

RC2:

The paper deals with tide gauge time series analysis using probabilistic method. Main contribution is the addition of an acceleration term (to address long-term record), and use of harmonic or CWT to calculate the ‘harmonics’. Extreme value analysis was done using a Copula function (Gumbel distribution). While these are worthy contributions, the presentation is particularly confusing, which makes the paper hard to follow; in addition, some key details are missing. I’ll wait for the clarification to my questions below first before proceeding to the result sections.

We thank the Reviewer for the detailed and constructive assessment of the manuscript and for pointing out the issues that require further clarification. We agree that the current presentation of the methodology may make it difficult to follow the different stages of the analysis and that some elements require a clearer description. Below, we address all the questions and comments raised by the Reviewer, explaining the data used, the individual methods, and how they are connected. If we are asked to prepare a revised version of the manuscript, we will incorporate these explanations by improving the structure of the methodology, clarifying the important details and presenting the overall research workflow more clearly.

- They used long-term tide gauge record, but it’s not until much later in the paper did I realize they already filtered out the tides. What’s the sampling frequency of the record, and how was tidal signal removed and daily anomaly calculated? Table 1 seems to show the results before the methodology was even presented.

We thank the Reviewer for this comment. The tide gauge data used in this study were obtained from the Institute of Meteorology and Water Management – National Research Institute (IMGW-PIB). The data included daily sea level values from five tide gauge stations located along the Polish Baltic coast: Swinoujscie, Kolobrzeg, Ustka, Wladyslawowo, and Gdansk. The data covered the period from 1951 to 2025. The daily sea level values were mean values averaged at 06:00 and were used in the form provided by the data provider, as metric sea level data. The analysis was performed using the prepared daily data provided by IMGW-PIB. The frequency of the time series used in the analysis was therefore one value per day. However, we do not attribute a specific sampling frequency of the original tide gauge observations to these data because the dataset used in our study already consisted of prepared daily values.

It should be noted, however, that the Baltic Sea is characterized by a relatively low astronomical tidal range compared with other factors controlling sea level variability, such as atmospheric pressure and wind forcing. Moreover, using daily-averaged values reduces the influence of short-term tidal fluctuations on the analyzed time series. Consequently, the potential effect of any uncorrected tidal components in the IMGW-PIB tide gauge data on the analysis of long-term sea level changes was considered to be

negligible. Nevertheless, this issue was recognized as a potential limitation arising from differences in data processing procedures among the individual data sources.

The procedure used in this study is also based on data and previous calculations presented in our previous study (DOI: 10.3390/rs18142398). In the revised version, we will also clarify the meaning of the term “daily anomalies” to clearly distinguish the daily tide gauge data from the sea-level anomalies derived from satellite altimetry data.

- Some important details are missing. A flow chart is missing in Section 4.2 to show the workflow. My understanding is that harmonics in eq (1) are derived from CWT first, then the trend is calculated by subtracting the harmonics from MSL? On line 130 they mentioned 2 frequencies (annual, semi-annual) but later they presented

We thank the Reviewer for this comment. The procedure is a hybrid method for analyzing changes in the amplitude of periodic components. It combines the estimation of the long-term sea level trend and acceleration with CWT analysis and the assessment of trends and acceleration in the amplitudes of selected periodic cycles.

CWT is used to identify and quantify the temporal changes in the amplitude of these components over time. It is not used to determine the trend by removing these components from the MSL time series.

In the revised version, we will clarify the order of the individual analysis steps and clearly specify the periods analyzed.

- In Section 4.3, need to provide some basic information on Gumbel to explain eq (8); would the eq need an acceleration term as well?

We thank the Reviewer for this comment. In the revised version of the manuscript, we will expand the description of the Gumbel distribution and clarify the components of Equation (8).

The acceleration term was not directly included in the extreme-value distribution equation because the acceleration term describes the long-term change in the mean sea level in the trend model, whereas the Gumbel model is used to statistically describe the distribution of extreme values.

Including the acceleration term in the extreme-value model would require a non-stationary extreme-value model, in which the location parameter of the Gumbel distribution would depend on time.

- Phase info seems missing in eq (9)?

We agree that the presentation of the phase requires further clarification. The harmonic components are characterized by both amplitude and phase.

In the revised version, we will add a clear notation for the phase parameter to the relevant equation and define its meaning.

- Section 4.4: would appreciate more description on how they did Monte Carlo sampling; for example, did they account for decorrelation scale? What's the standard deviation used in the normal distribution?

We thank the Reviewer for this comment. We agree that the Monte Carlo procedure should be described in more detail.

In this study, the deterministic part of the model includes the trend and periodic components, while the stochastic variability is represented by the residual component modeled using an AR(1) process. This means that the generated realizations retain the temporal dependence between successive residual values rather than assuming that they are independent.

In the Monte Carlo procedure, 1000 realizations of the future time series are generated. Based on these 1000 realizations, the mean trajectory and prediction intervals covering the 5th–95th percentiles are determined to represent the statistical uncertainty of the forecast. These realizations are also used to obtain the distribution of possible future values and the corresponding prediction intervals.

In the revised version, we will define more clearly the standard deviation used in the normal distribution.

At the same time, we emphasize that the forecast is a conditional probabilistic statistical forecast on statistical characteristics identified in the historical data. It should not be interpreted as a physical projection of future sea-level changes that accounts for all possible future changes in ocean dynamics, ice-mass balance, or atmospheric and oceanic circulation.

- Epsilon appears in both eq 1 and 2

We thank the Reviewer for pointing out this ambiguity. We agree that using the same symbol for different components may make the equations difficult to interpret.

In the revised version, we will use separate symbols for the residual component of the model and the innovation term of the AR(1) process, and we will define both symbols directly in the corresponding equations. This change does not affect the calculation results, but it improves the clarity of the mathematical notation and makes it easier to reproduce the procedure.

- Fig 3 etc: since both TG and SA are available at the gauges and Table 1 showed some differences between the two, how would the differences affect the interpretation?

We thank the Reviewer for this comment. The differences between TG and SA are important for the interpretation of the results because the two data sources represent different aspects of sea level.

Tide gauge data represent the local relative sea level at a specific coastal station, whereas satellite altimetry represents the absolute sea level over the open sea. Therefore, differences between these two data sources may result from local coastal processes, vertical land movements, spatial differences, bathymetry, and the different spatial representation of the two types of data.

We therefore do not automatically consider the TG–SA differences to be an error in one of the data sources. These differences can also provide information about differences in sea level variability between the coastal zone and the offshore area.

We would also like to clarify the role of Table 1. This table was prepared based on the results of our previous research (DOI: 10.3390/rs18142398). It provides a basic set of parameters and characteristics used as a reference for the present study. It does not present the final results of the new probabilistic model.

In the revised version, we will clearly indicate the source of these values and explain their role in the present study, so that readers do not interpret Table 1 as results obtained from the probabilistic model presented in the later part of the manuscript.