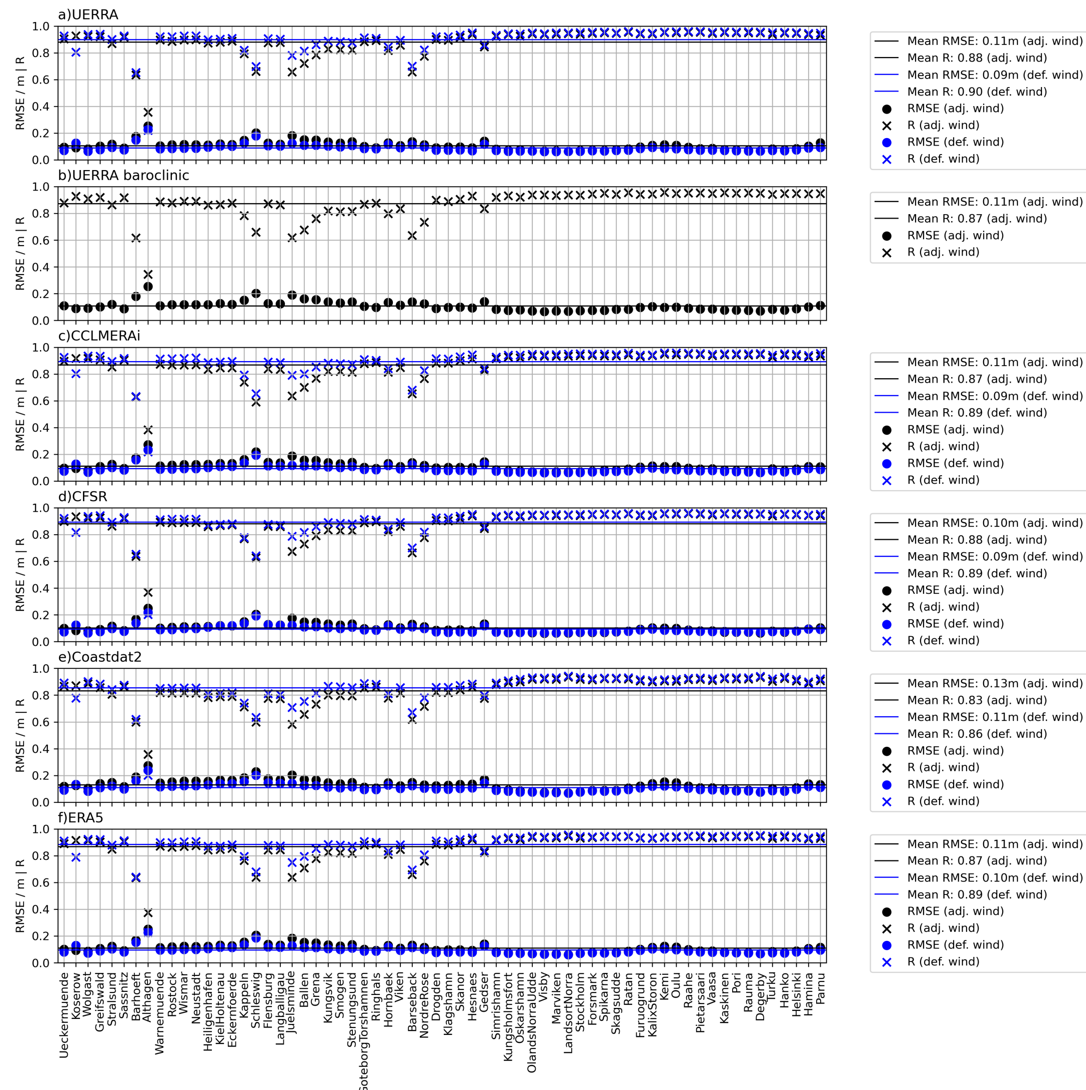


Figure 1: Comparison of the Root Mean Square Error (RMSE) and the Correlation coefficient between the different model runs for each tide gauge station. Note that the values between the adjusted wind speed simulations (black) and the default wind speed simulations (blue) are very similar. Therefore the black dots are hidden behind the blue dots.

> In the light of the given results, it seems that understanding the sensitivity of extreme sea level values to wind forcing is essential. Such sensitivities can be computed directly with adjoint models (e.g. Kärnä et al. (2023) and references therein). As such adjoint/inverse modeling could be an important asset in this research in addition to more traditional ensemble methods.

This seems indeed to be a good idea. We added this point to the discussion in two paragraphs: "Nevertheless, these would be strongly modified by the default increase in wind speeds and may lead to over- or underestimation of these with consequences for the general dynamics, e.g. stratification or the total salt import by Major Baltic inflows. One possible way of tackling this problem could be adjoint/inverse models (Errico, 1997) which can be used to calibrate coastal ocean input fields like bottom friction (Kärnä et al., 2023) or in this case wind speed factors."
"Still, the exploration of the effect of different drag coefficients formulations on ESLs in the Baltic Sea should be carried out in a future study. A calibrated drag coefficient formulation using an adjoint model could be a possible solution (Peng et al., 2013). However, depending on the wind input fields, different best fits have to be expected. Therefore, this inverse method cannot be expected to provide a one-size-fits-all solution."

> Minor comments: Which GETM version was used to run the simulations?

The version is GETM 2.5. A frozen version is now linked in the code availability statement. Now included in the main text.
"We use the General Estuarine Transport Model (GETM (version 2.5), Burchard and Bolding, 2002), a structured coastal ocean model (Klingbeil et al., 2018), to simulate the surface elevation in the Baltic Sea."

> line 98: Please elaborate. Are you only using the local atmospheric pressure information to calculate an additional offset to boundary SSH value? This would not include any atmospheric pressure-driven waves.

Only at the open boundary of the North Atlantic setup we prescribe the sea level by the inverted barometric pressure formula, i.e. the sea level is computed from the pressure above the respective water columns. You are right that we exclude waves which were generated outside the domain. We modified the sentence to: "Along its boundary, air pressure-induced water level changes are imposed using the atmospheric pressure from ERA5 to include large pressure systems from the Atlantic into the model chain (inverted barometric effect)."

> line 105: The atmospheric forcing is not the same for the North Atlantic and the Baltic Sea models. How can you ensure that SSH at the nesting boundary agree, e.g. if a low pressure system is located in a different place in ERA5? This could generate spurious strong wave fronts in the model and skew the ESL analysis.

This is true but we expect the effect of this to be very small because of following reasons: 1. The Danish Straits function as a low-pass filter thus the high-frequency shock waves would be filtered. The same reason why tides are negligible in the Western Baltic Sea. 2. The boundary is far away from the Baltic Sea and the North Sea provides enough surface area for the nested setup's forcing to overwrite these errors by their wind stresses and pressure fields, respectively. 3. ERA5 is the only dataset that spatially covers the domain completely.
We mention the first two arguments now in the main text: "This simulation prescribes boundary conditions for the one nautical mile North Sea / Baltic Sea domain (all simulations). This may introduce some inconsistencies in the sea level when using other atmospheric forcings for the North Sea / Baltic Sea nesting stage. However, we assume these errors to have little effect on Baltic Sea ESLs since the Danish Straits act as a low-pass filter. Thus, storm surges are generated inside the Baltic Sea."

> line 154: For the sake of clarity, please emphasize how you calculate ESLs from model/observation SSH. E.g. add an equation.

We rephrased these sentences and added additional information:
"For this comparison, we use the German Federal Maritime and Hydrographic Agency's definition of a storm surge in the Baltic Sea, which defines a storm surge as a sea level of more than 1 metre above mean sea level. Furthermore, we only consider events that are separated by more than 48 hours. We search for events that fulfil these criteria in the observed tide gauges and compare them with the modelled sea level for each station."

> Figure 4: I would not use a line plot here. There's no continuity between the stations (at least in the sense the figure suggests). A bar or violin plot for each station would be better.

Changed to a bar plot.

> Figure 7 caption: if you mention a) mention also b)

fixed

> line 288: typo FThe

fixed

> line 322: remove "therefore", there's no causality here

removed

> line 330: The reviewers should have access to the data/source code during the review.

We will do in the future!

# References

Arns, A., Wahl, T., Wolff, C., Vafeidis, A. T., Haigh, I. D., Woodworth, P., Niehüser, S., Jensen, J., 2020. Non-linear interaction modulates global extreme sea levels, coastal flood exposure, and impacts. Nature communications 11 (1), 1–9.

Burchard, H., Bolding, K., 2002. GETM: A General Estuarine Transport Model; Scientific Documentation. European Commission, Joint Research Centre, Institute for Environment and Sustainability.

Errico, R. M., 1997. What is an adjoint model? Bulletin of the American Meteorological Society 78 (11), 2577–2592.

Gräwe, U., Burchard, H., 2012. Storm surges in the Western Baltic Sea: the present and a possible future. Climate Dynamics 39 (1), 165–183.

Kara, A. B., Rochford, P. A., Hurlburt, H. E., 2000. Efficient and accurate bulk parameterizations of air–sea fluxes for use in general circulation models. Journal of Atmospheric and Oceanic Technology 17 (10), 1421–1438.

Kärnä, T., Ljungemyr, P., Falahat, S., Ringgaard, I., Axell, L., Korabel, V., Murawski, J., Maljutenko, I., Lindenthal, A., Jandt-Scheelke, S., Verjovkina, S., Lorkowski, I., Lagemaa, P., She, J., Tuomi, L., Nord, A., Huess, V., 2021. Nemo-nordic 2.0: operational marine forecast model for the baltic sea. Geoscientific Model Development 14 (9), 5731–5749.
URL `https://gmd.copernicus.org/articles/14/5731/2021/`

Kärnä, T., Wallwork, J. G., Kramer, S. C., 2023. Efficient optimization of a regional water elevation model with an automatically generated adjoint.

Klingbeil, K., Lemarié, F., Debreu, L., Burchard, H., 2018. The numerics of hydrostatic structured-grid coastal ocean models: state of the art and future perspectives. Ocean Modelling 125, 80–105.

Large, W., Yeager, S., 2009. The global climatology of an interannually varying air–sea flux data set. Climate dynamics 33, 341–364.

Mohrholz, V., Naumann, M., Nausch, G., Krger, S., Grwe, U., 2015. Fresh oxygen for the Baltic Sea  an exceptional saline inflow after a decade of stagnation. Journal of Marine Systems 148, 152–166.
URL `https://www.sciencedirect.com/science/article/pii/S0924796315000457`

Peng, S., Li, Y., Xie, L., 2013. Adjusting the wind stress drag coefficient in storm surge forecasting using an adjoint technique. Journal of Atmospheric and Oceanic Technology 30 (3), 590–608.

Smith, S. D., Anderson, R. J., Oost, W. A., Kraan, C., Maat, N., De Cosmo, J., Katsaros, K. B., Davidson, K. L., Bumke, K., Hasse, L., et al., 1992. Sea surface wind stress and drag coefficients: The HEXOS results. Boundary-layer meteorology 60, 109–142.

Suursaar, Ü., Sooäär, J., 2007. Decadal variations in mean and extreme sea level values along the Estonian coast of the Baltic Sea. Tellus A: Dynamic Meteorology and Oceanography 59 (2), 249–260.