

Author's response to the review

Tuomi, L., Johansson, M., Twelves, A., Rantanen, M., Lagemaa, P., Kanarik, H., Särkkä, J., Raudsepp, U., and Westerlund, A.: Record low sea levels in the Bothnian Bay in November 2024, State Planet Discuss. [preprint], <https://doi.org/10.5194/sp-2025-12>, in review, 2025.

We would like to thank the Reviewer for constructive comments and suggestions to improve our manuscript. We have revised our manuscript accordingly. Please find below our detailed answers to the questions and comments. The referee comments are shown with italic font and our replies with plain blue text. All changes are visible in the revised manuscript.

RC2, Anonymous Referee #2, 12 Nov 2025

This paper explores the accuracy of the Finnish Baltic Sea Monitoring and Forecasting Centre native model (named NRT physical system) in reproducing observed sea levels during a severe windstorm in November 2024. Model sea surface heights are compared against nearby tide gauge observations.

The manuscript assesses the NRT physical system model accuracy under different forecasting lead times (12-24 hrs / 36-48 hrs/ 60-72 hrs & 84-96 hrs), with the shortest lead times providing higher forecast accuracy. The authors found that the model tends to underestimate highs, and lows, thus not fully reproducing the extremes, particularly in the longer range forecasts. This is a known limitation of this version of the model, from prior literature.

With the negative storm surge being recorded in high resolution by the tide gauges, the return period of this record-breaking sea level low is estimated, using measurements from Kemi tide gauge (in the context of the longer 1977-2024 record). Land uplift and sea level changes over the period are known and were used to detrend the data. Return period estimates are demonstrably more accurate if observed extremes, like this storm, are included in the analysis.

We thank the Reviewer for constructive comments and suggestions to improve our manuscript. We have added time-series of observed annual minima from the Kemi tide gauge including also the detrended data in Figure 3. More detailed information about the general forecast accuracy of the BAL MFC NRT system has also been added to the text. Together with information about the FMI's warnings issued for low sea level events and how the accuracy of the forecasts fulfilled the criteria for those. We have also reformulated Table 1 to give better overview of the products and datasets used in this paper. Please also see our detailed answers below.

General comments and scientific issues:

This manuscript describing the modelling of a uniquely severe storm at the northern extremity of the Baltic Sea, in the Bothnian Bay and Gulf of Finland in November 2024,

reads as a technical report of the event - the purpose of which is validating the Baltic Monitoring and Forecasting Center's Near Real Time model for extreme events.

The topic of representing extreme sea levels due to storms should be of great interest for modellers, and is a topic of current interest in our changing climate. The goals of the paper are neatly set out and indeed addressed. The methodology and results are clear and well presented, although Table 1 may require some improvements to the headers, and description, for clarity (e.g. CMEMS is not defined in the report. The Product ID & Type descriptions are a little confusing).

We have made some clarifications for Table 1 to better describe the datasets used in this paper. We have also checked that all the acronyms have been explained either in the Table or in the text.

Also the role of the Automatic Weather Stations (lines 95-103) in the procedure may need clarifying. Data appears to be accessible for experiments to be reproduced.

The Automatic Weather Station (AWS) data was used in this study to give an overview of the measured wind conditions during the storm event. The 1-hourly data we have used in this study is available from FMI's open data server. We have added Data availability section to the manuscript, which describes in more detail how to reach this data and also other datasets used in the manuscript.

Results do certainly support the interpretations and conclusions. The entire manuscript is well structured and findings are concisely presented. As such it is a nice read.

However, the authors may wish to consider posing a scientific question or hypothesis to elevate the motivation of this manuscript for the readership. The innovation, or novelty, or new conclusions, presented by this paper was not clearly stated. I imagine few extreme negative storm surges are captured for model calibration and development, therefore a discussion of why the model had difficulty, and the things that could be considered to improve the model to better represent and forecast such events (even to longer lead times) would be an interesting contribution to the scientific knowledge. As such, it may benefit the manuscript if the discussion and conclusions sections are reworked to provide new insights, or discuss the unique problems of modelling such extreme events, alongside the unique tide gauge observation data and unique storm event.

We have added further information and discussion on the predictive capabilities of the NRT system in producing SSH. We have also clarified how forecast accuracy varies with lead time, distinguishing between contributions from the BAL MFC model setup and atmospheric forcing. To provide context for the reported accuracy metrics, we added discussion about the NRT system's ability to deliver sufficient information for issuing weather warnings.

There are a number of abbreviations and initialisms presented in the report, and as such, it may also be beneficial to consider an end of report Glossary for reference.

We have checked that revised version of the manuscript includes explanation for all the abbreviations used in the text, figures and tables.