Detecting anomalous sea-level states in North Sea tide gauge data using of auto associative Neural Network

Dear Agustín Sánchez-Arcilla, Dear Reviewer,

Thank you motivatiforng us to improve further our paper, for reading it so carefully and for your valuable comments. We have corrected the manuscript accordingly and provide the answers below.

Looking forward for any further instructions.

Best regards,

Kathrin Wahle

Review #1

The manuscript aims to investigate the skill of Auto Associative Neural Networks (AANNs), trained with different set of observation- and model-based data, to emulate sea-level and detect extremes in the North Sea. The method proposed by the authors relies on the capabilities of linear and non-linear models based on principal components (PCs) and AANNs, respectively, to reconstruct sea-level states. The departures between the reconstructed and observed sea states are then used to detect extreme events in the North Sea. The results, focused on two events characterized by low pressure systems and high wind speeds over the North Sea, show the potential of AANNs trained with tide-gauge records in detecting accurately the occurrence of extremes, both as a very localized and larger scale events, while underline the inaccuracies of the emulators based on PCs (false positives) and trained with model-based sea-surface height outputs due to model physics (Gaussian distribution of errors). The latter can be used to enhance representation of sea-level extreme events in ocean models. The authors highlight the high potential of their approach which can be easily extended to other sea-level observing networks.

The work is well developed, and the publication is recommended after a minor revision of the manuscript

Here follow some suggestions, which could help improve certain parts of the manuscript.

General Comments

1. A few sentences (e,g in the Introduction and Section 2) should be rephrased, and relevant references to the literature should be acknowledged (see Specific Comments).
We rephrased many sentences throughout the paper to improve its quality and readability and added relevant references.

2. Sometimes test, validation, and control are used as synonyms: check it and ensure consistency.
Thank you for the comment. For consistency, we used in the revised manuscript 'test'

3. The trained AANNs used in this work should be named appropriately (e.g. reference) and summarized in a dedicated Table giving a short description of the data sets used in training and testing phase.
We added to Table 1 the names of AANN's and their dataset. The training (2016/2017) and validation (2018) phases are the same for all NNs. We clarified and added this information, too:

“As for AANN_ref, training data in the latter three networks span the two year period 2016/17 and data from 2018 were kept for testing generalization abilities.”

4. Description of error distribution of AANNs (Figure 4): those results guide your analysis and conclusions, but it is hard to distinguish between the tail of error distributions in Figure 3 and Figure 4. This is probably due to poor quality of the graphics (see specific comments about Figures), which should be improved, but a more extended description of the results would help the reader to capture the actual meaning of the results.
Thank you for the comments. We have replotted this (and all other) Figure(s) with higher resolution and better quality; we enlarged the insets and provided an extended Figure caption.

5. The organization of text in the different Sections should be revised (see Specific Comments).
Specific Comments

Introduction
General comment to Introduction:
Lines 27—44: We have restructured this part of the introduction, separating now numerical models from satellite measurements and tide gauges.
Line 17: “tidal motion …. tidal motion” - rephrase.
We rephrased as: “. Theoretical prediction of tidal motion was pioneered by the application of Fourier analysis by Lord Kelvin (Thomson, 1880) and later improved by Doodsen (1921), who developed the tide-generating potential in harmonic form.”
Line 22-23: add relevant references.
We added the following recent relevant publications:
Sandery P.A., Sakov P. (2017)Ocean forecasting of mesoscale features can deteriorate by increasing model resolution towards the submesoscale. Nature Commun., 8, pp. 1-8
Stanev EV, F Ziemer, J Schulz-Stellenfleth J Seemann, J Staneva and KW Gurgel (2015) Blending surface currents from HF radar observations and numerical modelling: Tidal hindcasts and forecasts. Journal of Atmospheric and Oceanic Technology, Vol. 32, 256-281.
Mey-Frémaux, P. de, Ayoub, N., Barth, A., Brewin, R., Charria, G., Campuzano, F., Ciavatta, S., Cirano, M., Edwards, C.A., Federico, I., Gao, S., Hermosa, I.G., Sotillo, M.G., Hewitt, H., Hole, L.R., Holt, J., King, R., Kourafalou, V., Lu, Y., Mourre, B., Pascual, A., Staneva, J., Stanev, E.V., Wang, H., & Zhu, X. (2019): Model-Observations Synergy in the Coastal Ocean. Front. Mar. Sci., 23 July 2019, doi:10.3389/fmars.2019.00436
Jacobs, G., D’Addezio, J.M., Ngodock, H. and Souopgui, I. (2021). Observation and model resolution implications to ocean prediction. Ocean Modelling, 159, 101760, Ponte, R. M., Carson, M., Cirano, M., Domingues, C. M., Jevrejeva, S., Marcos, M., ... and Zhang, X.: Towards comprehensive observing and modeling systems for monitoring and predicting regional to coastal sea level. Frontiers in Marine Science, 6, 437, 2019.
Line 25-26: explain shortly the improvements you are mentioning.
We added, that by enhancing model resolution to 1.5 km, dynamical features such as coastal currents, fronts, and mesoscale eddies are better resolved, and model results improve, especially when compared to high spatial–temporal resolution observations.
Line 31-34: Rephrase (e.g. change the order of the sentences)
We rephrased: “However, advancements are underway, and new satellite missions characterized by better spatial and temporal sampling pave the way for improvements in coastal sea-level research (e.g. Dieng et al., 2021; Prandi et al., 2021; Dodet et al., 2020, Sanchez-Arcilla et al., 2021a,).

Tide gauge stations operating along the North Sea coast provide high-quality records of sea level observations over a long period (Wahl et al., 2013). Ponte et al. (2019), reviewing the state of science of coastal sea-level monitoring and prediction, outlined the importance of sea-level observations for studying sea-level variability. However, tidal gauges do not provide information about the basin-wide patterns of sea level. Furthermore, some of these data are not continuous; different gauges do not always operate simultaneously and there are gaps in many of the records.

Line 34: the sentence on numerical models can move further down in the text

We extended the paragraph on recent developments in numerical modelling (Lines 37ff) and added the information there: "Recent important evolution in predicting sea-level in the North Sea was achieved in the framework of the development of the Northwest European Shelf forecasting system (e.g. O'Dea et al., 2012, Tonani et al., 2019) by enhancing model resolution to 1.5 km. Thus dynamical features such as coastal currents, fronts, and mesoscale eddies are better resolved, and improve model results, especially when compared to high spatial–temporal resolution observations."

Line 38-42: Rephrase

We now state: "Therefore, a consistent dataset combining the gains of numerical model results with tide gauge measurements would be beneficial."

Line 42-43: This sentence can move e.g. to line 37: "Ponte et al., (2019) …. Sea-level variability. Therefore,..."

Yes. It now reads: "Tide gauge stations operating along the North Sea coast provide high-quality records of sea level observations over a long period (Wahl et al., 2013). Ponte et al. (2019), reviewing the state-of-the- of science of coastal sea-level monitoring and prediction, outlined the importance of sea-level observations for studying sea-level variability. However, tidal gauges do not provide information about the basin-wide patterns of sea level. Furthermore, some of these data are not continuous; different gauges do not always operate simultaneously; , and there are gaps in many of the records. Therefore,…"

Line 44: This sentence should be linked to the findings of previous work (e.g Zangh et al. 2020).

Yes. We link the two sentences, by:

"Zhang et al. (2020) use machine learning to reconstruct the sea-level variability in the North Sea using observations from 19 coastal tide gauges and data from numerical models. Noteworthy, they concluded that a relatively short-time record contains the most representative characteristics of sea level dynamics in the North Sea. While this was clear for the tides, it was not so obvious about changes in sea level caused by the atmosphere."

Line 47-51: Rephrase and add relevant references to the literature.

We rephrased:

"The coupling of the respective processes is, in most cases, nonlinear (Jacob et al., 2017), that is, one cannot easily consider the response to individual drivers in isolation. This happens, when either oscillatory motion has large amplitudes, e.g., tidal currents approaching 1 m/s, or wind driven current is of the same order. Thus, there is a need to use methods tailored to detect and reproduce nonlinear dynamics. The nonlinear processes are difficult to predict, even with sophisticated models; therefore, one also has to identify situations in which predictions fail (Ponte et al., 2019). Furthermore,…"

Jacobs, G., D'Addezio, J.M., Ngodock, H. and Souopgui, I. (2021). Observation and model resolution implications to ocean prediction. Ocean Modelling, 159, 101760,

Ponte, R.M., Carson, M., Cirano, M., Domingues, C.M., Jevrejeva, S., Marcos, M., Mitchum, G., Wal, R.S.W. van de, Woodworth, P.L., Ablain, M., Ardhuin, F., Ballu, V., Becker, M., Benveniste, J., Birol, F., Bradshaw, E., Cazenave, A., Mey-Frémaux, P. de, Durand, F., Ezer, T., Fu, L.-L., Fukumori, I., Gordon, K., Gravelle, M., Griffies, S.M., Han, W., Hibbert, A., Hughes, C.W., Idier, D., Kourafalou, V.H., Little, C.M., Matthews, A., Melet, A., Merrifield, M., Meyssignac, B., Minobe, S., Penduff, T.,

Picot, N., Piecuch, C., Ray, R.D., Rickards, L., Santamaría-Gómez, A., Stammer, D., Staneva, J., Testut, L., Thompson, K., Thompson, P., Vignudelli, S., Williams, J., Williams, S.D.P., Wöppelmann, G., Zanna, L., & Zhang, X. (2019): Towards Comprehensive Observing and Modeling Systems for Monitoring and Predicting Regional to Coastal Sea Level. Front. Mar. Sci. 6:437, doi:10.3389/fmars.2019.00437

Line 58-59: Add references.

We reformulated this paragraph and moved it to the end of the introduction. We added reference on AANN:

"In contrast to these applications, we will focus on the identification of situations in which spatial correlations between tide gauge measurements deviate greatly from the dominant principal ones. Usually, the entirety of tide gauge measurements in the North Sea will show specific spatial correlations (changing in time with the tide). However, in anomalous situations (e.g. localized storms) these correlations may drastically change. We use autoassociative neural nets (AANNs, Kramer, 1992) to detect such sea-level states. Atmospheric conditions related to such situations might aid further understanding and future developments in sea-level prediction."

Kramer, M. A.: Autoassociative neural networks. Computers & chemical engineering, 16(4), 313-328, 1992.

Line 83-84: Rephrase .

We rephrased as :

"The paper is structured as follows: In section 2 we present the observational and model data used throughout our study and introduce the concept of AANN. We then apply AANN onto the tide gauge array measurements in the North Sea. In section 3 we use AANN to detect anomalous events and examine the dependence of identification of such events from the AANN training data used. Two events are studied in detail, including atmospheric conditions. This is followed by a discussion and conclusions."

<u>Section 2.1</u>

Line 111: Rephrase.

We rephrased this and the following sentence:

"Time versus position diagrams of measured and detided (residual) water levels are shown in Figure 2 a, b. The former shows the propagation of the tidal wave (the slope of contours gives the speed of propagation) along the English coast with water level amplitudes increasing southward towards the channel. The specific feature between stations Cromer and Vilssingen identifies the small amfidrome in front of the English Channel. This feature is not present in the detided data (Figure 2 b), the later resembles the atmospheric forcing (Figure 2 c, d). In the presence of large gradients in wind tendency (Figure 2c), the water surface tilts considerably compared to the case of small gradients (Figure 2b)."

<u>Section 2.2</u>

Line 134: add references.

We added the reference to CMEMS tide gauge product: CMEMS global ocean in situ near-real-time observations, https://doi.org/10.48670/moi-00036

Line 143: avoid "very".

Lines 137-144 have been deleted and replaced by the following paragraph:

"Sea level data e.g. will form a 2D manifold. Different machine learning techniques, such as k-nearest neighbours algorithm, ensemble-based methods, and support vector machine (SVM) algorithms predict the posterior probabilities of a given dataset and are optimal for data compression. The different techniques are not equally well suited for detecting outliers, i.e. in deciding whether a given observation belongs to the same probability distribution. Outlier

detection in high-dimension, or without any assumptions on the distribution of the inlying data is very challenging. SVM algorithms work well if training data is not contaminated by outliers. Ensemble and k-nearest neighbour methods perform well for multi-modal data sets. Covariance estimators (in which category Principal Component Analysis, PCA, falls, too) degrade when the data is not unimodal. Autoassociative neural nets (AANN) combine the robust performance multi-modal data with the geometrical interpretability of PCA to identify these situations…"

Line 144: Rephrase**. (Lines 137-144 have been deleted; see answer to Line 143)**

Line 144: check punctuation. **(Lines 137-144 have been deleted; see answer to Line 143)**

Line 155: avoid repetitions: "Thus,…" **Done.**

Line 192: "presents" → "shows" **Done.**

Line 204: Is this choice based on a sensitivity analysis? The choice of the constraints/threshold errors should be explained in the text.

The choice is based on results within the training period of AANN. We added on constraints/thresholds the following: "Thus, we postulate an ocean state as anomalous when the relative error exceeds 0.035 in the position of at least 3 gauges simultaneously for at least 3 hours."

Line 207-211: "…, we trained the AANN based on the following data sets:

• sea-surface height outputs from CMEMS AMM15 (operational model) extracted at the tide-gauge positions (AANN_NEMO);

• observed total water levels from only 10 tide-gauge stations (with …) (AANN_less); and

• observed non-tidal water level residuals for the 14 gauge stations (AANN_resid). "

We added information on data sets to Table 1:

	training dataset	Spring (MAM)	Summer (JJA)	Autumn (SON)	Winter (DJF)
AANN_NEMO	NEMO model (14 data points)	2	2	6	13
AANN_less	tide gauge data (raw, 10 gauges)	1	0	6	13
AANN_resid	tide gauge data (de-tided, 14 gauges)	1	1	6	13
AANN_ref	**tide gauge data (raw, 14 gauges)**	**2**	**2**	**6**	**13**

Line 215: "areal" ?

We changed into "spatial"

Lines 225: "…a comparison between observed, modelled and AANN emulated…".

We show now rather the differences and thus reformulated: "A comparison of NEMO modelled and AANN emulated water level residuals with measured ones …"

Line 232: "the reference model" ?. I assume this refers to AANN in Table 1 (See General Comments).

Yes (see table 1 above), it is now referred to as 'AANN_ref' throughout the text.

Line 239: "In the following Section,…": in the current version of the manuscript this Section is followed by the Conclusion. This part of the text should be reorganized.

We removed 'section'.

Line 240: It is unclear which ANN model error you are referring to (ANN_resid ?). The same applies tp NEMO model error: is it the error in NEMO ssh outputs?

We clarified by referring to AANN names given in table 1 (see above) throughout the manuscript.

Comments about Figures

We increased resolution for all Figures to 400 dpi, made them more readable.

Figure 1: The description in the caption should be improved as well as the labels in the figure (see wind tendency). **We splitted this Figure into 2 figures and changed Figure caption, too.**

The captions now read:

"Figure 1: (a) Location of tide gauge stations used in this study and M2 cotidal chart. Additionally, snapshots of NEMO model water levels for two selected times are shown together with water level residuals at tide gauge locations (black lines) and a plane fitted linearly to these residuals (grey areas): (b) 7. January 2017, 16 UTC and (c) 11. January 2017, 20 UTC."

"Figure 2: Time versus position diagrams series of (a) measured water levels, (b) water level residuals, (c) wind tendency (defined as change in wind speed within one hour, here) and (d) pressure anomaly for January 2017 at tide gauge positions. "

Figure 3: It is hard to read the labels.

We have replotted the Figure with readable labels.

Figure 4:

- Panels a, b, d, e (see previous comment on Figure 3).

We have replotted the Figure with readable labels.

- Panel c: It is very hard to read this panel, both time-series and labels. Also: show hours on x-axis.

Fig.4c is now a separate figure showing a time versus position diagram of differences of NEMO model and tide gauge data, respectively. Labels are readable now.

Figure 5: The panels should be displayed bigger as well as the labels in each of them.

We splitted Figure 5 into two parts. Panels are bigger and readable now.

Caption:

- "…a 14-day…" **Changed to actual periods: 1.-15. June 2017**
- "NEMO model error" – is it the AANN_NEMO ? Clarify. **Done. All reconstruction models are now shown in one Figure.**

Figure 6: Make sure the labels on top of the panels correspond with the description in the caption.

Done. We have extended this Figure to contain all AANN reconstructions as well as those based on PCA. Additionally, a comparison with NEMO modelled results are shown.

Figure 7: see comments on Figure 5.

We have replotted as described for Figure 5.

Figure 8: see comments on Figure 6.

We have replotted as described for Figure 6.

Figure 9: The panels can be displayed better. Avoid legend.

We changed this Figure into a time versus position diagram.