Response to reviewer

GENERAL COMMENTS

This manuscript presents a hybrid machine-learning (ML) framework for short-term storm-surge nowcasting in the southern North Sea. Rather than relying exclusively on observations or numerical modelling, the authors combine three complementary sources of information: tide-gauge observations from 22 stations, ERA5 atmospheric forcing, and storm-surge simulations from the operational COHERENS hydrodynamic model. The implemented ML emulators predict surge residuals at 10-minute intervals using a 12-hour input history and initially generate forecasts for the subsequent two hours. Longer forecast horizons, extending up to 12 hours, are obtained recursively.

The results indicate that combining numerical-model outputs with observations can provide more accurate and operationally useful nowcasts than using either source of information independently. Overall, the study lies at the intersection of operational coastal forecasting, hybrid modelling, scientific machine learning, and coastal-hazard prediction.

The manuscript addresses an important operational problem through an interesting hybrid modelling strategy. The work is generally well written, technically sound, and potentially suitable for publication. However, in its current form, several methodological choices would benefit from further justification, and the scientific contribution could be demonstrated more convincingly.

In my view, the principal limitation is not the ML methodology itself, but rather the relatively limited scientific analysis of why the hybrid framework improves the predictions, under which conditions these improvements occur, and to what extent the proposed approach can be expected to generalize beyond the present application. Addressing these aspects, particularly in the Discussion, would substantially strengthen the manuscript.

Answer:

We sincerely thank you for the thorough and constructive assessment of our manuscript. We greatly appreciate the time and expertise invested in evaluating our work. Your observation and suggestions have helped us identify several opportunities to improve both methodological clarity and scientific rigor of the manuscript. We have carefully considered each comment and our responses to the individual comments detail the methodological and scientific enhancements that will be implemented in the revised manuscript.

SPECIFIC COMMENTS

L03–04:

“while purely data-driven models suffer from the limited duration of tide-gauge records.”

This statement may give the impression that data-driven models can only be trained using observations. I suggest slightly reframing the problem statement to acknowledge that outputs from numerical models can also provide training targets for ML emulators, particularly where long observational records are unavailable.

Answer:

Thanks for your recommendation and we agree that we should acknowledge that outputs from numerical models can also provide training targets for ML. Hereby, we will rephrase the sentence as *"Operational hydrodynamic models remain too computationally demanding to deliver on-demand, high-frequency forecasts at the required resolution, purely data-driven models suffer from the limited duration of tide-gauge records while numerical model simulation can be helpful for providing the high-quality ML training inputs."* in the revision.

L10–11:

"… and modestly better than the pure data-driven baseline (0.061 m; 0.043–0.100 m)."

The meaning of "pure data-driven baseline" was not immediately clear to me. Does this refer to an emulator trained using only observations, or to a model that excludes COHERENS information? I suggest defining this terminology explicitly when it is first introduced.

Answer:

We agree with the reviewer, and your comment refers to the abstract, which is by default highly condensed. The definitions of all the concepts are carefully introduced in the core text. However, the abstract shall be adapted to clarify the meaning of this sentence at the first reading. We agree that the abstract shall be totally rewritten to become self-contained. It will clearly express the objectives of the article, the tested approaches and define the different baselines in our revision.

L13:

"the COHERENS baseline."

Here, COHERENS is described as the baseline, whereas the preceding lines also refer to a "pure data-driven baseline." Please consider distinguishing these baselines more explicitly and consistently throughout the manuscript.

Answer:

We agree with the reviewer, and we will revisit the abstract to make these concepts clearer at the first reading.

L14–15:

"The approach should generalize to other coastal regions with comparable observational coverage."

This statement appears stronger than can be demonstrated from the experiments presented. The study investigates a single geographical region, and no transfer-learning, cross-region, or leave-one-station-out experiments are presented. I therefore suggest softening this conclusion or clearly presenting it as a hypothesis for future investigation rather than a demonstrated property of the framework.

Answer:

We agree with the reviewer. As this is an important argument in the abstract, then the sentence must correctly be introduced and supported in the discussion and conclusion. We will soften this conclusion in the revision.

L38–39:

"This article investigates how data-driven models and artificial intelligence can improve this final, yet crucial, stage of the decision-making process."
This sentence appears to express an important objective of the study. I wonder whether it would be more effective if positioned later in the Introduction, together with the specific objectives and contributions of the manuscript.

Answer:

In lines 32-39, our intention was to introduce the decision-making gap that the proposed nowcasting tool aims to address. Bridging this gap constitutes the overarching objective of the exploratory work presented in this manuscript. In lines 38-39, we simply formulate the hypothesis that AI/ML could provide a suitable methodological framework for the development of such a nowcasting tool. This naturally leads into the review of the state of the art presented in lines 40-59. Following this review, lines 56-65 define the specific objectives of the study.

We believe that this structure provides a logical progression from the identification of the problem, through the assessment of existing knowledge, to the formulation of the study objectives. Nevertheless, if it improves the readability and flow of the introduction, we are willing to move the statement of the overarching objective to a position after the state-of-the-art review.

L43–50:

Several references concerning ML approaches for storm-surge modelling are introduced here. However, the discussion does not sufficiently cover previous data-driven and machine-learning storm-surge applications in the North Sea or closely related areas. This makes it difficult to fully assess the regional and methodological novelty of the present work. I encourage the authors to consider relevant studies such as:

- Krieger et al. (2025): https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024GL111558
- Schaffer et al. (2025): https://doi.org/10.5194/nhess-25-2081-2025
- Royston et al. (2013): https://www.sciencedirect.com/science/article/pii/S0165011412004289
- Kristensen et al. (2026): https://arxiv.org/html/2601.02090

Including these and other relevant regional studies would help place the proposed methodology within the existing North Sea literature and clarify more precisely how the present contribution differs from previous work.

Answer:

We agree that the four references suggested by the reviewer would strengthen our description of the current state of the art, and we will incorporate them into the revised manuscript. For example, the corresponding paragraph could be revised as follows:

"Several studies have recently explored the application of machine learning (ML) techniques for storm-surge prediction in the North Sea region. Royston et al. (2013) proposed a linguistic decision-tree approach that uses historical water-level and meteorological data to predict storm surges in the Thames Estuary up to 8 hours in advance, achieving an accuracy of approximately 0.1 m. Krieger et al. (2025) demonstrated that neural networks can translate large-scale atmospheric pressure forecasts into local predictions of storm-surge conditions in the German Bight, with the highest skill obtained when predictions are averaged over multiple years rather than applied to individual years. Schaffer et al. (2025) used a multiple linear regression model based solely on 10 m effective wind speed to predict storm surges in the German Bight, achieving an F1 score of 0.49 for storm surges exceeding 250 cm. More recently, Kristensen et al. (2026) developed a graph neural network (Flo) for predicting storm-surge water levels in the Norwegian and Barents Seas using atmospheric forcing data. Trained on 16 years of observations, the model achieved a root mean square error (RMSE) of 12.4 cm. Although these studies demonstrate the potential of data-driven approaches, most models have been developed for specific geographic locations and require large training datasets to capture complex relationships, particularly in the North Sea region. Their applicability can therefore be limited when observations are scarce or when only short historical records are available, as is often the case for extreme storm-surge events. Furthermore, model generalization may be compromised when training datasets are geographically restricted or contain systematic biases. Incorporating physical knowledge into data-driven frameworks offers a promising avenue to address these limitations."

- ***Krieger, D., Weisse, R., Baehr, J., & Borchert, L. F. (2025). Machine learning-driven skillful decadal predictions of German Bight storm surges. Geophysical Research Letters, 52(4), e2024GL111558.***
- ***Schaffer, L., Boesch, A., Baehr, J., & Kruschke, T. (2025). Development of a wind-based storm surge model for the German Bight. Natural Hazards and Earth System Sciences, 25(6), 2081-2096.***
- ***Royston, S., Lawry, J., & Horsburgh, K. (2013). A linguistic decision tree approach to predicting storm surge. Fuzzy Sets and Systems, 215, 90-111.***
- ***Kristensen, N. M., Matuszak, M., Tedesco, P., Kullmann, I. K. B., & Röhrs, J. (2026). Flo: A data-driven limited-area storm surge model. arXiv preprint arXiv:2601.02090.***

L51–52:

"*Neural networks typically require large training datasets to learn complex feature relationships, which is difficult when observations are scarce or only short historical records are available*."
This is a reasonable point. However, I suggest also acknowledging that numerical-model simulations can provide training targets for ML emulators when sufficiently long observational records are unavailable. Observations are generally preferable when the objective is to reproduce measured conditions, since an emulator trained exclusively on numerical-model output may also learn systematic biases present in the parent numerical model. Nevertheless, numerical simulations can provide valuable training data where observational coverage is limited.

Answer:

We fully agree with the reviewer's comment. The sentence in question was intended as a transition to the discussion of Physics-Informed Neural Networks (PINNs) presented in the state-of-the-art review. However, we acknowledge that this intention was not sufficiently clear. To avoid any ambiguity, we propose to remove this sentence from the revised manuscript.

In addition, and in response to comments from both Reviewers #1 and #2, we will add a new paragraph to the Methodology section explaining the rationale behind the selection of the training period used in this study.

L59–60:

"*In this paper, rather than replacing the physics with a machine-learning surrogate, we exploit the physics at the data level*."
Does "the physics" refer specifically to information provided by the physics-based COHERENS model? If so, I suggest stating this more explicitly. The current wording is somewhat abstract and could be interpreted in different ways.

Answer:

The reviewer has correctly understood our intention. We note this valuable comment and will make express our statement in a more explicit way in the revision.

L64–65:

"(iii) *we carry out a systematic comparison against both the operational COHERENS hydrodynamic baseline and a pure data-driven LSTM+Transformer baseline*."
This would be an appropriate place to explicitly define what is meant by "pure data-driven," particularly which predictors are included or excluded relative to the hybrid configuration.

Answer:

We agree with the reviewer's observation. The text in lines 56-65 should be revised to more clearly articulate the specific objectives of the study and to provide a stronger rationale for the different approaches investigated. This revision will help readers better understand the motivation behind the methodological choices considered in this work. We will implement this in the revised manuscript.

Introduction:

More generally, I suggest revisiting the flow of the final three paragraphs. A clearer progression from (i) existing approaches and limitations, to (ii) the research gap, and finally to (iii) the objectives and specific contributions of the present study would strengthen the motivation.

Answer:

This comment just reiterates the comments previously stated for lines 38-65. We agree with all these comments. In the revised manuscript, these 3 paragraphs will be significantly improved.

L84–85:

"*flagged values were discarded and replaced by linear interpolation when the gap was shorter than 2 hours (12 observations)*."
Could the authors provide additional information on the extent of this preprocessing? For example, it would be useful to report the percentage of observations that were interpolated, the amount of missing data at each station, and whether interpolation affected periods containing extreme surge events. This information would help assess the potential influence of preprocessing on model training and evaluation.

Answer:

This is a valuable suggestion. In the revised manuscript, we intend to add the following table.

Table 1 Gauge stations with the corresponding percentages of flagged dataset during year 2019 -– 2025. When flagged data cover periods shorter than 2 hours, the missing surge data have been replaced with linear interpolation.

Station ID	Station name	Percentage of flagged dataset						
		2019	2020	2021	2022	2023	2024	2025
1	Dunkerque	0	0.04%	1.53%	0.71%	2.45%	0.82%	0
2	Nieuwpoort	0	0.22%	0.09%	0.14%	0	4.09%	1.73%
3	Oostende	0	0.05%	0.09%	0.04%	0	0.03%	0.03%
4	Zeebrugge leopoldii dam	0	0.06%	0.20%	0.05%	0	0.05%	0.05%
5	ZeebruggeWielingen dok	0	0.06%	0.06%	0.05%	0	4.76%	0.22%
6	VlakteVdRaan	0	0	0	0	0	0	0
7	Westkapelle	0	0	0	0	0	0	0
8	RoompotBuiten	0	0	0	0	0	0	0
9	Brouwershavensegat8	0	0	0	0	0	0	0
10	Haringvliet10	0	0	0	0	0	0	0
11	Hoek van holland	0	0	0	0	0	0	0
12	Scheveningen	0	0	0	0	0	0	0.20%
13	Ijmondstroompaal	0	0.01%	2.43%	0	0	0	4.14%
14	Ijmuiden	0	0	0	0	0	0	0
15	Den helder	0	0	0	0	0	0	0

Fig. 1:

I understand that the M2 cotidal and cophase information is shown because of the importance of tidal dynamics in the study area. However, its relevance to the subsequent analysis is not entirely clear, particularly because it does not appear to be discussed further. It may be more useful to improve the visibility and identification of the tide-gauge locations, for example by adding station IDs or labels. This would also facilitate interpretation of later references to specific stations; for instance, Oostende harbour is mentioned at L87 but cannot be readily identified in Fig. 1.

Answer:

In the revised manuscript, the map of Fig 1. will be redrawn in such a way as the reader will be able to locate the different stations with IDs (with counter-clock order) that are shown in the figure (see the following example). In table 1 of appendix A, we will show the station IDs and names in column 1 and 2 to enable readers to find the stations correspondingly.

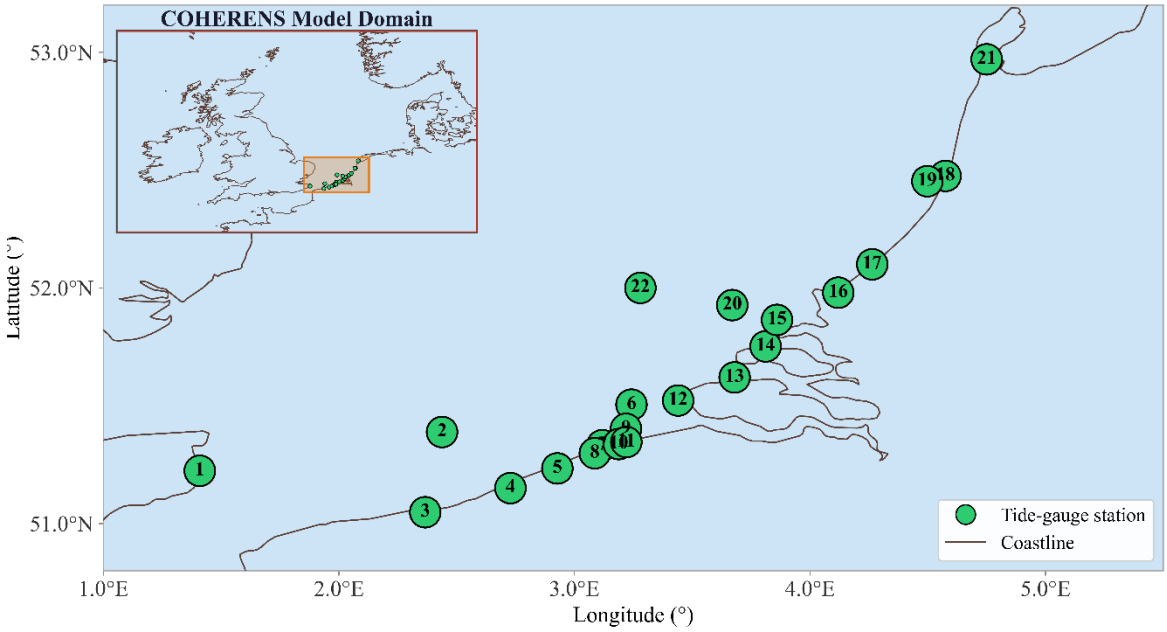


Figure 1. Distribution of the tide gauge stations (green dots) over the southern bight of the North Sea

Section 2.3 – Methodology:

In general, additional information regarding hyperparameter selection would improve the reproducibility of the study. Several architectural choices appear to be largely empirical, including the use of six Transformer layers, three LSTM layers, the selected hidden dimensions, and the number of training epochs. Please clarify whether a systematic hyperparameter search or sensitivity analysis was conducted and, if so, describe the procedure and selection criterion. The choice of a 12-hour input history also requires further justification. It would be useful to know whether alternative temporal windows were evaluated.

Answer:

This is an important remark from the reviewer.

Although we did not perform an exhaustive grid search, we conducted targeted ablation studies to validate key architectural choices. Specifically, we varied the number of Transformer layers from 5-8 and LSTM layers from 2–4 on our LSTM+Transformer model configuration. Regarding the hidden dimensions, we also tried a few, including dimensions of [128,64,32], [256,128,64], [512, 128, 64] and [1024, 512, 256]. Apart from it, we also tried to vary the learning rate with the number of 0.00001,0.00002,0.00005, 0.002,0.005 as well as the dropout rate with 0.05 and 0.2. The error metrics of corresponding experiments based on the validation dataset in the training procedure are summarized in the following table.

Table 2. Error metrics of hybrid LSTM +Transformer by using various hyperparameters

		RMSE (m)	R^2 (surge residual)	R^2 (total height)	Bias (m)
Baseline (6 Transformer layers, 3 LSTM layers, Learning rate of 0.0001, hidden dimensions [512, 256, 128])		**0.059**	**0.935**	**0.996**	**0.012**
Transformer layers	5	0.065	0.928	0.994	-0.013
	7	0.058	0.937	0.996	0.022
	8	0.066	0.918	0.995	0.026
LSTM layers	2	0.066	0.924	0.993	-0.003
	4	0.062	0.933	0.995	-0.019
drop out rate	0.05	0.057	0.942	0.996	-0.004
	0.2	0.059	0.938	0.995	-0.003
Learning rate	0.00002	0.057	0.941	0.996	0.015
	0.00005	0.054	0.946	0.996	-0.002
	0.0002	0.060	0.935	0.995	0.004
	0.0005	0.083	0.885	0.990	-0.005
Hidden dimension	[256,128,64]	0.076	0.902	0.992	-0.012
	[512, 128, 64]	0.057	0.938	0.996	0.010
	[1024, 512, 256]	0.072	0.913	0.993	-0.003

It can be shown that by varying the hyperparameters slightly, the validation performance doesn't change dramatically (with the RMSE with respect to the baseline hyperparameter setting is less than 0.01 m). This implies that the chosen values are in a stable region, which is reasonable for later application. The training epochs were determined by early stopping with a patience of 11 based on validation loss, making that choice principled rather than purely empirical. We acknowledge that a full hyperparameter sweep was not performed due to computational constraints.

For the choice of a 12-hour input history, we selected it as the baseline and tested the 6-hour interval in section 3.5.1, and a degradation of performance is observed by using 6-hour look-back table rather than 12-hour look-back table. That's why we propose the 12-hour interval in the paper. When we further increase the look-back window, this will further increase our training period as well as data load. Hereby, we did not do it at this moment, and we will test longer look-back window (18 or 24-hour) when the resource is approachable.

In the revised manuscript, we will add a specific sub-section with these elements of information.

L117:

"*we augment the gauge observations with COHERENS*."

This appears to be a central methodological aspect of the study and possibly one of its main contributions. I therefore suggest providing a more explicit description of what is meant by "augmentation" and how the COHERENS information is incorporated into the ML pipeline. At this point in the manuscript, it was not sufficiently clear to me how the observational and numerical-model information are combined.

Answer:

In the revised manuscript, we will provide a clearer description of the augmentation procedure.

Augmenting the gauge observations with COHERENS output means increasing the dimensionality of the input vector by appending the storm-surge values simulated at the 102 COHERENS grid cells located within the study area. Thus, for each time step, the model receives both the gauge observations and the corresponding COHERENS-simulated surge values as input features.

COHERENS information that is added in the input vector provides a geographical description of the surge waves in the domain, what is hardly available in the tide gauge time series. The resulting input vector therefore contains much more physics and therefore could help the model to better predict the short term evolution of the system, considering the atmospheric forcings.

L131–132:

"*This enables the Transformer to handle long time series more naturally than long short-term memory (LSTM) networks, which must process sequences in order, and allows for efficient parallelisation during training*."

This is an interesting methodological statement. Please provide an appropriate reference supporting this comparison between Transformer and recurrent architectures.

Answer:

Several key studies support the use of Transformer-based models for time-series analysis, highlighting their ability to capture long-term temporal dependencies. We will add these references in the revision.

- ***Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30.***
- ***Wen, Q., Zhou, T., Zhang, C., Chen, W., Ma, Z., Yan, J., & Sun, L. (2023). Transformers in time series: A survey. In Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence (pp. 6778–6786).***
- ***Zhao, J., Chu, F., Xie, L., Che, Y., Wu, Y., & Burke, A. F. (2026). A survey of transformer networks for time series forecasting. Computer Science Review, 60, 100883.***

L137–138:

"*F = 583 input features: 459 atmospheric-forcing values, 102 COHERENS surge values, and 22 gauge observations*."

The model uses a relatively large number of input features (583), yet no dimensionality-reduction or feature-redundancy analysis appears to be performed. Please provide some justification for retaining the complete feature set and, if possible, discuss whether redundancy or collinearity among the predictors could influence model performance.

Answer:

The number of input features (583) is relatively large. However, we retain the forcings and COHERENS simulation for all the grids as the storm surge is governed by spatially distributed hydrodynamic processes and we would like to make sure all the input feature domain can include most of the information affects the storm-surge. In specific, our input features come from physical and meteorological variables, each feature carries domain-meaningful information and removing them a priori without evidence of redundancy would be an arbitrary choice. This is similar to Hermans et al (2025) and Tiggeloven et al (2021) which use the box of 5° by 5° (20 by 20 grid cells) and a centered box of 1.25° around the station (i.e. 5 × 5 cells) respectively under the assumption of remote winds and sea-level pressure to be relevant with the storm-surge. " Similarly, Kristensen et al (2026) also use the forcing variables, 10-meter winds as x- and y-components (U10m and V10m) and mean sea level pressure (MSLP) from NORA3. Moreover, Transformers are inherently designed to handle high- dimensional inputs by the self-attention mechanism which learns to weight features by relevance, effectively performing implicit feature selection internally. Forcing a hard dimensionality reduction beforehand would discard information that the model itself could learn to ignore or down-weight.

- ***Hermans, T. H., Ben Hammouda, C., Treu, S., et al.: Computing extreme storm surges in Europe using neural networks, Natural Hazards and Earth System Sciences, 25, 4593–4612, 2025.***
- ***Tiggeloven, T., Couasnon, A., van Straaten, C., et al.: Exploring deep learning capabilities for surge predictions in coastal areas, Scientific Reports, 11, 17 224, 2021.***
- ***Kristensen, N. M., Matuszak, M., Tedesco, P., Kullmann, I. K. B., & Röhrs, J. (2026). Flo: A data-driven limited-area storm surge model. arXiv preprint arXiv:2601.02090.***

L173–175:

"*This dataset was then split 80/20 into training and validation sets. The full year 2019, which is independent of the training period, was held out as the test set to assess the model's robustness on unseen data*."

Although the test set is independent of the training data, training the model using years chronologically later than the test year is somewhat unusual for a forecasting application. I recommend discussing the rationale for this choice explicitly. In particular, readers may question whether the experimental design adequately represents the intended operational use case, even if no direct temporal leakage occurs.

Answer:

Thanks for your comment. When designing our workflow, we planned to train the dataset for 5 years (2020-2024) and then do the test with the following year-2025. However, at the time when we trained the model, full time series of tide gauge datasets for year 2025 at these 22 stations were not available to download. This is the reason why we selected year 2019 for validation in our manuscript. We think using 2019 datasets to evaluate the model performance is fair and does not weaken the conclusion of our manuscript as 2019 is also an independent year and this dataset is not being used in the training.

L205:

"*All metrics are computed on the absolute sea level (surge residual plus astronomical tide)*."

Could the authors please justify this choice? Since the ML models are designed to predict surge residuals, it would be useful to explain why performance is ultimately assessed using absolute sea level and whether the conclusions differ when metrics are calculated directly for the surge residual.

Answer:

As in *Hermans et al. (2025)*, we chose to present validation metrics only for the absolute sea level height, as this is the most relevant parameter for assessing coastal flooding risk. However, we agree with the reviewer that presenting validation metrics for both surge and absolute sea level height would provide a more comprehensive evaluation. As also noted by Reviewer 2, the RMSE and bias for surge and absolute sea level height must be nearly identical. This is because the errors in absolute sea level height are primarily driven by errors in the surge estimates, whereas the error on astronomical tidal component contributes negligibly to the overall prediction error. Consequently, the main difference between the two evaluations is expected to be reflected in the coefficient of determination (R^2). In the revised manuscript, we therefore suggest reporting four metrics: RMSE, bias, and R^2 for surge, together with R^2 for total sea level height. Table 2 hereabove provides an example on how these metrics could be reported.

L220–221:
"*We define an extreme event as a period during which the surge residual exceeds +1.0 m or falls below −1.0 m*."

Please provide further justification for the ±1.0 m threshold. For example, is it based on a percentile of the local surge distribution, previous literature, or an operational threshold used by relevant

authorities in the study area? Establishing the physical or operational basis of this definition would strengthen the extreme-event analysis.

Answer:

The ±1 m threshold was originally defined in consultation with an expert in operational oceanography. We acknowledge that this choice is somewhat arbitrary and, in the revised manuscript, we may replace it with a data-driven criterion. For example, the 99th percentile of surge values over the 2020-2024 reference period corresponds to thresholds of -0.52 m and +0.75 m for negative and positive surges, respectively.

Section 3.5 – Influence of model setting:

In my view, this analysis is more closely related to model calibration or methodological sensitivity than to the main results. I suggest considering whether it would be more appropriately placed in the Methodology section or presented before the principal performance results, since the selected model configuration presumably informs the subsequent analyses.

Answer:

The section 3.5.2 can be moved to section 2.4. In addition, we already confirmed that section 2.4 shall be expended to present the hyperparameter tuning in our revision.

Section 4 – Discussion:

I consider the Discussion to be the section that would benefit most from further development. At present, substantial parts reiterate findings already presented in the Results. I encourage the authors to expand the scientific interpretation of the results. In particular, the manuscript would benefit from a deeper discussion of:

- The physical interpretation of why incorporating COHERENS information improves the ML predictions.
- The operational implications of the reported improvements.
- Limitations of the proposed framework and data configuration.
- Geographical and temporal transferability.
- Sources of uncertainty and their propagation through the hybrid system.
- Directions for future research.

Such discussion would help move the manuscript beyond a performance comparison and provide greater insight into when and why the proposed hybrid framework is expected to be beneficial.

Answer:

We warmly thank reviewer for these suggestions and we will reorganize our discussion to address all these points in the revision.

TECHNICAL CORRECTIONS

L01:

"*so accurate short-term predictions*". "so" sounds somewhat informal in this context. Please consider alternative wording.

Answer:
This correction will be incorporated into the revised manuscript.

L114–115:

"*For example, Naeini and Snaiki (Naeini and Snaiki, 2024)*". This should be revised to "For example, Naeini and Snaiki (2024)."

Answer:
This correction will be incorporated into the revised manuscript.

Fig. 5:

"COHERNS" should be corrected to "COHERENS."

Answer:
This correction will be incorporated into the revised manuscript.

L243:

"*a understating*" should be corrected to "an understating."

Answer:
This correction will be incorporated into the revised manuscript.

L380:

"*infomation*" should be corrected to "information."

Answer:
This correction will be incorporated into the revised manuscript.