Responses to Reviewer 2

The manuscript investigates the capability of high-resolution regional coupled models to predict meteotsunamis, focusing on the event that occurred in June 2022 and affected Ireland, the UK, and France. The event generated waves of up to 1 meter and caused substantial damage to coastal community.

Study used two different high resolution regional climate models, the Met Office's UKC4 - a coupled atmosphere–ocean–wave system, and Météo-France's AROBASE - a coupled atmosphere–ocean system.

This study is significant because it demonstrates that a kilometer-scale, high-frequency coupled regional system can successfully forecast meteotsunamis. However forecast is shown only for the UK4 model while AROBASE only shows the representation of the meteotsunamis and therefore the manuscript requires a major revision to be accepted for publication.

Moreover, at times it gets hard to follow due to poor referencing of the figures and the figures themselves are not well explained in the captions.

Thank you for taking the time to review our paper, it is highly appreciated and we believe the quality of the manuscript is greatly improved thanks to your comments.

Major comments

1.) Since the manuscript title "Meteotsunami prediction in km-scale regional systems coupled at high frequency" is about meteotsunami prediction in km-scale regional systems (plural) the concerning part is why AROBASE forecast is not shown for 1 and 3 day prediction as UK4 model such as for Milford Heaven station. This seems to be a part of the manuscript (Fig. 19) on which conclusions about forecast for 1 and 3 day are based.

Thank you for your comment and for raising this important point.

We acknowledge that, given the original title, it is reasonable to expect a comparison of forecasts (including 1- and 3-day lead times) for both modelling systems. However, in the first part of the manuscript comparing the best possible deterministic forecast, our results show that the AROBASE system has very limited skill in capturing this event. As a result, extending the analysis to ensemble forecasts (e.g., 1- and 3-day predictions) would provide limited additional scientific insight, as the signal is not meaningfully represented. Additionally, ensemble forecasting capability is not yet available in the AROBASE system.

For this reason, we chose to focus the forecast evaluation on the UKC4 system, where the meteotsunami signal is sufficiently well captured to allow a meaningful assessment of predictability at different lead times.

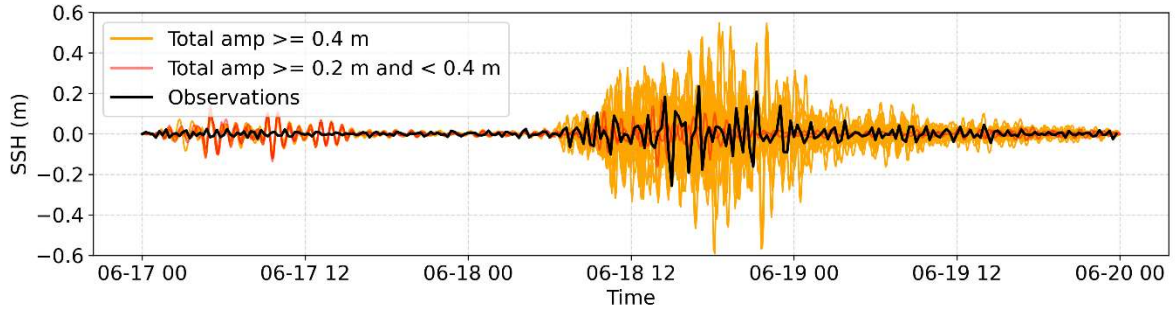
To better reflect the actual scope and balance of the manuscript, we have revised the title to: ***"Representation of meteotsunamis in km-scale regional simulations coupled at high frequency."*** This change clarifies that the study primarily evaluates system performance rather than providing a comparison of ensemble forecasts for each model.

We have also clarified this point in the manuscript in the ensembles section, explicitly stating the limitations of the AROBASE system and the rationale for not extending the ensemble forecast analysis to it. Texts reads: *"At the same time, since AROBASE did not exhibit satisfactory performance in deterministic mode, its evaluation in ensemble mode is not presented."*

2.) There is a lack of uncertainty estimation for forecast. This should be included in the Fig. 19.

Thank you, we added a new section to the manuscript, where we additionally present the percentage of members able to capture various amplitudes of meteotsunamis, the duration spread and start time spread generated by the ensemble in Table 2.

Additionally, we colour-coded the members able to capture a minimum meteotsunami threshold of 0.2m and a larger threshold of 0.4m (observed threshold) in the figure. We believe that showing each individual member is more useful than showing an ensemble spread since the oscillatory behaviour of the model is more important to be captured than the actual inter-member spread. This now highlights increased performance at 1-day lead time (fewer red members).



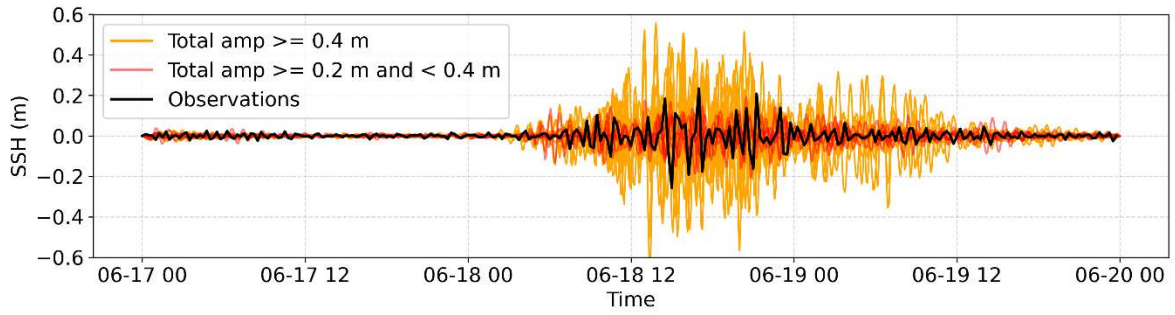


Table 2. UKC4 forecasting performance at 1-day lead time for selected locations. Reported values include the modelled (UKC4) and observed peak signals, along with the probability of exceeding three thresholds: 0.2 m, the activity threshold (act_thresh), and the observed peak. Event characteristics are summarised by the duration of consecutive hours above the activity threshold and the event start time. For UKC4, mean duration and start time across ensemble members are shown, with minimum and maximum values provided in parentheses. When minimum and maximum values are not reported, the timings are identical across all ensemble members.

Location	UKC4	Obs.	UKC4 Prob. > 0.2	UKC4 Prob. >act_thresh	UKC4 Prob. >obs	Obs Duration Consec. Hrs	UKC4 Mean Duration (min/max) Consec. Hrs	Obs Start time	UKC4 Mean Start Time (min-max)
Milford Haven	0.9 m	0.5 m	100%	100%	88.9%	6.3h	8.2h (2.4 - 13.5 h)	18/06 15:00	18/06 12:55 (10:15 - 16:54)
Union Hall	0.1 m	1 m	16.7%	100%	0%	14.2h	37h (33-45h)	18/06 14:20	17/06 02:45 (2:30 - 3:10)
Dieppe	0.1 m	0.8 m	22.2%	100%	0%	18.1h	8.0h (5.8 - 11.7)	18/06 20:23	19/06 00:47 (18/06 20:45 - 19/06 03:10)
Le Crouesty	0.1 m	0.4 m	100%	100%	0%	2.9h	14.7h (5.5 - 56.4)	19/06 00:48	19/06 02:05 (18/06 21:10 - 19/06 04:30)

Table 3. Same as Table 2 but for 3-day lead time.

Location	UKC4	Obs.	UKC4 Prob. > 0.2	UKC4 Prob. >act_thresh	UKC4 Prob. >obs	Obs Duration Consec. Hrs	UKC4 Mean Duration (min/max) Consec. Hrs	Obs Start time	UKC4 Mean Start Time (min-max)
Milford Haven	0.9 m	0.5 m	100%	100%	100%	6.3h	23.2 (3.5 - 29.9 h)	18/06 15:00	18/06 08:20 (3:35 - 19:30)
Union Hall	0.1 m	1 m	16.7%	100%	0%	14.2h	35.6h	18/06 14:20	18/06 09:05 (7:05-11:20)
Dieppe	0.1 m	0.06 m	5%	100%	0%	18.1h	29.8h	18/06 20:23	18/06 11:45
Le Crouesty	0.1 m	0.4 m	22.2%	77%	0%	2.9h	15.4h (5.5 - 56 h)	19/06 00:48	19/06 02:10 (18/06 21:00 - 19/06 02:10)

"To assess forecast uncertainty, we include a set of ensemble-based statistics from the UKC4 system at each country's location showing the strongest observed signals: Milford Haven, Union Hall, Dieppe, and Le Crouesty. Tables 2 and 3 summarise the statistics for forecasts issued 1 day and 3 days in advance, respectively. Each table reports the maximum signal simulated by UKC4 alongside the corresponding observed values.

We further estimate the probability that the UKC4 signal exceeds three predefined thresholds.

(a) the observed signal at each location. This provides a direct measure of the model's ability to reproduce or surpass the magnitude of the observed event.

(b) the minimum meteotsunami threshold (0.2 m)

(c) Given the model's limitation to sometimes capture the full amplitude of the meteotsunami signal, but still showing a significant signal in comparison to its own baseline of variability, we define a third threshold based on a scaled increase in model variability (hereafter referred to as the activity threshold), defined as:

$$\text{activity\_threshold} = \sigma_{\text{model, day1}} \times \left(\frac{\sigma_{\text{obs, meteotsunami duration}}}{\sigma_{\text{obs, day1}}} \right)$$

where σ denotes the standard deviation of the filtered signal. "Day 1" refers to the first forecast day, during which no meteotsunami activity is present, used to calculate "baseline variability".

"Model" corresponds to each UKC4 ensemble member, and "obs" denotes observations. The observed meteotsunami duration at each location is defined as first and last time step when the absolute value of the filtered signal is above 0.1m (half the meteotsunami threshold).

The duration of the event for the model at a given location is then defined as first and last time step when the absolute value of the filtered signal is above the activity threshold. We add a condition that duration is only counted for members which maximum amplitude still exceed the minimum theoretical meteotsunami signal threshold of 0.2m. These values are reported in the tables alongside the corresponding observed duration and start time.

Overall, the one-day lead forecasts presented in Table 2 indicate that the model is able to show the possibility of a signal which would qualify as meteotsunami (>0.2m) for all the sites, ranging from 16.7% in Union Hall (Ireland) to 100% in Milford Haven (UK) and Le Crouetsy (France). All the members also show considerable increase in filtered SSH variability compared to the 17/06 for all sites (>activity threshold), indicating a strong likelihood of abnormal high-pass sea level variations. The model is only able to predict the observed values of meteotsunami amplitude for Milford Haven.

Regarding the start time of the event, the observed start time is within the interval predicted by the ensemble for two sites (Milford Haven and Le Crouetsy), and slightly off by 20mn for Dieppe. In Union Hall, however, the model starts the event 12h earlier than observed. The duration of the event is more difficult to predict: the range of durations indicated by the model is large, indicating weak skills to determine the exact event duration. Nevertheless, the ensemble mean duration shows absolute errors ranging from 1.9h (Milford Haven) to 11.8h (Le Crouesty).

At longer lead times (Table 3), the overall skill decreases, as reflected by reduced probabilities of exceeding the minimum criteria for meteotsunami amplitude (0.2m) and increased spread in both duration and timing estimates. Nevertheless, the model retains high probabilities of enhanced filtered SSH variability on the 18/06 and 19/06 compared to 17/06 suggesting that while amplitude prediction remains challenging, the model can still provide useful indications of potential meteotsunami occurrence on 18/06 and 19/06."

Minor comments

Line 11-13:

The sentence does not read well and needs paraphrasing. If the Met Office system is more successful over certain area the sentence implies that the Météo-France system is worse or less successful in representation over those area. But this is not completely true since in lines 356-357 is written that in the Bay of Biscay as the forecast progresses the Météo-France system (AROBASE) outperforms the Met Office (UK4) system.
Thank you for highlighting this point and for requesting clarification.

The apparent inconsistency arises because the two statements refer to different variables. Lines 11–13 describe the model performance in representing the meteotsunami signal itself (i.e., sea surface height), whereas Lines 356–357 refer to atmospheric pressure perturbations rather than the resulting ocean response.

We agree that the previous wording could be misleading. In particular, stating that AROBASE "outperforms" UKC4 in this context is not appropriate, as no direct comparison with observations is made for the pressure perturbations, and therefore no assessment of relative skill can be robustly established.

We have revised the text in this part accordingly. The updated wording now simply describes the differences observed between the systems (e.g., larger pressure perturbations in AROBASE) without implying superior performance. This ensures consistency and avoids overstating the results.

Line 23:

Sentence "Their period is between 2 and 120 min." is missing a reference.
Thank you for highlighting this, now corrected.

Line 35-36:

In sentence "Even though these phenomena can cause significant disturbances and even be high risk for coastal infrastructures, property and human life, they are usually underestimated and overlooked."

change "be high risk" to "be of high risk". Also, the word "even" is repeating. Paraphrase for clarity and smoothness.
Thank you for the suggestions – included.

Line 47:

"However, their ..." change to "However, the study of meteotsunamis remain rare ..."

Thank you for the suggestion – included.

Line 50-52:

Sentence: "In 2013, with an update in 2018, O'Brien et al. attempted to give a catalogue of extreme wave events in Ireland that included meteotsunamis (O'Brien et al., 2013, 2018)."

Paraphrase the sentence to something like:

"A catalogue by O'Brien et al. (2013), with an updated version in 2018 (O'Brien et al., 2018), documented extreme wave events in Ireland, including meteotsunamis." since in the original "O'Brien et al." it reads like a citation with missing year.
Thank you for highlighting this and for the suggestion – included.

Also "attempted"? Were they not successful? If not, write why.
Thank you for spotting this. The use of "attempted" was unintended and does not accurately reflect the outcome, so we removed it.

Line 54-55:

Sentence: "From their analysis,..." change to "According to their analysis,...."
Thank you for the suggestion – Now included.

Line 56-57:

"It was only until very recently..." only + until seems redundant. Keep it simple, "It was only very recently that their importance and frequency around the UK have begun to be studied

in more detail, with studies trying to understand and update the records of their occurrence."
Thank you for the suggestion – Now included.

Line 135:

Add long name before SSH since abbreviation is mentioned for the first time in the text but its long name "sea surface height (SSH)" is for the first time written in the line 247.
Thank you for spotting this – Now corrected.

Line 268:

Repentance of long name "sea surface height (SSH)" which is written in line 247 but should be already in 135.
Thank you – also corrected.

Line 337-339:

There are open water locations written for Celtic Sea and Bay of Biscay while English Channel is missing the locations and it is only written that those are locations from the lightships. What were the locations of the lightships? Write those as well.
Thank you for spotting this – now included.

Line 343-344:

"In the Celtic Sea (Fig. 9), UKC4 consistently shows higher disturbance amplitudes, reaching around 200Pa at all three locations."

Figure 9 only shows two locations. Also, higher than what? Write than you mean higher than AROBASE.
Thank you for spotting these. Three locations was a typo, while we have now added that we compare with AROBASE for this.

Line 351:

"The plots ..." write "the Figs. 9-12"

Thank you for the suggestion – included.

Line 355-356:

"At the two locations in the Bay of Biscay (Fig. 11), the AROBASE system shows a noticeable improvement in capturing high-frequency pressure disturbances, with amplitudes approaching 400Pa, comparable to those produced by UKC4."

Is the UK4 a reference?

Thank you for seeking clarification. UKC4 is not used as a reference in this context. The comparison is intended to provide context for the magnitude of the pressure perturbations, as UKC4 generally shows stronger and more consistent performance across the study. We agree that the original wording could be misleading and have revised the sentence to avoid implying that UKC4 is a benchmark.

The sentence now reads:

"At the two locations in the Bay of Biscay (Fig. 11), the AROBASE system shows a noticeable increase in the amplitude of high-frequency pressure disturbances, reaching values of up to ~400 Pa, which are of a similar order of magnitude to those simulated by UKC4."

Adding as well the observations used in Renzi et al. (2023) and referring to these plots in lines: *"To further assess the ability of the two models to capture atmospheric perturbations, additional validation of the atmospheric pressure forcing is provided in Appendix Fig. A4. This analysis is based on the same Irish stations used in Renzi et al. (2023), with data processed using consistent filtering. The results corroborate the findings discussed above 380 and offer further insight into why UKC4 outperforms AROBASE. As shown in Fig. A4, UKC4 exhibits very good agreement with the observations, clearly reproducing the 2 hPa pressure jump. Although minor discrepancies in timing are present, the overall representation remains adequate. In contrast, AROBASE shows weaker performance at the two stations within its domain: the pressure jump is less pronounced, which likely contributes to a reduced amplification of the meteotsunami signal."*

Line 356-357:

How is it proved that AROBASE is outperforming UK4 if no observations were used?

Thank you, as mentioned we have revised the wording there.

Line 361-363:

Please refer to the figures which show meteotsunami signal (Brest in Fig. 5 and Le Havre in Fig. 4).
Thank you for spotting this, the figures are now properly stated.

In addition, UK4 does not have doubled the AROBASE amplitude in Le Havre, where actually around 09:00 UTC the AROBASE is showing higher amplitude.

Thank you for spotting this. Wording is now corrected.

Line 366:

"which" change to "whose"
Thanks for spotting this – now corrected.

Line 413:

Add Fig. 14 (bottom panel) after "higher than 4mmh−1".
Thank you for the suggestion – now added.

Add Fig. 7 after "In this area, isobars are not smooth" for better readability.

Thank you for the suggestion – now added.

Line 434:

"...first in the Celtic Sea ..." no need for first if second is not mentioned later on.
Thank you for the correction – "first" is now removed.

Line 439:

"up" write updrafts instead
Thank you for the suggestion – now corrected.

Line 462-465:

The pressure disturbance observation for Dieppe station is not shown in any of the figures. It can hardly be referred as Greenwich location noted with blue star in Fig. 8 whose pressure disturbance is shown only from model simulation in Fig. 10. (right-hand side).
Thank you for seeking clarification here. We added "…while comparing to available observations in Le Havre its pressure disturbance was…" to make clear for which location we talk about.

Line 513:

"...even if high resolution..." change "if" with "...with high resolution..." or "...of high resolution..."
Thank you for the suggestion – now corrected.

Line 517:

"not sent back the atmosphere" change to: "not send back to the atmosphere"
Thank you for spotting this – now corrected.

Line 548:

"either" requires second parameter. Either one or the other.

Thank you for spotting this. "either" is now replaced by "the"

Line 551:

"to have" change to "having"
Thank you – now corrected.

Line 553-554:

"Finally, both systems show a weak signal in the Bay of Biscay (Fig. 5)"

Write which variable shows a weak signal and is shown in Fig. 5. Which is SSH.

Thank you very much for spotting this. We have added "meteotsunami" before "signal" to be clear.

Line 576:

"This can conclude". Write, "From this it can be concluded ..."

Thank you for the suggestion – now included.

Line 589-590:

Indicate that pressure data are only from model (except Brest and Le Havre). Clearly state this here.
Thank you for spotting this, now corrected.

Line 610-612:

Give reference for "one significant event every five years"
Thank you for spotting this. Lewis et al.,2023 is now included as reference.

Line 621:

In "The importance of high-frequency observational data is further illustrated in Figure A2" write "is further illustrated for Union Hall station (Fig. A2)" since in the original writing the Union Hall station is not linked with any figure in the text.
Thank you for spotting this – now corrected.

Line 622:

If the model has temporal resolution of 10 min than in "... a 5 min observation ...." should be "... at 5 min observation ...".
Thank you for spotting this – now corrected.

Comments regarding figures

Most of the figures have very small labels, especially Fig. 13. Increase the labels i Figs 9-13. and Figs. 17-19.

Thank you for your suggestion, the labels and general font has been now increased in the above figures.

Also it would be better if longitude is written in SW/NE notation not 0-360 and without noting it is degrees on x axis.

Figure 2. Add the meaning of the station colours in the caption.
Thank you for seeking clarification, now included together with a statement for recording frequency.

Figure 8. Add the meaning of the colours in the caption.

Thank you for seeking clarification, colours now included.

Figs. 9-12 Clearly state that the MSLP data are from models, not lightships, as indicated in the text on line 247. Only Fig. 12 shows the MSLP observations.

As currently written, it may give the impression that all MSLP time series are based on observations.

Thank you for this comment. Legend has now been added to the plots for clarity.

Figure 10. is missing coordinates in the caption and in the text.

Thank you for spotting this, now corrected.

Figure 15. In caption: upper level warm front is marked with full semi-circles not empty.

This is correct, but we added more description in the caption as this feature is quite small.

The upper level warm front is marked on the 18:00\,UTC chart **(right panel)** with empty semi-circles along its line, **it is located over England**.

Figure 16: there is something like watermark above the figure showing a) 12:00 UTC - can this be removed? Probably an artifact of figure cutting and pdf conversion.

Figure 19: Add the meaning of the colours.
Thank you for spotting this, legend had now been included for clarity.

Fig. A1 is missing the station name. while A2 has it in the tile. Write the station name in caption for both.
Legend is not explained anywehere.

Thank you for spotting these, now both corrected including improvements to readability (font size)

MO data in Fig. A1 as well as 5 min data in Fig. A2 seems to be cut by the y axis limits. Show full range of 5 min data.

Thank you for the suggestion, now corrected.

Similar is cut in Fig.3 also for Milford Heaven station.
Thank you for this comment. However, as explained in our response in the main comments, a consistent axis scale was used across all panels to facilitate comparison and to ensure that locations with weaker signals remain clearly visible. Expanding the y-axis range would compromise the readability of these features. Nevertheless, we agree that the reviewer's comment is well taken, and for completeness we have included these results as a standalone figure in the Appendix (now presented as the first figure).

Mixing of "Fig. X" and "Figure X" in text. Use Figure in full only at the beginning of a sentence and in figures captions.

Thank you for your comment. This has now been corrected throughout the manuscript.

Comments regarding references:

Angove, M., Kozlosky, L., Chu, P., Dusek, G., Mann, G., Anderson, E., Gridley, J., Arcas, D., Titov, V., Eble, M., McMahon, K., Hirsch, B.,and Zaleski, W.: Addressing the meteotsunami risk in the United States, Natural Hazards, 106, 1467–1487, https://doi.org/10.1007/s11069020-04499-3, 2021.

Kain, J. S. and Fritsch, J. M.: A One-Dimensional Entraining/Detraining Plume Model and Its Application in Convective Parameterization.,

Journal of the Atmospheric Sciences, 47, 2784–2802, https://doi.org/10.1175/1520-0469(1990)047<2784:AODEPM>2.0.CO;2, 1990.

Also some papers do have functioning links but those are not included, i.e.:

O'Brien, L., Dudley, J., and Dias, F.: Extreme wave events in Ireland: 14680 BP to 2012, Nat. Hazards Earth Syst. Sci., 13, 625–648, 2013.

O'Brien, L., Renzi, E., Dudley, J., Clancy, C., and Dias, F.: Catalogue of extreme wave events in Ireland: revised and updated for 14680 BP to 2017, Nat. Hazards Earth Syst. Sci., 18, 729–758, 2018.

This will be corrected in the updated manuscript! Thank you.