ADVANCING EARTH AND SPACE SCIENCES

JGR Machine Learning and Computation

RESEARCH ARTICLE
10.1029/2025JH001192

Drivers of Interannual to Decadal Sea Level Variability in Northern Europe—Data Driven Approach

Lea Poropat[1] and **Céline Heuzé**[2]

[1]National Centre for Climate Research, Danish Meteorological Institute, Copenhagen, Denmark, [2]Department of Earth Sciences, University of Gothenburg, Gothenburg, Sweden

Key Points:

- We analyze sea level variability in northern Europe using neural networks and linear regression combined with permutation feature importance
- Sea level is mainly driven by local atmospheric forcing followed by NAO in northern Baltic and sea surface temperature in the North Sea
- Optimal model is region-specific with neural networks performing best in the Baltic and linear regression in the rest of the study area

Correspondence to:

L. Poropat,
lep@dmi.dk

Citation:

Poropat, L., & Heuzé, C. (2026). Drivers of interannual to decadal sea level variability in northern Europe—data driven approach. *Journal of Geophysical Research: Machine Learning and Computation*, *3*, e2025JH001192. https://doi.org/10.1029/2025JH001192

Received 13 DEC 2025
Accepted 21 MAY 2026

Abstract

Coastal sea level reflects not only global trends but also complex regional and local processes acting across multiple temporal and spatial scales, which are often missed by large-scale climate models. To address this, we use a data-driven approach to examine potential drivers of interannual to decadal sea level variability in northern Europe. We train neural network and linear regression models to simulate monthly mean sea level from 45 tide gauges using 11 potential drivers and quantify their contributions using permutation feature importance. To include possible lag or memory in the system, models include forcing history. Models explain approximately 70% of observed variability, except in the Danish Straits, where lower skill (25%–50%) suggests missing drivers. In the Baltic neural networks perform best, revealing highly nonlinear relationships between sea level and its drivers, while in other areas linear regression works better, indicating predominantly linear driving mechanisms. Most locations reveal a memory of at least one previous month, often longer. Primary sea level drivers are local wind and atmospheric pressure, followed by the North Atlantic Oscillation and sea surface temperature (both local and globally averaged), with minor influence from precipitation. Greenland and Antarctic mass loss, as well as regional runoff and evaporation do not affect sea level variability on these time scales. Regional analysis reveals clear spatial patterns, with different driving mechanisms in the North, Baltic, and Norwegian Seas, and the Danish Straits. These findings enhance our understanding of regional sea level variability and offer additional tools for improving coastal flood risk assessments.

Plain Language Summary

With rising sea level, coastal areas face increasing risk from flooding. Local sea level is influenced by a mix of global changes, regional climate variability, and local conditions. Understanding all these factors is essential for making accurate coastal risk assessments. In this study, we use machine learning to find what contributes most to year-to-year sea level fluctuations in northern Europe. We predict sea level from other atmospheric, ocean and climate variables using neural networks and linear regression models and analyze each variable's impact on the prediction using permutation feature importance method, which reveals its influence on the sea level. We find that these slow sea level fluctuations are mainly caused by a combination of local wind and atmospheric pressure, with some influence from a climate pattern called North Atlantic Oscillation, sea surface temperature variations, and precipitation. There are notable regional differences between the North, Baltic, and Norwegian Seas, and the Danish Straits. Our findings improve understanding of how sea level behaves locally, which can improve climate projections in this area, help with flood risk assessments and support better planning and adaptation in vulnerable coastal communities.

1. Introduction

Global mean sea level is rising, and its rise has accelerated to 3.7 mm y^{-1} over the period 2006–2018 (Fox-Kemper et al., 2023). Much of the world's population, economic activities and critical infrastructure are concentrated near the sea, and because of the sea level changes face very high risks (Glavovic et al., 2023). But sea level is not rising uniformly, there are significant regional differences (Cazenave & Le Cozannet, 2014; Slangen et al., 2014). Superimposed to the global sea level rise caused by temperature expansion, melting of the ice sheets and glaciers, and changes in the land-water storage are the regional variations caused by changes in ocean currents and density, as well as gravitational, rotational and deformational effects resulting from changes in the loading of ice and water masses and vertical land movement (Durand et al., 2022). Sea level is also driven by atmosphere, through changes in air pressure captured by the inverse barometer effect, and wind which transfers momentum across the air-sea interface (Rus et al., 2023), on all scales from hourly to decadal. The presence of significant interannual to decadal variations in sea level linked to internal climate variability hampers the early detection of sea level accelerations due to anthropogenic climate change (Frederikse et al., 2016; Haigh et al., 2014).

Investigating the sources of this variability is therefore crucial to determine whether recent or possible future accelerations in sea level represent temporal fluctuations related to internal climate variability or are the response to anthropogenic forcing (Dangendorf et al., 2014). However, climate models have relatively low spatial resolution, which limits their ability to fully capture the complex interactions that influence coastal sea level.

To address these limitations, there has been growing interest in data-driven approaches that can better resolve local variability. Supervised machine learning methods have been extensively used for sea-level modeling due to their ability to learn from labeled data and make predictions based on established relationships (Ayinde et al., 2024). Already more than 25 years ago Hsieh and Tang (1998) describe the first usages of neural networks for prediction and data analysis in meteorology and oceanography and the associated challenges. Since then it has been increasingly applied in sea level studies (e.g., Bruneau et al., 2020; French et al., 2017; Guillou & Chapalain, 2021; Hieronymus et al., 2019; Lai et al., 2019; Lee, 2006; Monahan et al., 2025; Nieves et al., 2021; Pashova & Popova, 2011; Sithara et al., 2020; Tiggeloven et al., 2021; Tur et al., 2021). Machine learning methods advanced so much in ocean science that state-of-the-art machine learning models manage to match or even surpass numerical storm surge models (Hermans et al., 2025; Rus et al., 2023). Interestingly, comparisons between various machine learning methods show remarkably different results. In the study by Bruneau et al. (2020) Long Short-Term Memory (LSTM) networks, which take into account the temporal dependencies of the data, did not provide a significant improvement compared to feed-forward networks for predicting extreme sea levels, but Tiggeloven et al. (2021) found the opposite for a similar problem. Guillou and Chapalain (2021) as well as Balogun and Adebisi (2021) determined that neural networks outperform simpler models such as multilinear regression and support vector machines, while Ayinde et al. (2023) discovered that LSTMs are often outperformed by simpler random forest regression and gradient boosting machine, though they are usually better than multiple linear regression or feed-forward neural networks. This shows that which model performs best highly depends on the location and the temporal scale and cannot be known a *priori*. Model architecture and input data also have a strong influence on the results, and need to be carefully selected depending on the problem at hand.

While many studies use machine learning models, especially neural networks, as “black boxes”, focusing on the prediction without investigating the internal decision-making process of the model, a common end-goal in geoscience is to discover the physical pathways affecting the patterns or relationships in the data (Karpatne et al., 2019). Therefore, there have been attempts to explore the reasoning behind the model's output using various machine learning models and different atmospheric and ocean variables as input (e.g., Balogun & Adebisi, 2021; Nieves et al., 2021; Sithara et al., 2020). Results are varied and region-dependent, with atmospheric variables sometimes driving the sea level prediction (e.g., Balogun & Adebisi, 2021), while in other cases ocean variables play a larger role (e.g., Sithara et al., 2020), further emphasizing the need to have separate regional studies.

European coasts are highly populated (Neumann et al., 2015) and contain many crucial financial and logistic hubs. It is therefore not surprising that sea level on the European continental shelf has attracted considerable scientific attention. Previous studies have investigated the drivers of interannual sea level variability using statistical methods (e.g., Chafik et al., 2017; Dangendorf et al., 2014; Frederikse et al., 2016; Mangini et al., 2021), model sensitivity experiments (e.g., Hermans et al., 2020; Tinker et al., 2020; Wise et al., 2024), and machine learning (Hermans et al., 2025; Heuzé et al., 2025). In this study we focus on the drivers of interannual sea level variability in northern Europe: North Sea, Baltic Sea, the transition zone between them, and the southern part of the Norwegian Sea. Sea level in the southern half of the North Sea rises faster than global average, with main contributions for the rise being ocean warming and glacial isostatic adjustment (Carson et al., 2016). It is also a region with low-lying coasts (Figure 1) and high population density, so even a relatively small sea level rise could cause devastating floods and affect millions of people. The Baltic Sea coasts are not as affected by sea level rise because land uplift from the glacial isostatic adjustment (e.g., Vestøl et al., 2019) compensates for it, but due to the limited transport capacity through the narrow passages that connect it with the rest of the ocean, its dynamics differ from most other coastal seas. The transition zone between the Baltic and the North Sea, consisting of Skagerrak, Kattegat and the Danish Straits, is especially interesting due to its complexity. An advantage of these areas is that they contain many long tide gauge records. Even though the Norwegian Sea contains fewer records, its southern part is included to ensure that the analysis captures the full diversity of the dynamical regimes in the area.

10.1029/2025JH001192

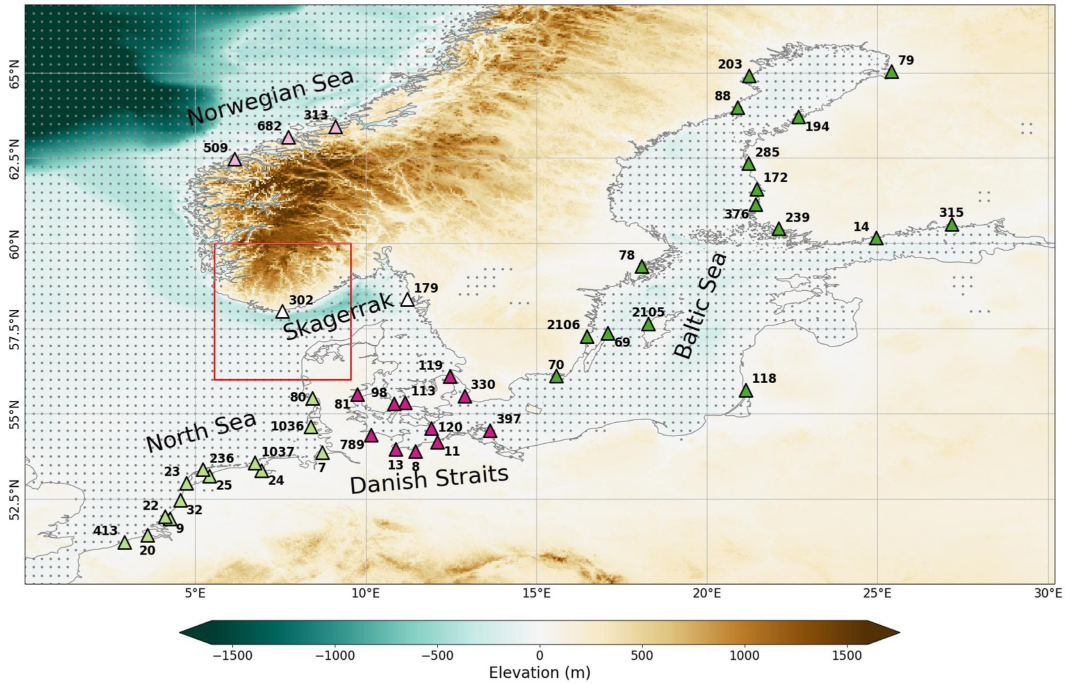

Figure 1. Map of the study area with elevation and bathymetry from the General Bathymetric Chart of the Ocean (GEBCO, 2022). Triangles represent the locations of the tide gauge stations used for the experiments, with their PSMSL station IDs (PSMSL, 2023), separated by color into regions discussed in Section 3.3. Dots show the grid points of the ERA5 data set (Hersbach et al., 2023) for which sea surface temperature is available. Red rectangle represents the area around a sample station over which regional drivers are integrated.

Here we apply a machine learning approach similar to the method used by Heuzé et al. (2025) to quantify the contribution of potential sea level drivers to interannual to decadal sea level variability on a local scale. We train two types of machine learning models, neural networks and linear regression, to simulate sea level based on other variables, and quantify the models' reliance on the input variables with the permutation feature importance method. This paper is organized as follows: Section 2 contains the explanation of the methods applied in this study, description of all used data sets and the outline of the experiment design, Section 3 discusses the results and their implications, followed by the conclusions and outlook in Section 4.

2. Methods and Data

We simulate sea level time series from 45 tide gauges with data available for the period 1960–2016 based on 11 atmospheric, ocean, and hydrological variables as predictors using two types of neural networks and a linear regression model. We choose these two methods to contrast one of the most complex machine learning models with the simplest traditional regression method and to investigate whether using complex methods gives additional insights. We then use the permutation feature importance method to quantify the contribution of each driver to the prediction. In this section we describe the applied machine learning models (Sections 2.1 and 2.2) and the permutation feature importance method (Section 2.3), followed by the descriptions of the data sets and data processing steps (Section 2.4), and the experiment design (Section 2.5).

2.1. Neural Networks

The main machine learning method used in this study is the artificial neural network (NN), a type of supervised learning algorithm that models relationships in data by mimicking how the brain processes information. More detailed explanation of neural networks can be found in, for example, MacKay (2005), Nielsen (2015), Schmidhuber (2015), Yu et al. (2019), Shen (2018), Van Houdt et al. (2020) or Ishida et al. (2020).

The basic feed-forward neural network, sometimes also called multi-layer perceptron, consists of an input layer which contains the input data, one or more hidden layers with neurons and the output layer in which the prediction is made. In machine learning the term prediction refers to the output of a trained model and does not necessarily imply forecasting future states like in ocean science. In each hidden layer the value of a neuron is computed by applying an activation function to the weighted sum of the outputs from the previous layer or, in case of the first hidden layer, to the weighted sum of the input data. Activation function introduces non-linearity and transforms the output into a specific range, allowing the network to model complex relationships in the data (Dubey et al., 2022; MacKay, 2005). The output neuron, representing the predicted value, is calculated as the weighted sum of the neurons in the last hidden layer, without the activation function in regression problems such as this one. Neural networks are trained by adjusting the weights in order to obtain the prediction closest to the target by evaluating the loss function, that is, the function describing the difference between the target variable and the predictions, and backpropagating the changes through all the layers (Rumelhart et al., 1986).

A feed-forward neural network treats each sample as independent. However, samples in our study represent consecutive time steps in climate time series, which are typically not independent; sea level is likely influenced not only by the current ocean and atmospheric state, but also by conditions in preceding months. To capture this temporal dependence, we implement a Long Short-Term Memory (LSTM; Van Houdt et al., 2020) network, which is designed to model sequential relationships in the data and is a commonly used neural network is sea level studies. Since LSTM networks rely on drivers from both the concurrent and the preceding time steps, the number of time steps in the input sequence of an LSTM model is normally referred to as sequence length. However, in this study we define sequence length as the number of preceding months used as input (i.e., one less than in the standard definition), to be able to extend the term to also describe the input depth for feed-forward networks, where the sequence length of zero denotes the use of only concurrent predictors. Since the characteristic timescale of temporal dependence is not known a priori, we use both network types to evaluate multiple sequence lengths (0, 1–6 and 12 months) to capture potentially delayed influences on sea level variability. If sea level relies solely on concurrent information, the feed-forward network is a better choice, whereas if it is influenced by previous states, an LSTM network is required.

For all networks we use the adaptive moment estimation optimizer (Adam; Kingma & Ba, 2014) with starting learning rate of 0.001 and batch size of 1, as they consistently yielded good performance in the initial tests. We also fix the number of layers in the LSTM networks to two, as preliminary tests showed better performance compared to the more typical single-layer setup, while deeper networks significantly increase training time without notable gains in accuracy. The loss function is mean squared error (MSE), as is common for regression problems. Rather than setting the number of training epochs in advance, we use early stopping based on performance on a separate validation set to prevent overfitting. The remaining hyperparameters—number of units and activation function for both network types, and number of layers for the feed-forward networks—are optimized individually for each tide gauge station by performing a grid search and selecting the best hyperparameter combination based on validation set performance. The tested activation functions are sigmoid, hyperbolic tangent, and rectified linear unit; the number of units per layer is selected from 2, 3, 5, 10, 20, 30, 50, 100, 200, and 300; and the feed-forward networks are tested with 1–5 hidden layers. While the sigmoid function generally performs best, optimal numbers of units and layers vary considerably between stations and sequence lengths, highlighting the importance of station-specific hyperparameter tuning.

2.2. Linear Regression With Temporal Dependence

Linear regression (LR) is the simplest supervised machine learning method, often used for sea level (Ayinde et al., 2024). It performs well if the relationship between sea level and its drivers is linear, which is often the case. It has already been applied in the North Sea (e.g., Dangendorf et al., 2014), proving that it can capture a significant portion of the sea level variability in this region. Therefore, we use it alongside neural networks to assess whether the more complex models offer any additional insights into sea level driving mechanisms. To account for temporal dependence, as with the LSTM models, we apply linear regression using sea level drivers from both current and preceding months, with sequence lengths matching those used in the neural networks. The predicted sea level at each time step can then be computed as:

$$\hat{y} = \sum_{k=1}^{D} \sum_{j=0}^{n} w_{kj} x_{kj} + \Theta, \tag{1}$$

10.1029/2025JH001192

where x_{kj} is the k-th driver of the j-th month before the prediction, w_{kj} the corresponding weight, and Θ the bias. The first sum refers to the usual linear regression over D drivers, while the second one accounts for the sea level driver information from the current and preceding months up to sequence length n.

2.3. Permutation Feature Importance

Permutation feature importance is a method used to quantify the contribution of each input variable to the prediction (Fisher et al., 2019). It is simple to use and model-agnostic (Fisher et al., 2019; Mandler & Weigand, 2024), that is, it can be used with any supervised machine learning model regardless of its internal structure, as it is applied after model training, which allows comparison of the results between different model types. It works by permuting each input variable independently to remove its influence on the prediction, generating new predictions with such permuted variables and evaluating the resulting change in model performance, which is typically reduced. The magnitude of the reduction, that is, the difference between the baseline prediction and the prediction with a permuted driver, indicates how strongly the model relies on that variable.

Correlated input variables may influence the attribution results. Because neural networks are trained from different initializations, individual models may achieve similar predictive performance while relying on different correlated predictors. Krell et al. (2025), however, showed that although this increases the uncertainty of the derived attribution, it does not substantially affect its mean value.

It should be noted that permutation feature importance measures the explainability of the model, without considering physical causality, that is, variables correlated with the true physical drivers may be assigned high importance even if they are not directly connected. However, if the potential drivers are selected based on our knowledge of physical processes affecting sea level, and signals that could induce artificial correlations between variables are removed, this method can provide meaningful insight into the relative importance of the considered drivers for sea level variability.

2.4. Data

As target of our models we use monthly mean sea level observed by tide gauges downloaded from Permanent Service for Mean Sea Level (Holgate et al., 2013; PSMSL, 2023). Only time series in the Revised Local Reference (RLR) are used, in which the monthly means have been reduced to a common datum using the tide gauge datum history provided by the supplying authority (Holgate et al., 2013). Figure 1 shows the locations of all tide gauge stations included in this work, together with the bathymetry of our area of interest from GEBCO (2022). We use 45 time series that span the common period across all data sets (1960–2016; 57 years) and contain no gaps or up to 3 years (36 months) of missing data. Additionally, we exclude any monthly means calculated with more than 7 days of missing data. Gaps in the sea level time series do not affect the experiments beyond reducing the number of available samples. Unfortunately, the available in situ time series that meet these criteria are geographically unevenly distributed. There is a high concentration of stations along the southern coast of the North Sea, but only a few along the eastern coast and none on the western. Part of the Norwegian Sea included in our study area contains only three suitable time series. In contrast, Baltic Sea and its transition zone with the North Sea—particularly the complex Danish Straits—are well covered.

The choice of potential drivers is based on a combination of our knowledge of the underlying physical processes and data availability. We also perform a series of preliminary experiments with different combinations of drivers and discard those whose inclusion always results in lower model skill. That is necessary as generally the number of drivers is limited by the number of samples (Alwosheel et al., 2018); if too many inputs are used, the model is unable to generalize and its prediction skill is reduced. Included potential sea level drivers are, broadly speaking, grouped into local, regional, and remote categories based on the spatial scale at which they affect sea level. Local drivers include zonal and meridional wind, mean sea level pressure (MSLP), and sea surface temperature (SST). Wind influences sea level by driving water toward or away from the coast through Ekman transport (e.g., Chafik et al., 2019; Diabaté et al., 2025), local MSLP affects it through the inverse barometer effect (e.g., Hermans et al., 2020), and SST reflects the contribution of thermal expansion. Sea level can also potentially be influenced by adding or removing freshwater through precipitation, evaporation and runoff, which might affect local water density (Piecuch et al., 2018). As the possible influence of precipitation, evaporation and runoff is not limited just to the nearest point but could depend on a larger area, we consider them regional drivers. Finally, we investigate

10.1029/2025JH001192

whether remote processes such as thermal expansion due to global temperature changes, mass loss from Greenland and Antarctica, and the North Atlantic Oscillation (NAO), the dominant climate mode in northern Europe, influence local sea level variability. Although the ice mass loss from glaciers is expected to be more important for northern Europe than Greenland and Antarctica (Carson et al., 2016), we did not manage to find a suitable glaciers mass loss data set that we could include. Based on preliminary experiments we do not include ocean heat content (i.e., vertical ocean temperature profiles) and salinity (Good et al., 2013), as their inclusion consistently reduced models' predictive skill. Even though ocean heat content is more representative for thermal expansion than SST, it is possible that these data sets do not contain enough signal for the models to capture due to limited availability of the observations of the whole water column, especially in the earlier part of the training set period.

All local and regional variables considered in this study, as well as the global SST, representing thermal expansion due to global temperature changes, are from the ERA5 reanalysis data set (Hersbach et al., 2023), a global product with 0.25° spatial resolution covering the whole required time span. The NAO index is downloaded from Climate Prediction Center (2023), while Antarctic runoff (Hansen & Olesen, 2024) is from the surface mass balance model for Antarctica described in Hansen et al. (2021) forced by ERA5 reanalysis data at the lateral boundaries. The Greenland time series is derived from the Greenland ice sheet runoff and discharge data set (Bamber et al., 2018), described in Bamber et al. (2012), which is the shortest of the utilized data sets, available for 1958–2016, and determines the time span in our study. Solid ice discharge from it is also the only variable with yearly temporal resolution, which might influence the results, but we include it regardless as it is the only one available.

Local drivers, that is, MSLP, wind and SST time series, are extracted from the sea grid point in the ERA5 data set closest to each tide gauge location. To account for the potential regional influence and spatial variability of precipitation, evaporation, and runoff, time series for these variables are constructed by integrating over a $4° \times 4°$ box centered around each station (Figure 1). While the exact size and shape of the area of influence most likely varies depending on the location, this box size approximately covers the drainage basins of the larger rivers in the study area (e.g., Rhein River) and the local variability in precipitation without losing signal due to spatial smoothing. There is some overlap in the time series constructed in this manner, for example, signal from the precipitation on land is contained in both the precipitation and the runoff series, but we wish to investigate which of the variables has a stronger influence without going through additional processing to further separate the signals. Using a spatial sum is a somewhat arbitrary choice, but comparison with spatial mean and median showed very similar results, differing only in magnitude, which is irrelevant since all drivers are standardized before model training (see below). Global SST time series is obtained by globally averaging SST from ERA5. The Greenland data set provides discharge and runoff separately for each drainage basin, but we aggregate them to produce a single time series representing the total water input to the ocean from the Greenland ice sheet. We obtain the time series representing the Antarctic contribution by integrating the Antarctic runoff over the whole continent.

Since this study focuses on long-term variability, we remove other dominant signals that could otherwise attract the focus of the machine learning models, specifically, seasonal cycle and trend. Seasonal cycle is removed by subtracting the seasonal mean and dividing by seasonal standard deviation. Linear trend is fitted and subtracted from the time series, which removes both the global mean sea level rise and the vertical land motion signal from the glacial isostatic adjustment present in a large portion of our study area. A segment of the data from 2011 to 2016 is set aside as a test set, while the remaining period (1960–2010) is used for model training and validation. Test set is reserved exclusively for the final evaluation and analysis and is not involved in any part of the model development process, ensuring that the model's performance is assessed on its ability to generalize to truly unseen data, rather than being inadvertently optimized for the test period. This test set makes only 10% of data, which is below the usual recommendations, but is necessary due to the relatively small amount of data and the experiment design (see Section 2.5). Finally, drivers in both the training and the test set are standardized by subtracting the mean and scaling to unit variance. Sea level time series are not scaled in this step, as that is required only for the drivers, but they are nonetheless also scaled to some extent, due to removing the variance of the seasonal cycle, and thus have a similar magnitude across all stations, even though the actual magnitude differs significantly between some areas.

Many of the potential sea level drivers are correlated to each other, as expected from a complex climate system. The strongest correlation is between precipitation and MSLP at most stations, followed by precipitation and one of the wind components. Some stations have additional strongly correlated pairs. Somewhat surprisingly, even though the NAO index is based on atmospheric pressure, its correlation with the local MSLP is relatively low, never exceeding 0.3, which allows us to separately investigate the influences of the dominant mode of large-scale climate variability and the local atmospheric conditions. While neural networks are typically not very sensitive to correlated input, it can affect stability and generalization ability of linear regression (Wilks, 2011). However, methods specifically made for correlated inputs would also cloud the investigation of feature importance, which is the aim of this study. Therefore, we keep all potential drivers and use an ensemble approach (Section 2.5) to mitigate the impact of correlated input variables. Even though feature importance of an individual model, be it neural network or linear regression, can vary due to correlated input, that mainly increases the uncertainty in the ensemble results, not the obtained ranking (Krell et al., 2025).

2.5. Experiment Design

This section describes the experiment setup and the evaluation metrics used in the study. Machine learning models used in this study are very different, but all steps are applied to them equally, unless specified otherwise. Neither of the models applied here have a spatial component; each of the 45 locations used in this study has its own set of neural networks and linear regression models. Separate models are also created for each of the tested sequence lengths.

As target of the machine learning models we use the sea level observations from tide gauges, while other variables described in Section 2.4 are employed as inputs or predictors (Figure 2). Sea level is strongly autocorrelated in both space and time, and many machine learning studies utilize that by including it as a feature, either through past (e.g., Rus et al., 2023; Tur et al., 2021) or concurrent sea level observations from nearby locations (e.g., Hieronymus et al., 2019; Passaro & Juhl, 2023). However, we do not do that in this study because the focus is on its connections to other climate variables and the influence they have. Including past or nearby sea level would most likely improve the predictions, but it would limit our ability to understand the underlying drivers.

In this study, we rely on two evaluation metrics, each serving a different purpose. The first is mean squared error (MSE), defined as:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i}\left(y_i - \widehat{y}_i\right)^2, \tag{2}$$

where y_i are the observations, $\widehat{y}_i$ the predictions, and N the number of samples (i.e., time steps). MSE is used as the loss function during neural network training. It is a standard choice for training regression models, as it directly penalizes larger errors and aligns with the objective of minimizing the difference between predicted and observed values. However, MSE is not ideal for assessing model performance, as its value depends heavily on the range of the target variable, without an upper limit for what would be considered a "good" model (Chicco et al., 2021). Therefore, for the analysis we instead use the coefficient of determination $\left(R^2\right)$, as is common in similar sea level studies (e.g., Richter et al., 2012; Roberts et al., 2016; Wise et al., 2024). To facilitate the evaluation, we multiply R^2 by 100%, obtaining relative explain variance:

$$R^2_{\%} = \left(1 - \frac{\sum_i \left(y_i - \widehat{y}_i\right)^2}{\sum_i \left(y_i - \bar{y}\right)^2}\right) \times 100\%, \tag{3}$$

where $\bar{y}$ is the mean of the observations. Explained variance quantifies the proportion of variance in the observations that is explained by the model. For a perfect model, it would be 100%, while a value of zero or below implies that the model is not able to explain any variability in the data. Since R^2 is monotonically related to MSE, ordering models based on the two metrics is identical (Chicco et al., 2021).

To improve data utilization, increase robustness and stability of the models and capture a range of possible outcomes we employ an ensemble method with cross-validation and bootstrapping. We introduce an additional source of randomness, besides random initialization of weights and biases, by using a different, randomly

JGR: Machine Learning and Computation
10.1029/2025JH001192

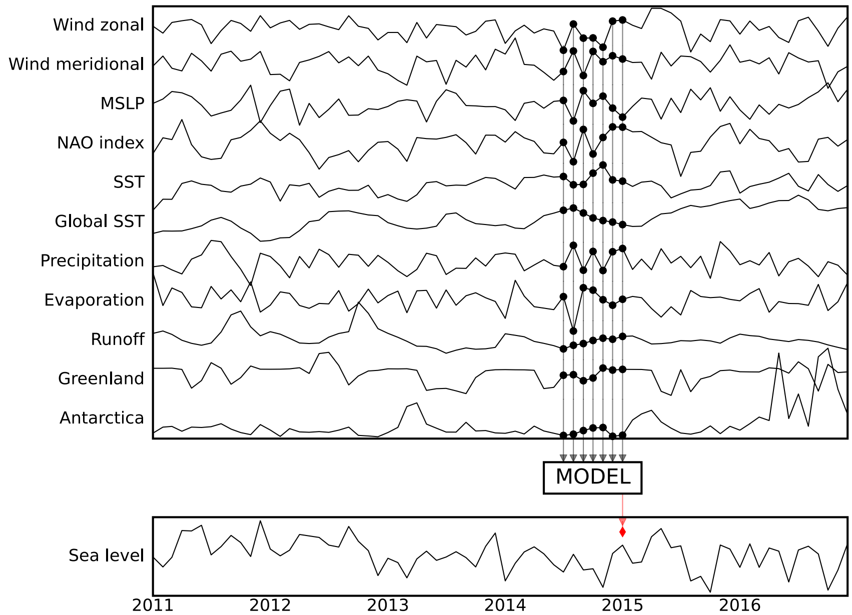

Figure 2. Schematic illustration of the inputs and output of machine learning model with sequence length of 6 months. Sea level prediction (red diamond) is made using drivers from current and six preceding months (black circles). Black lines represent the "true" values for both features and sea level.

selected, 6-year segment of data for validation in each of the ensemble members, with the remaining data used for training. In this way, even though each individual model is trained on a shorter data set, the ensemble has the opportunity to learn about all processes observed in the whole 1960–2010 time span, as the importance of individual drivers can change over time (e.g., Chafik et al., 2019; Wise et al., 2024). A disadvantage of this method is that it introduces some data leakage between the training and the validation set for models with sequence length of one or more, as these models rely on the drivers' information from preceding months, which near the boundary between the sets is used by both. We nevertheless allow it in order to not lose samples, as the data set is already small. To mitigate the impact that correlated inputs have on linear regression and to remove biases that might arise from treating the models differently, we also create linear regression ensembles, with a subset of data corresponding to the validation set removed before fitting each ensemble member.

After preparing the data, we train a small ensemble of neural networks (10 members) with each hyperparameter combination and select the best hyperparameter combination based on the ensemble median explained variance of the validation set. Model performance on the validation set is affected by its time span, which may result in uncharacteristically poor performance that does not necessarily represent general model skill, but rather the characteristics of that particular time span. In a small ensemble such outliers could significantly skew the median and point to an overall worse hyperparameter combination. To avoid that, the three worst performing ensemble members are discarded prior to calculating the median and comparing hyperparameter combinations. We perform this hyperparameter tuning separately for each of the 45 stations and 8 sequence lengths, to make sure that we find the best model for each station, as model performance can vary across different regions, requiring regional adaptation and fine-tuning (Ayinde et al., 2024). The models are then trained using the ensemble method described above with 99 ensemble members, an ensemble size that ensures generalization and robust performance, while the odd number allows direct extraction of the median ensemble member for some analyses. Using the models with the complete set of drivers in the test set produces the baseline predictions—the best prediction these models are able to make. Then, applying the permutation feature importance method, the time series of each driver is permuted to remove its influence, and the test set is re-evaluated, resulting in 11 more predictions. The difference in relative explained variance between the baseline prediction and the prediction with a permuted driver reveals how much that driver contributes to the prediction.

We use Python library scikit-learn (Pedregosa et al., 2011) to prepare the data. All neural networks are created and trained with the Python package for deep learning Tensorflow (Abadi et al., 2015), while the Keras deep learning package (Chollet et al., 2015) is used for early stopping of the training process to avoid overfitting. All computations are performed on a personal computer using only CPU in approximately 2 months.

3. Results and Discussion

3.1. Model Evaluation and Selection

To assess model performance and find the best model at each location, as well as to gain some insight into the nature of the processes governing interannual sea level variability, we analyze the baseline predictions of the test set. We compare neural network and linear regression models, with all tested sequence lengths (0–6, 12), resulting in a total of 16 models per location. Models without temporal input (NN0 and LR0) correspond to feed-forward neural networks and standard linear regression, respectively, while models with higher sequence lengths (NN1–12 and LR1–12) represent LSTM networks and linear regression with temporal dependence. Although the test set covers 2011–2016, only the 2012–2016 span is used for evaluation to ensure consistent comparison across models. Since recurrent models require historical input, predictions can only begin after the initial sequence window, thus excluding 2011 (corresponding to the largest sequence length of 12 months) avoids biases due to unequal data coverage.

Whether neural networks or linear regression perform better depends on several factors. Linear regression typically requires smaller data sets, is less sensitive to noise in the input, and it might generalize better, while neural networks are able to model much more complex relationships but their performance can be affected by initialization and overfitting. Assuming that overfitting risk, initialization effects, and data set size are similar across locations, then which model type performs best most likely reflects the (non-) linearity of the relationships between sea level and its drivers. The optimal sequence length of the best performing model reveals the memory of the processes controlling sea level.

From Figure 3, which shows the relative explained variance of the test set for all models, it is obvious that there is no universally best model type or appropriate sequence length for sea level predictions. At most stations there is some overlap between the ensembles and several locations have multiple virtually identical models. At stations for which several models perform similarly, typically the whole predicted time series is closely matching (not shown), even when the models perform poorly, that is, different models fail in nearly identical way. This suggests that the main limitation stems less from model architecture and more from the input data. To simplify further analyses we focus on the best performing model at each location, determined by the ensemble median of the test set relative explained variance.

There are some distinct spatial patterns in model performance, as well as in the best performing model type and sequence length (Figure 4). Model performance (Figure 4a) is relatively similar across large open areas, with models explaining between 60% and 84% of the observed variability, and typically slightly increases from west to east in each basin. Conversely, in the more complex region characterized by constrained geometry—the Danish Straits—models generally perform worse. With a few exceptions, the interquantile range of the ensembles (Figure 4b) is typically below 3%, demonstrating the consistency in model performance across ensemble members and revealing that small changes in temporal coverage of the training data do not significantly affect it. Best performing model type also exhibits a clear spatial divide (Figures 4c and 4d), with Baltic Sea generally relying on neural networks (always LSTMs), while in most other regions linear regression outperforms them. Previous studies of this region found that the relationship between interannual sea level variability and its main drivers is largely linear (e.g., Chafik et al., 2019; Dangendorf et al., 2014; Frederikse et al., 2016; Hermans et al., 2020; Kolker & Hameed, 2007), but it is nonetheless surprising that linear regression surpasses the skill of the much more complex neural networks to this extent, contrary to many other studies (e.g., Ayinde et al., 2024; Balogun & Adebisi, 2021; Guillou & Chapalain, 2021). At most locations (26) models have a memory of 1 month, 16 include a longer history, and only three rely solely on concurrent data, highlighting the need for recurrent models.

A key challenge in sea level prediction with machine learning is the limited availability and quality of training data (Ayinde et al., 2024). While sea level observations are sufficient, most drivers here rely on reanalysis or model data, which can introduce errors and limit prediction accuracy. Despite this, the models are able to adequately capture the relationships between sea level and its drivers and simulate 70% or more of the observed

Figure 3. Relative explained variance for baseline predictions of the test set (2012–2016) with all models at all stations. Boxes represent the ensembles of each model type and sequence length. Best linear regression and neural network models at each location are indicated using filled boxes and their medians additionally emphasized with horizontal lines. Overall best model at each location is highlighted with a gray background. Each plot is labeled with the station name and ID, as provided by Permanent Service for Mean Sea Level (PSMSL, 2023), together with the name of the best model at that station, and color denoting the region station belongs to.

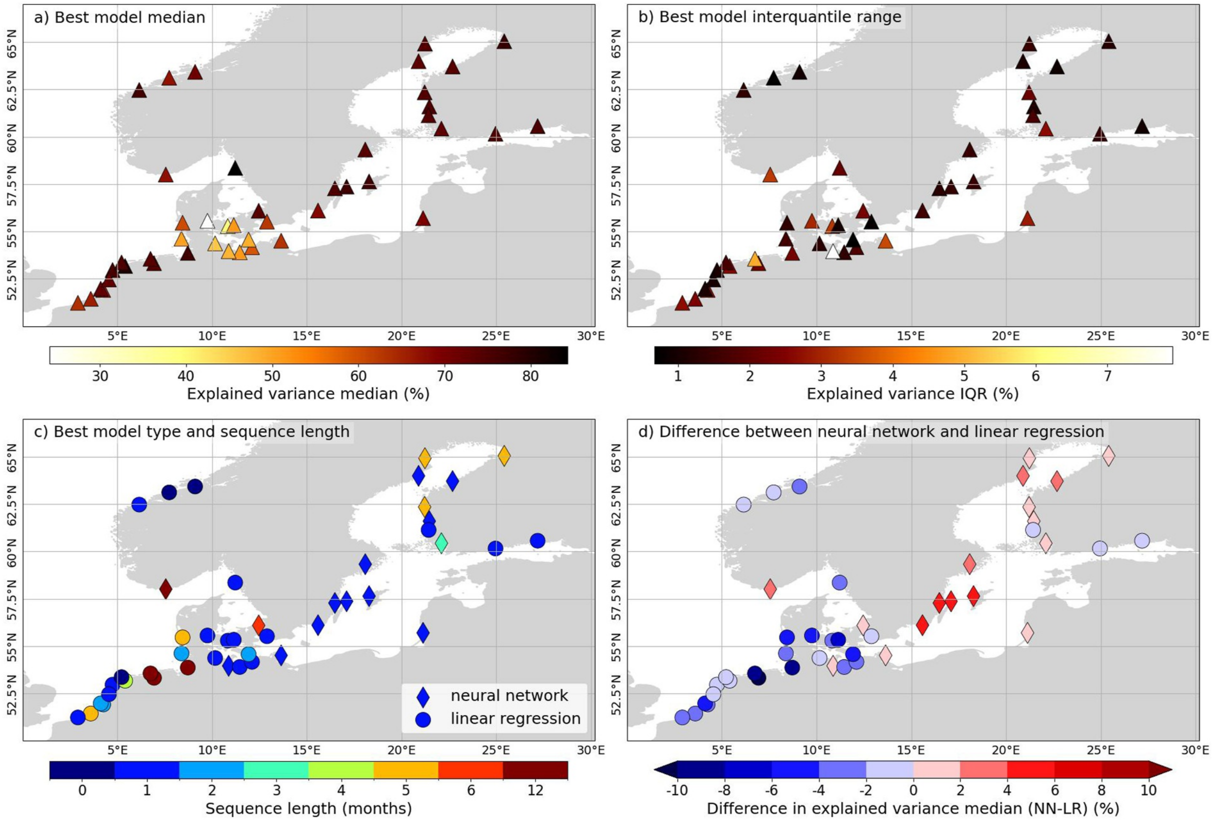

Figure 4. Baseline prediction on the test set (2012–2016) across all locations. (a) Ensemble median and (b) interquantile range of the relative explained variance of the best-performing model at each location. (c) Best-performing model at each location, where marker shape denotes model type and color indicates sequence length. (d) Difference in performance between the best neural network and the best linear regression model (ensemble medians), where positive values indicate locations at which neural networks perform better and negative stations at which linear regression is better, while marker shape again denotes model type.

variability across most of the study area. In areas where models explain less than half of the observed variability this likely reflects that some of the sea level drivers are missing from the input data sets.

3.2. Contribution of the Potential Sea Level Drivers

The median difference between the baseline and the prediction with the permuted driver using the best available model at each station (Figure 5) reveals the potential driver's contribution. It should be noted that this method quantifies the relevance of the drivers to the models' predictions, and does not necessarily reflect the full extent of the complex physical relationship between sea level and its drivers. Contributions of individual drivers are not expected to add up to 100% due to the correlations between them and because they may interact to produce compound effects. The effect of correlated inputs can be seen in the interquantile ranges of the contributions, as individual ensemble members rely on them to different extents depending on the random initialization. While only the best model at each station is shown, permutation feature importance results are typically very similar for all models with sufficient predictive skill at a given location. Even though a few percent of explained variance might seem very low, the models are nonetheless improved by including these drivers. That is especially the case for locations where models are only able to capture around 50% of the variability or less, making a reduction of 2%–5% non-negligible. It is immediately obvious that there are pronounced spatial patterns in drivers' contribution, confirming the unsupervised classification of the region based on interannual sea level variability (Poropat et al., 2024).

Zonal wind (Figure 5a) is by far the most important driver on the southern and eastern coast of the North Sea, most of the transition zone, and the southern part of the Baltic Sea; its removal from the models results in a complete loss of predictive skill (negative explained variances). It also influences sea level in the northern part of the Baltic, where it contributes to around 5%–30% of explained variance. Meridional wind component (Figure 5b) is

10.1029/2025JH001192

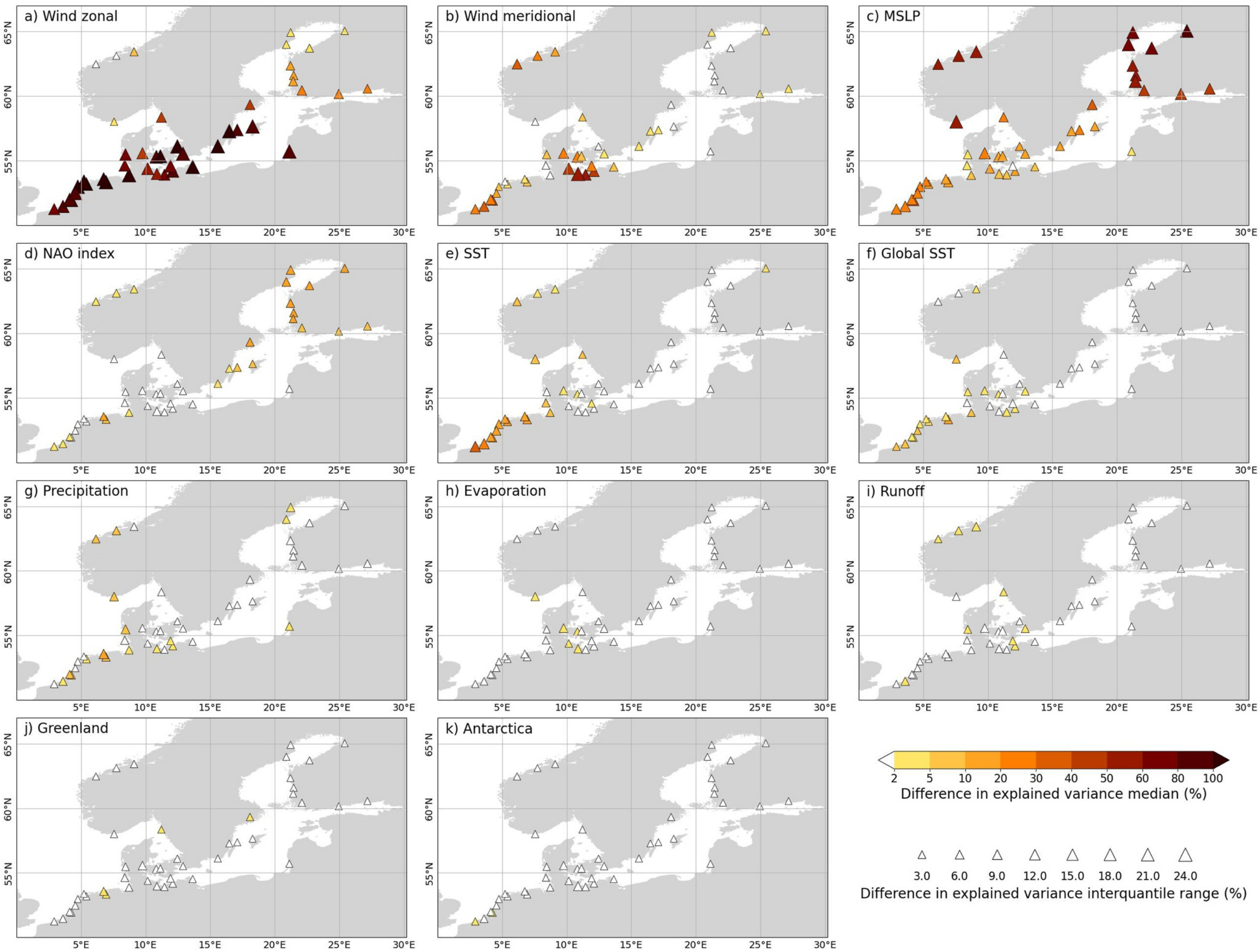

Figure 5. Contribution of each potential sea level driver to sea level prediction with the best machine learning model at each location, expressed as the difference between relative explained variance of the baseline prediction and the relative explained variance of the prediction with the driver permuted. Difference is calculated for the whole ensemble and here shown as median (color) and interquantile range (marker size).

significantly less influential than zonal, but it nevertheless affects sea level at distinct locations. In western Baltic it contributes around 50% of explained variance to the prediction. As the baseline prediction in this area is also only around 50% (Figure 4a), without meridional wind models are not able to capture any sea level variability at all. It is also important in the Norwegian Sea, western part of the North Sea, and in the Danish Straits, where it explains 10%–30% of the variability.

MSLP (Figure 5c) is an important driver of sea level variability on interannual to decadal timescales throughout the whole study area. It is the dominant driver in the northern part of the Baltic Sea, as well as in the Norwegian Sea and at the entrance to Skagerrak, where it contributes between 40 and almost 100%, depending on the location. It also contributes 20%–30% along the western side of the North Sea and 5%–20% at most other locations.

The North Atlantic Oscillation, represented by the NAO index (Figure 5d) is most influential in northern Baltic, where it contributes 10%–20% of explained variance, and its influence decreases going southward and eastward. It also affects sea level along the southern coast of the North Sea, where its influence grows toward the east, as the contribution of the local MSLP diminishes.

It is well established that temperature, that is, thermal expansion, is one of the main mechanisms of sea level rise due to climate change (Fox-Kemper et al., 2023). Its signal on the interannual and decadal timescales is also detected by the machine learning models. Local SST (Figure 5e) mainly affects sea level in the North Sea, especially on its western side, where it typically contributes around 10%–20% to the explained variance of the prediction, exceeding 30% at the westernmost station. It also has some influence in Skagerrak and the Norwegian Sea, but it does not contribute to sea level variability in the Baltic Sea and Danish Straits, most likely because the local water volume increase is negligible due to the shallow water depth. Global surface temperature (Figure 5f) generally contributes less than the local temperature variations, but it nonetheless improves the prediction by 2%–10% along the southern North Sea coast.

Precipitation, evaporation, and runoff represent the possible effects of water fluxes between the atmosphere, ocean and land at a regional scale. Precipitation (Figure 5g) is the most influential of them, contributing up to 20% of relative explained variance at certain locations, mainly in the North Sea, entrance to Skagerrak, and Norwegian Sea. In the North Sea and Skagerrak its influence is most pronounced at stations where models have a longer than average sequence length, while in the Norwegian Sea models rely on precipitation despite including only short or no history, which indicates that the processes connecting precipitation with sea level might be different in the two regions. Evaporation and runoff (Figures 5h and 5i) show no significance for sea level prediction except at a few locations, mainly in the Danish Straits, where their removal deteriorates the prediction by 2%–5%.

The last two remote drivers, the mass loss from Greenland (Figure 5j) and Antarctica (Figure 5k) are not relevant for sea level prediction at these time scales. They contribute only 2%–5% at just a few locations (four and one, respectively), and less than 2% at all others, with no discernible spatial pattern that could explain their contribution. It should be noted that Greenland and Antarctic contributions are represented by single time series, without considering regional patterns through sea-level fingerprints (e.g., Camargo et al., 2022). Since spatial uncertainties associated with these contributions are concentrated close to the source (Camargo et al., 2022), we do not expect this to significantly affect the Antarctic contribution, and our preliminary tests showed that separating the Greenland contribution based on its drainage basin does not impact the results. It is, however, possible that the models are not able to detect a notable signal from these series because of the characteristics of the input data sets, that is, missing Antarctic and only yearly resolution in the Greenland ice discharge.

Overall, we find that sea level variability on the interannual to decadal time scales is predominantly driven by local atmospheric drivers —wind and MSLP— throughout the study area, confirming the findings of previous statistical (Chafik et al., 2019; Sturges & Douglas, 2011) and model sensitivity studies (Hermans et al., 2020; Tinker et al., 2020; Wise et al., 2024). Local SST also has an influence, albeit significantly smaller and limited to specific regions. The only significant remote driver is the NAO, while global SST plays a minor role in some areas. Of the regional drivers, only precipitation has a noticeable contribution at some locations. It should be noted that our models do not include ocean forcings besides SST, nor are they able to properly represent the unforced ocean variability. Earlier studies (e.g., Hermans et al., 2020; Wise et al., 2024) have shown that on the northwestern European continental shelf these contributions are minor compared to atmospheric forcing, but Danish Straits were not part of their study domains, so it is possible that ocean forcing plays a more important role there than on the rest of the shelf (Section 3.3).

3.3. Regional Analysis

We find that sea level is driven by spatially variable mechanisms, with our study area split into four distinct regions with similar patterns: Norwegian Sea, North Sea, Baltic Sea, and the Danish Straits, including the adjacent western part of the Baltic Sea. Additionally, Skagerrak is considered separately because it differs from the regions it is connected to, but since the two tide gauge stations it contains exhibit pronounced differences, it is not a coherent region in terms of sea level.

In the part of the Norwegian Sea included into our study sea level is mainly driven by fluctuations in MSLP, confirming the findings from Richter et al. (2012). We also discover a notable influence from meridional wind, and minor contributions from precipitation and local SST. Meridional wind influence decreases from southwest to northeast, as zonal wind impact increases, reflecting the coastal orientation and the alongshore winds driving upwelling and raising local sea level (Chafik et al., 2019; Sturges & Douglas, 2011). At all three locations linear regression outperforms neural networks. It is also worth noting that two out of only three stations at which the best model relies solely on the concurrent drivers are located here, while at the third the difference between LR0 and LR1 is small (Figure 3; ID: 509). This suggests that the processes driving sea level in this area are predominantly linear and generally not impacted by the past conditions. As the models are able to explain 67%–76% of observed variability, they likely capture the main sea level driving mechanisms in the region.

The North Sea consists of two regions with distinctly different sea level driving regimes, with wind dominating on the south and east and MSLP controlling the west and north (Dangendorf et al., 2014; Hermans et al., 2020). Here we focus on the former, while the one station in our study belonging to the MSLP regime (Tregde; ID: 302) is discussed later, as it is geographically located within Skagerrak. Along the southern coast models explain 63%–81% of sea level variability, with their predictive skill generally increasing from west to east, while at the two stations on the eastern coast model skill is somewhat lower (50% and 60%). At all stations the best model is linear regression, often significantly outperforming neural networks, suggesting that the underlying dynamics controlling sea level in this area are predominantly linear. At virtually all locations sea level depends on both present and past state of the atmosphere and ocean, with western locations relying on 1–5 and eastern on up to 12 previous months. Dominant driver is zonal wind, followed by substantial contributions from meridional wind and MSLP. NAO and SST, both local and globally averaged, also play a role. Our findings agree with previous studies about the importance of atmospheric forcing (Plag & Tsimplis, 1999; Roberts et al., 2016; Tinker et al., 2020) and specifically wind (Dangendorf et al., 2014; Diabaté et al., 2025; Gerkema & Duran-Matute, 2017; Hermans et al., 2020; Wise et al., 2024) for driving the interannual sea level variability in the North Sea, as well as with Dangendorf et al. (2013) about the characteristics of sea level variability in Cuxhaven (ID: 7). Our results generally agree better with studies based on ocean model sensitivity experiments than those using statistical methods. Statistical methods typically consider only concurrent relationships between sea level and its drivers, which indicate only zonal wind as important in this region (Dangendorf et al., 2013). Our approach, similar to ocean modeling studies (e.g., Hermans et al., 2020), incorporates the memory of the system, enabling the detection of delayed atmospheric effects, which reveals a significant contribution from meridional wind and MSLP in southern North Sea. Relative importance of the two wind components is again linked to coastal orientation, most likely detecting coastal upwelling due to Ekman transport, which plays an important role for sea level variability on the European continental shelf (Chafik et al., 2017; Hermans et al., 2020; Mangini et al., 2021). The minor role NAO plays agrees well with previous studies (Chafik et al., 2017; Kolker & Hameed, 2007; Wahl et al., 2013; Wakelin et al., 2003). Lastly, consistent with Dangendorf et al. (2014), our analysis confirms a dependence of sea level on ocean temperature.

There is a distinct difference between the western Baltic Sea and the rest of the basin, consistent with Poropat et al. (2024), who found that sea level there is more similar to the Danish Straits than to the broader Baltic. Here we focus on the bulk of the Baltic Sea, while western Baltic is discussed later, together with the Danish Straits. Models manage to achieve exceptionally good performance in the Baltic Sea, with relative explained variances consistently increasing from southwest (63%) to northeast (78%). At most locations neural networks perform better than linear regression, indicating that sea level is driven by non-linear processes. Prediction is also markedly improved when data from the previous month is included (Figure 3), revealing that these processes are slow. As Pham et al. (2024), we find that Baltic sea level is mainly driven by local atmospheric forcing, with a smaller contribution from NAO. We also discover a difference between the north, which is mainly driven by MSLP, and the south, where contribution from zonal wind prevails. Even though the Baltic Sea has a very low salinity due to the influx of freshwater from river runoff, we find no notable influence on sea level from said runoff, consistent with Samuelsson and Stigebrandt (1996). This is not that surprising, as sea level in the Baltic depends significantly more on how much water passes through the Danish Straits (Gräwe et al., 2019). Unlike Hünicke and Zorita (2006), we do not find any contribution from SST and precipitation, possibly because their influence is season-dependent, so our models are not able to detect it from the deseasoned input time series. While NAO is a good indicator for the influence of the atmospheric circulation on the Baltic sea level variability, it is most likely not the only relevant atmospheric mode (Weisse et al., 2021). External forcing via in- and outflows through the Danish Straits is also affected by regional circulation (e.g., Andersson, 2002; Karabil et al., 2018; Samuelsson & Stigebrandt, 1996), so including MSLP or wind at the basin mouth might further improve the prediction.

The Danish Straits, a complex system of narrow, shallow passages connecting Kattegat with the western Baltic, are the most complex and varied region in our area of interest. Zonal wind is the dominant sea level driver in the region, as found by Passaro et al. (2015), followed by a considerable contribution from MSLP. Additionally, its southern part, especially the western Baltic, is strongly affected by meridional wind, as expected given that the north–south aligned Danish Straits lead directly into the north-facing western Baltic coast. Preceding month's atmospheric conditions significantly affect the prediction (Figure 3), reflecting the inertia of the processes involved. There is a notable east–west gradient in model skill. While on the east side the models are able to explain 62%–74% of the observed variability, comparable with the rest of the study area, model skill significantly drops moving westward, explaining only 24%–51%. This suggests that sea level in the Danish Straits, particularly on their western side, is affected by additional processes not included in the input data sets, and thus not incorporated into the models. As the straits connect the brackish Baltic Sea with the much saltier North Sea, water in them is characterized by large differences in salinity, which could affect sea level. Major Baltic inflows, that is, events during which large volumes of dense salty water enter the Baltic Sea through the Danish Straits, temporarily alter circulation and stratification, which can also affect sea level. They are driven by specific atmospheric conditions over the Baltic and the North Seas (e.g., Löptien et al., 2025; Pham et al., 2024), which are not necessarily visible in the local wind and MSLP time series. Furthermore, they are short-lived, typically lasting only days to weeks, so monthly mean time series most likely do not fully capture this dramatic shift of the conditions. This is to some extent confirmed by Heuzé et al. (2025), who found no such decrease in predictive skill for high-frequency variability in the Danish Straits. Linear regression works better at most locations in the straits, but here this may be a consequence of low signal-to-noise ratio, indicated by the consistently poor performance of all models, rather than due to linearity of the relationships between the variables. Complex models such as neural networks are more prone to overfitting when the available signal is weak. The best performing model in the region (Hornbaek; ID: 119) being a neural network also points toward that conclusion.

Finally, Skagerrak contains only two tide gauge stations, Tregde (ID: 302) bordering the North Sea, and Smogen (ID: 179) on the opposite side, with markedly different driving mechanisms. Sea level in Tregde is mainly driven by MSLP, with a small contribution from temperature (both local and global), and precipitation. This generally aligns with the results from Dangendorf et al. (2014), who found that Tregde belongs to the MSLP-driven regime of the North Sea. The best model in Tregde is NN12, which indicates that the processes driving sea level are predominantly non-linear and have a very long memory. On the other end of the channel, in Smogen, sea level is mainly controlled by zonal wind, with MSLP playing only a secondary role. Considering that the best model is LR1, explaining striking 84% variability, the processes affecting it are almost exclusively linear, with only a minor dependence on the past state.

4. Conclusions

In this study we use machine learning to quantify the contribution of the potential drivers of interannual to decadal sea level variability in the coastal seas of northern Europe. We apply neural networks and linear regression to predict monthly mean sea level at 45 tide gauge locations in the Baltic and North Sea, the Baltic-North Sea transition zone and the southern part of the Norwegian Sea from atmospheric, ocean and hydrological variables. We then use permutation feature importance to identify which drivers the models rely on and quantify their contribution, thereby revealing the key factors driving the local sea level. Analyzing the best model type and the sequence length it uses indicates the structural and temporal characteristics of the relationships between sea level and its drivers.

We identify four regions with broadly similar sea level driving mechanisms: Norwegian, North, and Baltic Seas, and the Danish Straits (including western Baltic), as well as Skagerrak, an area separate from others but not internally consistent. In the Norwegian Sea sea level variability is mainly driven by MSLP with some influence from the meridional wind and precipitation, through largely linear and concurrent processes. In the majority of the North Sea, sea level is predominantly driven by zonal wind, with a significant contribution from meridional wind and MSLP and a minor influence from temperature. As in the Norwegian Sea, the relationships between sea level and its drivers are predominantly linear, but there is a strong dependence on the prior states of the drivers. The processes driving Baltic Sea level are largely nonlinear and relatively slow, and the main driving mechanism is the local MSLP (north) and zonal wind (south), followed by the NAO. The Danish Straits are affected strongly by zonal wind, meridional wind in the south, and by MSLP, and typically there is some memory in the system. In the western part of the Danish Straits, models are able to explain less than half of the observed variability, indicating that a significant portion of it is driven by processes not included into the models. Finally, the two locations in Skagerrak differ both from the adjacent regions and from each other, with sea level on the North Sea side being mainly driven by MSLP through predominantly nonlinear and slow processes, while on the other end largely linear forcing by zonal wind dominates, with only a small dependence on the previous month.

Overall, the results of this study support four main conclusions: (a) sea level variability is driven by a range of dynamic processes that differ significantly across regions, leading to distinct spatial patterns; (b) there is no single best machine learning model for this purpose, sometimes complex models are needed, while other times linear regression works better; (c) choosing adequate predictors is even more important than model selection; and (d) in northern Europe sea level on the interannual to decadal timescales is mainly driven by local atmospheric pressure and wind, while remote drivers such as global mean SST and NAO have a more limited impact. A comprehensive study across 45 locations using two vastly different machine learning methods and 11 potential sea level drivers allows us to provide new insights into the sea level driving mechanisms along the northern European coasts. Contrasting linear regression and neural networks allows us to better understand the nature of the relationships between sea level and its drivers, while comparing different sequence lengths in the models indicates the response time. Using permutation feature importance to uncover the model dependence on the individual inputs reveals the sea level drivers' contribution. Our data-driven approach confirms the influence of meridional wind in driving the North Sea variability, which has previously been seen only in model sensitivity experiments, and reveals the split between the predominantly wind driven south and MSLP driven north of the Baltic Sea. Although neural networks are generally regarded as superior and increasingly applied in ocean and climate research, our results demonstrate that sometimes linear regression performs as well or even better. While neural networks are able to represent far more complex relationships in the data, if the simulated processes are predominantly linear, and particularly if the predictive signal is weak or data are limited, neural networks can introduce unnecessary complexity without improving predictive skill and may increase the risk of overfitting. This highlights the importance of applying multiple machine learning methods, including the simple ones, to obtain the best predictions and a more comprehensive understanding of the studied processes. Finally, our feature importance experiments, and especially the poor model performance in the Danish Straits, demonstrate the importance of selecting appropriate drivers for sea level prediction.

The need for different methods, network architectures, and input sequence lengths highlights the complexity of sea level variability and its relationships with other variables, even across relatively small geographical distances. Given this complexity, it is unsurprising that low-resolution climate models struggle to accurately represent coastal sea level variability, underscoring the benefits of data-driven approaches to complement the traditional climate modeling. These results can be used to improve local sea level projections by downscaling climate model projections to the locations included into the study, identify the most endangered areas and help us determine the best ways to protect the Baltic and the North Sea coastlines from sea level changes. The same methodology can be applied to investigate sea level changes in other regions of the world or to continue the study from Heuzé et al. (2025) focused on the high-frequency signals to reveal future risks of flooding caused by storm surges.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Availability Statement

Data sets used in this study are publicly available from various sources, all listed and described in Section 2.4. Tide gauge data is available from https://psmsl.org/ (PSMSL, 2023), ERA5 data can be obtained from https://doi.org/10.24381/cds.f17050d7 (Hersbach et al., 2023), the Greenland data set is available at https://doi.org/10.5285/643aa9bc-bcd6-45ad-e053-6c86abc07da0 (Bamber et al., 2018), Antarctic data set at https://download.dmi.dk/Research_Projects/PROTECT/HIRHAM5_ERA5_ANT/PPdata/ (Hansen & Olesen, 2024), NAO index at https://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/nao.shtml (Climate Prediction Center, 2023), and the GEBCO bathymetry used for creating Figure 1 can be found at https://doi.org/10.5285/e0f0bb80-ab44-2739-e053-6c86abc0289c (GEBCO, 2022). Code created for this study is available at https://doi.org/10.5281/zenodo.19070975 (Poropat, 2026a), while neural networks created in this study, input data, and results of the study can be found at https://doi.org/10.5281/zenodo.19204512 (Poropat, 2026b).

Acknowledgments
This work is part of the project "Would the Northern European Enclosure Dam really protect Sweden from sea level rise? (NEEDS)" funded by FORMAS through a Grant awarded to Céline Heuzé (dnr 2020-00982). Lea Poropat was funded by the Danish state through the National Center for Climate Research (NCKF) and the Danish Climate Atlas. We thank the three anonymous reviewers for their insightful comments and constructive feedback, which helped improve the quality of this work.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., et al. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems [Software]. Retrieved from https://www.tensorflow.org

Alwosheel, A., van Cranenburgh, S., & Chorus, C. G. (2018). Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis. *Journal of Choice Modelling*, *28*, 167–182. https://doi.org/10.1016/j.jocm.2018.07.002

Andersson, H. C. (2002). Influence of long-term regional and large-scale atmospheric circulation on the Baltic Sea level. *Tellus A: Dynamic Meteorology and Oceanography*, *54*(1), 76–88. https://doi.org/10.3402/tellusa.v54i1.12125

Ayinde, A. S., Huaming, Y., & Kejian, W. (2024). Review of machine learning methods for sea level change modeling and prediction. *Science of the Total Environment*, *954*, 176410. https://doi.org/10.1016/j.scitotenv.2024.176410

Ayinde, A. S., Yu, H., & Wu, K. (2023). Sea level variability and modeling in the Gulf of Guinea using supervised machine learning. *Scientific Reports*, *13*(21318), 21318. https://doi.org/10.1038/s41598-023-48624-1

Balogun, A.-L., & Adebisi, N. (2021). Sea level prediction using ARIMA, SVR and LSTM neural network: Assessing the impact of ensemble ocean-atmospheric processes on models' accuracy. *Geomatics, Natural Hazards and Risk*, *12*(1), 653–674. https://doi.org/10.1080/19475705.2021.1887372

Bamber, J., Tedstone, A., King, M., Howat, I., Enderlin, E., van den Broeke, M., & Noel, B. (2018). Freshwater fluxes for land ice in the Arctic and N Atlantic, 1958–2016 [Dataset]. *British Oceanographic Data Centre—Natural Environment Research Council*. https://doi.org/10.5285/643aa9bc-bcd6-45ad-e053-6c86abc07da0

Bamber, J., van den Broeke, M., Ettema, J., Lenaerts, J., & Rignot, E. (2012). Recent large increases in freshwater fluxes from Greenland into the North Atlantic. *Geophysical Research Letters*, *39*(19). https://doi.org/10.1029/2012GL052552

Bruneau, N., Polton, J., Williams, J., & Holt, J. (2020). Estimation of global coastal sea level extremes using neural networks. *Environmental Research Letters*, *15*(7), 074030. https://doi.org/10.1088/1748-9326/ab89d6

Camargo, C. M. L., Riva, R. E. M., Hermans, T. H. J., & Slangen, A. B. A. (2022). Trends and uncertainties of mass-driven sea-level change in the satellite altimetry era. *Earth System Dynamics*, *13*(3), 1351–1375. https://doi.org/10.5194/esd-13-1351-2022

Carson, M., Köhl, A., Stammer, D., Slangen, A. B. A., Katsman, C. A., van de Wal, R. S. W., et al. (2016). Coastal sea level changes, observed and projected during the 20th and 21st century. *Climatic Change*, *134*(1–2), 269–281. https://doi.org/10.1007/s10584-015-1520-1

Cazenave, A., & Le Cozannet, G. (2014). Sea level rise and its coastal impacts. *Earth's Future*, *2*(2), 15–34. https://doi.org/10.1002/2013EF000188

Chafik, L., Nilsen, J. E. i., & Dangendorf, S. (2017). Impact of North Atlantic teleconnection patterns on Northern European sea level. *Journal of Marine Science and Engineering*, *5*(3), 43. https://doi.org/10.3390/jmse5030043

Chafik, L., Nilsen, J. E. I., Dangendorf, S., Reverdin, G., & Frederikse, T. (2019). North Atlantic Ocean circulation and decadal sea level change during the altimetry era. *Scientific Reports*, *9*(1041), 1041. https://doi.org/10.1038/s41598-018-37603-6

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, *7*, e623. https://doi.org/10.7717/peerj-cs.623

Chollet, F., et al. (2015). Keras [Software]. Retrieved from https://keras.io

Climate Prediction Center (CPC). (2023). Monthly mean NAO index since January 1950 [Dataset]. *National Oceanic and Atmospheric Administration*. Retrieved from https://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/nao.shtml

Dangendorf, S., Calafat, F. M., Arns, A., Wahl, T., Haigh, I. D., & Jensen, J. (2014). Mean sea level variability in the North Sea: Processes and implications. *Journal of Geophysical Research: Oceans*, *119*(10). https://doi.org/10.1002/2014JC009901

Dangendorf, S., Mudersbach, C., Wahl, T., & Jensen, J. (2013). Characteristics of intra-inter-annual and decadal sea-level variability and the role of meteorological forcing: The long record of Cuxhaven. *Ocean Dynamics*, *63*(2–3), 209–224. https://doi.org/10.1007/s10236-013-0598-0

Diabaté, S. T., Fraser, N. J., White, M., Berx, B., Marié, L., & McCarthy, G. D. (2025). On the wind-driven European shelf sea-level variability and the associated oceanic circulation. *Continental Shelf Research*, *291*, 105466. https://doi.org/10.1016/j.csr.2025.105466

Dubey, S. R., Singh, S. K., & Chaudhuri, B. B. (2022). Activation functions in deep learning: A comprehensive survey and benchmark. *Neurocomputing*, *503*, 92–108. https://doi.org/10.1016/j.neucom.2022.06.111

Durand, G., van den Broeke, M. R., Le Cozannet, G., Edwards, T. L., Holland, P. R., Jourdain, N. C., et al. (2022). Sea-level rise: From global perspectives to local services. *Frontiers in Marine Science*, *8*, 709595. https://doi.org/10.3389/fmars.2021.709595

Fisher, A., Rudin, C., & Dominici, F. (2019). All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously. *Journal of Machine Learning Research*, *20*(177), 1–81. Retrieved from http://jmlr.org/papers/v20/18-760.html

Fox-Kemper, B., Hewitt, H., Xiao, C., Aðalgeirsdóttir, G., Drijfhout, S., Edwards, T., et al. (2023). Ocean, cryosphere and sea level change. In V. Masson-Delmotte (Ed.), *Climate change 2021: The physical science basis. Contribution of working Group I to the sixth assessment report of the intergovernmental panel on climate change* (pp. 1211–1362). Cambridge University Press. https://doi.org/10.1017/9781009157896.011

Frederikse, T., Riva, R., Kleinherenbrink, M., Wada, Y., van den Broeke, M., & Marzeion, B. (2016). Closing the sea level budget on a regional scale: Trends and variability on the Northwestern European continental shelf. *Geophysical Research Letters*, *43*(20), 10864–10872. https://doi.org/10.1002/2016GL070750

French, J., Mawdsley, R., Fujiyama, T., & Achuthan, K. (2017). Combining machine learning with computational hydrodynamics for prediction of tidal surge inundation at estuarine ports. *Procedia IUTAM*, *25*, 28–35. (IUTAM Symposium on Storm Surge Modelling and Forecasting). https://doi.org/10.1016/j.piutam.2017.09.005

GEBCO Compilation Group. (2022). GEBCO_2022 grid [Dataset]. https://doi.org/10.5285/e0f0bb80-ab44-2739-e053-6c86abc0289c

Gerkema, T., & Duran-Matute, M. (2017). Interannual variability of mean sea level and its sensitivity to wind climate in an inter-tidal basin. *Earth System Dynamics*, *8*(4), 1223–1235. https://doi.org/10.5194/esd-8-1223-2017

Glavovic, B., Dawson, R., Chow, W., Garschagen, M., Haasnoot, M., Singh, C., & Thomas, A. (2023). Cities and settlements by the sea. In *Climate change 2022 – Impacts, adaptation and vulnerability: Working Group II contribution to the sixth assessment report of the intergovernmental panel on climate change* (pp. 2163–2194). Cambridge University Press. https://doi.org/10.1017/9781009325844.019

Good, S. A., Martin, M. J., & Rayner, N. A. (2013). EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates [Dataset]. *Journal of Geophysical Research: Oceans*, *118*(12), 6704–6716. https://doi.org/10.1002/2013JC009067

Gräwe, U., Klingbeil, K., Kelln, J., & Dangendorf, S. (2019). Decomposing mean sea level rise in a semi-enclosed basin, the Baltic Sea. *Journal of Climate*, *32*(11), 3089–3108. https://doi.org/10.1175/JCLI-D-18-0174.1

Guillou, N., & Chapalain, G. (2021). Machine learning methods applied to sea level predictions in the upper part of a tidal estuary. *Oceanologia*, *63*(4), 531–544. https://doi.org/10.1016/j.oceano.2021.07.003

Haigh, I. D., Wahl, T., Rohling, E. J., Price, R. M., Pattiaratchi, C. B., Calafat, F. M., & Dangendorf, S. (2014). Timescales for detecting a significant acceleration in sea level rise. *Nature Communications*, *5*(3635), 3635. https://doi.org/10.1038/ncomms4635

Hansen, N., Langen, P. L., Boberg, F., Forsberg, R., Simonsen, S. B., Thejll, P., et al. (2021). Downscaled surface mass balance in Antarctica: Impacts of subsurface processes and large-scale atmospheric circulation. *The Cryosphere*, *15*(9), 4315–4333. https://doi.org/10.5194/tc-15-4315-2021

Hansen, N., & Olesen, M. (2024). HIRHAM5 (forced by ERA5) model output for Antarctica [Dataset]. Retrieved from https://download.dmi.dk/Research_Projects/PROTECT/HIRHAM5_ERA5_ANT/PPdata/

Hermans, T. H. J., Ben Hammouda, C., Treu, S., Tiggeloven, T., Couasnon, A., Busecke, J. J. M., & van de Wal, R. S. W. (2025). Computing extreme storm surges in Europe using neural networks. *Natural Hazards and Earth System Sciences*, *25*(11), 4593–4612. https://doi.org/10.5194/nhess-25-4593-2025

Hermans, T. H. J., Le Bars, D., Katsman, C. A., Camargo, C. M. L., Gerkema, T., Calafat, F. M., et al. (2020). Drivers of interannual sea level variability on the Northwestern European shelf. *Journal of Geophysical Research: Oceans*, *125*(10), e2020JC016325. https://doi.org/10.1029/2020JC016325

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horanyi, A., Muñoz Sabater, J., et al. (2023). ERA5 monthly averaged data on single levels from 1940 to present [Dataset]. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. https://doi.org/10.24381/cds.f17050d7

Heuzé, C., Carlstedt, L., Poropat, L., & Reese, H. (2025). Drivers of high-frequency extreme sea levels around northern Europe – Synergies between recurrent neural networks and random forest. *Ocean Science*, *21*(4), 1813–1832. https://doi.org/10.5194/os-21-1813-2025

Hieronymus, M., Hieronymus, J., & Hieronymus, F. (2019). On the application of machine learning techniques to regression problems in sea level studies. *Journal of Atmospheric and Oceanic Technology*, *36*(9), 1889–1902. https://doi.org/10.1175/JTECH-D-19-0033.1

Holgate, S. J., Matthews, A., Woodworth, P. L., Rickards, L. J., Tamisiea, M. E., Bradshaw, E., et al. (2013). New data systems and products at the Permanent Service for Mean Sea Level. *Journal of Coastal Research*, *29*(3), 493–504. https://doi.org/10.2112/JCOASTRES-D-12-00175.1

Hsieh, W. W., & Tang, B. (1998). Applying neural network models to prediction and data analysis in meteorology and oceanography. *Bulletin of the American Meteorological Society*, *79*(9), 1855–1870. https://doi.org/10.1175/1520-0477(1998)079<1855:ANNMTP>2.0.CO;2

Hünicke, B., & Zorita, E. (2006). Influence of temperature and precipitation on decadal Baltic Sea level variations in the 20th century. *Tellus A: Dynamic Meteorology and Oceanography*, *58*(1), 141–153. https://doi.org/10.1111/j.1600-0870.2006.00157.x

Ishida, K., Tsujimoto, G., Ercan, A., Tu, T., Kiyama, M., & Amagasaki, M. (2020). Hourly-scale coastal sea level modeling in a changing climate using long short-term memory neural network. *Science of the Total Environment*, *720*, 137613. https://doi.org/10.1016/j.scitotenv.2020.137613

Karabil, S., Zorita, E., & Hünicke, B. (2018). Contribution of atmospheric circulation to recent off-shore sea-level variations in the Baltic Sea and the North Sea. *Earth System Dynamics*, *9*(1), 69–90. https://doi.org/10.5194/esd-9-69-2018

Karpatne, A., Ebert-Uphoff, I., Ravela, S., Babaie, H. A., & Kumar, V. (2019). Machine learning for the geosciences: Challenges and opportunities. *IEEE Transactions on Knowledge and Data Engineering*, *31*(8), 1544–1554. https://doi.org/10.1109/TKDE.2018.2861006

Kingma, D., & Ba, J. (2014). Adam: A method for stochastic optimization. In *International Conference on learning representations*. https://doi.org/10.48550/arXiv.1412.6980

Kolker, A. S., & Hameed, S. (2007). Meteorologically driven trends in sea level rise. *Geophysical Research Letters*, *34*(23). https://doi.org/10.1029/2007GL031814

Krell, E., Mamalakis, A., King, S. A., Tissot, P., & Ebert-Uphoff, I. (2025). The influence of correlated features on neural network attribution methods in geoscience. *Environmental Data Science*, *4*, e29. https://doi.org/10.1017/eds.2025.19

Lai, V., Ahmed, A. N., Malek, M., Abdulmohsin Afan, H., Ibrahim, R. K., El-Shafie, A., & El-Shafie, A. (2019). Modeling the nonlinearity of sea level oscillations in the Malaysian coastal areas using machine learning algorithms. *Sustainability*, *11*(17), 4643. https://doi.org/10.3390/su11174643

Lee, T.-L. (2006). Neural network prediction of a storm surge. *Ocean Engineering*, *33*(3), 483–494. https://doi.org/10.1016/j.oceaneng.2005.04.012

Löptien, U., Renz, M., & Dietze, H. (2025). Major Baltic inflows come in different flavours. *Communications Earth & Environment*, *6*(235), 232. https://doi.org/10.1038/s43247-025-02209-0

MacKay, D. J. (2005). *Information theory, inference, and learning algorithms*. Cambridge University Press.

Mandler, H., & Weigand, B. (2024). A review and benchmark of feature importance methods for neural networks. *ACM Computing Surveys*, *56*(12), 1–30. https://doi.org/10.1145/3679012

Mangini, F., Chafik, L., Madonna, E., Li, C., Bertino, L., & Nilsen, J. E. Ø. (2021). The relationship between the eddy-driven jet stream and Northern European sea level variability. *Tellus A: Dynamic Meteorology and Oceanography*, *73*(1), 1886419. https://doi.org/10.1080/16000870.2021.1886419

Monahan, T., Tang, T., Roberts, S., & Adcock, T. A. A. (2025). RTide: Automating the tidal response method. *Journal of Geophysical Research: Machine Learning and Computation*, *2*(2), e2024JH000525. https://doi.org/10.1029/2024JH000525

Neumann, B., Vafeidis, A. T., Zimmermann, J., & Nicholls, R. J. (2015). Future coastal population growth and exposure to sea-level rise and coastal flooding—A global assessment. *PLoS One*, *10*(3), 1–34. https://doi.org/10.1371/journal.pone.0118571

Nielsen, M. A. (2015). *Neural networks and deep learning*. Determination Press.

Nieves, V., Radin, C., & Camps-Valls, G. (2021). Predicting regional coastal sea level changes with machine learning. *Scientific Reports*, *11*(7650), 7650. https://doi.org/10.1038/s41598-021-87460-z

Pashova, L., & Popova, S. (2011). Daily sea level forecast at tide gauge Burgas, Bulgaria using artificial neural networks. *Journal of Sea Research*, *66*(2), 154–161. https://doi.org/10.1016/j.seares.2011.05.012

Passaro, M., Cipollini, P., & Benveniste, J. (2015). Annual sea level variability of the coastal ocean: The Baltic Sea-North Sea transition zone. *Journal of Geophysical Research: Oceans*, *120*(4), 3061–3078. https://doi.org/10.1002/2014JC010510

Passaro, M., & Juhl, M.-C. (2023). On the potential of mapping sea level anomalies from satellite altimetry with random forest regression. *Ocean Dynamics*, *73*(2), 107–116. https://doi.org/10.1007/s10236-023-01540-4

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in python. [Software]. *Journal of Machine Learning Research*, *12*, 2825–2830. Retrieved from https://scikit-learn.org/

Permanent Service for Mean Sea Level (PSMSL). (2023). Tide gauge data [Dataset]. Retrieved from https://psmsl.org/

Pham, N. T., Staneva, J., Bonaduce, A., Stanev, E. V., & Grayek, S. (2024). Interannual sea level variability in the North and Baltic seas and net flux through the Danish straits. *Ocean Dynamics*, *74*(8), 669–684. https://doi.org/10.1007/s10236-024-01626-7

Piecuch, C. G., Bittermann, K., Kemp, A. C., Ponte, R. M., Little, C. M., Engelhart, S. E., & Lentz, S. J. (2018). River-discharge effects on United States Atlantic and Gulf coast sea-level changes. *Proceedings of the National Academy of Sciences*, *115*(30), 7729–7734. https://doi.org/10.1073/pnas.1805428115

Plag, H.-P., & Tsimplis, M. (1999). Temporal variability of the seasonal sea-level cycle in the North Sea and Baltic Sea in relation to climate variability. *Global and Planetary Change*, *20*(2), 173–203. https://doi.org/10.1016/S0921-8181(98)00069-1

Poropat, L. (2026). Code for "drivers of interannual to decadal sea level variability in northern Europe—Data driven approach [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.19070975

Poropat, L. (2026). Data associated with "drivers of interannual to decadal sea level variability in northern Europe—Data driven approach [Dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.19204512

Poropat, L., Jones, D., Thomas, S. D. A., & Heuzé, C. (2024). Unsupervised classification of the northwestern European seas based on satellite altimetry data. *Ocean Science*, *20*(1), 201–215. https://doi.org/10.5194/os-20-201-2024

Richter, K., Nilsen, J. E., & Drange, H. (2012). Contributions to sea level variability along the Norwegian coast for 1960–2010. *Journal of Geophysical Research*, *117*(C5). https://doi.org/10.1029/2011JC007826

Roberts, C. D., Calvert, D., Dunstone, N., Hermanson, L., Palmer, M. D., & Smith, D. (2016). On the drivers and predictability of seasonal-to-interannual variations in regional sea level. *Journal of Climate*, *29*(21), 7565–7585. https://doi.org/10.1175/JCLI-D-15-0886.1

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536. https://doi.org/10.1038/323533a0

Rus, M., Fettich, A., Kristan, M., & Ličer, M. (2023). HIDRA2: Deep-learning ensemble sea level and storm tide forecasting in the presence of seiches – The case of the northern Adriatic. *Geoscientific Model Development*, *16*(1), 271–288. https://doi.org/10.5194/gmd-16-271-2023

Samuelsson, M., & Stigebrandt, A. (1996). Main characteristics of the long-term sea level variability in the Baltic Sea. *Tellus A: Dynamic Meteorology and Oceanography*, *48*(5), 672–683. https://doi.org/10.3402/tellusa.v48i5.12165

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, *61*, 85–117. https://doi.org/10.1016/j.neunet.2014.09.003

Shen, C. (2018). A transdisciplinary review of deep learning research and its relevance for water resources scientists. *Water Resources Research*, *54*(11), 8558–8593. https://doi.org/10.1029/2018WR022643

Sithara, S., Pramada, S., & Thampi, S. (2020). Sea level prediction using climatic variables: A comparative study of SVM and hybrid wavelet SVM approaches. *Acta Geophysica*, *68*(6), 1779–1790. https://doi.org/10.1007/s11600-020-00484-3

Slangen, A. B. A., Carson, M., Katsman, C. A., van de Wal, R. S. W., Köhl, A., Vermeersen, L. L. A., & Stammer, D. (2014). Projecting twenty-first century regional sea-level changes. *Climatic Change*, *124*(1–2), 317–332. https://doi.org/10.1007/s10584-014-1080-9

Sturges, W., & Douglas, B. C. (2011). Wind effects on estimates of sea level rise. *Journal of Geophysical Research*, *116*(C6), C06008. https://doi.org/10.1029/2010JC006492

Tiggeloven, T., Couasnon, A., van Straaten, C., Muis, S., & Ward, P. J. (2021). Exploring deep learning capabilities for surge predictions in coastal areas. *Scientific Reports*, *11*(17224), 17224. https://doi.org/10.1038/s41598-021-96674-0

Tinker, J., Palmer, M. D., Copsey, D., Howard, T., Lowe, J. A., & Hermans, T. H. J. (2020). Dynamical downscaling of unforced interannual sea-level variability in the North–West European shelf seas. *Climate Dynamics*, *55*(7–8), 2207–2236. https://doi.org/10.1007/s00382-020-05378-0

Tur, R., Tas, E., Haghighi, A. T., & Mehr, A. D. (2021). Sea level prediction using machine learning. *Water*, *13*(24), 3566. https://doi.org/10.3390/w13243566

Van Houdt, G., Mosquera, C., & Nápoles, G. (2020). A review on the long short-term memory model. *Artificial Intelligence Review*, *53*(8), 5929–5955. https://doi.org/10.1007/s10462-020-09838-1

Vestøl, O., Ågren, J., Steffen, H., Kierulf, H., & Tarasov, L. (2019). NKG2016LU: A new land uplift model for Fennoscandia and the Baltic region. *Journal of Geodesy*, *93*(9), 1759–1779. https://doi.org/10.1007/s00190-019-01280-8

Wahl, T., Haigh, I., Woodworth, P., Albrecht, F., Dillingh, D., Jensen, J., et al. (2013). Observed mean sea level changes around the North Sea coastline from 1800 to present. *Earth-Science Reviews*, *124*, 51–67. https://doi.org/10.1016/j.earscirev.2013.05.003

Wakelin, S. L., Woodworth, P. L., Flather, R. A., & Williams, J. A. (2003). Sea-level dependence on the NAO over the NW European continental shelf. *Geophysical Research Letters*, *30*(7). https://doi.org/10.1029/2003GL017041

Weisse, R., Dailidienė, I., Hünicke, B., Kahma, K., Madsen, K., Omstedt, A., et al. (2021). Sea level dynamics and coastal erosion in the Baltic Sea region. *Earth System Dynamics*, *12*(3), 871–898. https://doi.org/10.5194/esd-12-871-2021

Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (3rd ed.), (Vol. 100). Elsevier. Retrieved from https://www.sciencedirect.com/bookseries/international-geophysics/vol/100

Wise, A., Calafat, F. M., Hughes, C. W., Jevrejeva, S., Katsman, C. A., Oelsmann, J., et al. (2024). Using shelf-edge transport composition and sensitivity experiments to understand processes driving sea level on the Northwest European shelf. *Journal of Geophysical Research: Oceans*, *129*(5), e2023JC020587. https://doi.org/10.1029/2023JC020587

Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. *Neural Computation*, *31*(7), 1235–1270. https://doi.org/10.1162/neco_a_01199