Abstract  The analysis of past regional climate-related sea level variations has important implications for diagnosing changes in future sea level driven by climate fluctuations. As the climate changes, there is a need for new explanatory variables of within-region climate factors and for more complex methods able to identify nonlinear relationships, such as machine learning algorithms. This study demonstrates the application of a new machine learning-based methodology to reconstruct historical sea level tide gauge records from proxy data (i.e., upper-ocean temperature estimates in open ocean regions), which provide a reasonably good dynamical representation of coastal sea level variations linked to slow and persistent natural processes like internal climate variability. The learning performance of our method was evaluated against observations of multiple stations and across a variety of model reconstructions, as shown and evidenced by the results.

Plain Language Summary  Coastal sea level changes in many regions of the world's oceans are still dominated by offshore (natural) processes such as internal climate variability. These local climate factors are reflected in the temperature changes within the upper-ocean layers of each region, which can be used to reconstruct regional sea level variability from incomplete tide gauge records. However, modeling the complex nonlinear dynamics of these variables requires advanced statistical techniques to produce such responses. We introduce a machine learning framework that allows including the regional ocean's dynamic properties and accounts for such complexity. Our ML-based reconstructions bring the opportunity to improve analysis of past climate-driven sea level variability, while exploring multiple variables for automatically modeling and reconstructing coastal sea level changes.

1. Introduction

Historical sea level records offer the opportunity to have a wider perspective on past climate-driven sea level variations and change on multi-year to decadal timescales. This knowledge is essential to understand current changes and improve future sea level predictions in response to climatic shifts. In fact, a recent study suggests that internal climate variability is still substantially impacting sea level in many coastal regions around the globe (Nieves et al., 2021). However, unfortunately, sea level records from tide gauges around the world's coastlines frequently include gaps (with over 50% missing data points at many stations) and biases (due to for example vertical land movements) (Holgate et al., 2013), which limits the study of past local and regional sea level variations. Another limiting factor is that some regions (e.g., coastal regions of Africa and South America) are poorly sampled. At some specific locations, tide gauge data may also represent a variety of regional processes other than climate variations (such as tides or storm surges, among other local factors) (Holgate et al., 2013). Thus, gap filling options for segmented, incomplete or uncertain tide gauge records continue to remain an important challenge (Wenzel & Schröter, 2010), especially for the study of “internally-induced” regional sea level changes (i.e., changes due to internal climate variations) over the past decades (Sérazin et al., 2016).

Of all the historical reconstruction methods that have been developed over the years (Carson et al., 2017), machine learning (ML) is a sound strategy for capturing the complex nonlinear behavior in oceanic (but not only) dynamical systems (Brunneau et al., 2020; Martinez et al., 2020; Vieira et al., 2020; Wenzel & Schröter, 2010, 2014). The only minor drawback of these techniques is that they require a training data set to construct the models that will be applied to “forecast” new data (in this case, the values at all tide gauge positions where data is missing). Current strategies for training the ML models include fitting the tide gauge data to climatological means, to a statistical analog (e.g., Empirical Orthogonal Function-based estimates) or to atmospheric reanalyses (Brunneau et al., 2020; Wenzel & Schröter, 2010, 2014). However, these estimates may not always fully represent the underlying nonlinear physics to ocean dynamics (Carson et al., 2017).
As a practical alternative for reconstructing the missing values, in this paper, we propose the use of region-specific sea level proxy data from upper-ocean temperature estimates. These estimates have previously proven to be successful in revealing observed coastal sea level changes associated with natural climate fluctuations (Nieves et al., 2021). Our further goal is to provide a novel modeling framework, developed with ML methods and physical knowledge (tide gauge observations and proxy data) to accommodate the different situations (described above and) generally encountered in the reconstruction of past sea levels on regional scales. State-of-the-art methods of ML include the recurrent neural network (RNN), which contain layers and interconnected nodes (or neurons) so that each data point can be influenced by every other data point to adjust the relationships that it sees according to a specific architecture. In particular, RNN was implemented in the present study using gated recurrent unit (GRU) for both univariate and multivariate scenarios (times series with a single/more than one time-dependent variable, respectively). GRU-RNN is a gating mechanism in RNN with two vectors (the so-called, update and reset gate) acting as a threshold for helping the network to learn the temporal dependencies of sequential data that carry all relevant information on the past of the process (Ha et al., 2021). Analysis with Gaussian Processes (GPs), a non-parametric regression ML method frequently used in modeling spatial and time series data, was also included for additional experiments (Camps-Valls et al., 2016).

The proposed methodology was tested in the coastal regions surrounding the Pacific, Indian and Atlantic basins against tide gauge and satellite observations. As an illustrative example, three regions (one for each ocean basin) are only shown here. Our models are, however, applicable to any coastal region on the globe most impacted by internal climate variability (Nieves et al., 2021). For each region, we chose one station with significant gaps (a station with at least a missing observation segment of several consecutive years) to reconstruct the incomplete record.

2. Materials and Methods

2.1. Sea Level Data and Proxies

The tide gauge monthly mean sea level records (relative to the Revised Local Reference datum) from the PSMSL (Permanent Service for Mean Sea Level) were used to reconstruct their missing values (https://www.psmsl.org/data/). Since the focus is on relatively open-ocean sites, the locations in narrow bays or river estuaries were avoided here. The analysis was also restricted to the same period (the overlapping years of the majority of observations) in each region to have at least a few independent observations for comparison with the regional models. In addition, we used the (delayed-time “all sat merged”) daily mean sea level anomalies product from satellite altimetry (https://resources.marine.copernicus.eu/?option=com_csw&view=details&product_id=SEALEVEL_GLO_PHY_L4_REP_OBSERVATIONS_008_047) to examine the regional confidence of our results post-1993. All time-series were resampled at seasonal resolution for consistency with the resolution of the proxy variables.

Aside from sea level measurements, we used sea level proxy-data from temperature at-depth (from Nieves et al., 2021 historical datasets). The temperature observations of the upper layers of open ocean regions contain climate information associated with internal changes (due to heat exchange between the ocean and the atmosphere) (Liu et al., 2016; Nieves et al., 2015; Roemmich et al., 2015), which modulates sea level variability along many coastlines of the world (Nieves et al., 2021). In fact, they both (the upper-ocean temperature changes in the open ocean and sea level variability in the respective coastal regions) exhibit similar patterns on time scales of a few years, as it was shown via ML (Nieves et al., 2021). Note that the temperature estimates (with a temporal resolution of 3 months) are derived from Levitus et al. (2012) data at different depth layers and defined regionally over large open ocean areas to properly reflect the signature of internal climate variability, as explained in Nieves et al. (2017). Records are also detrended by removing the long-term trend to highlight variations on multi-year timescales. Then, a 1-year running mean filter (Nieves et al., 2017, 2021) was applied to the time-series to smooth out inter-annual variability.

It should be noted as well that temperature-based estimates have many advantages over other proxy data based on (for example) the shorter satellite altimetry sea surface height record such as the memory length of the climate system that goes back to 1955. Moreover, a temperature estimate is also less sensitive to local effects (such as atmospheric forcing) that can affect coastal sea levels, particularly, in semi-enclosed basins and along shallow areas (Dangendorf et al., 2014, 2015).
2.2. GRU Artificial Recurrent Neural Networks

GRU is an artificial RNN architecture widely used in machine learning to learn complex patterns and train the model by using back-propagation through time (Ha et al., 2021). This is, the GRU decides how much of the past information to forget or to keep, and what new information to add to update the neural network's weights (the factors applied to each individual input of a neuron) by using the reset and update gate. Similar to the long short-term memory (LSTM), GRUs adjust a fully-connected nonlinear hierarchical structure made of neurons by minimizing the training loss function in the output layer. Here, the loss function is the mean-square error function (or MSE) and it characterizes the quality of the network resulting from training. Compared to that of LSTM, GRU-based networks have been shown to demonstrate better performance on smaller to medium quantity datasets (Ayzel & Heistermann, 2021), a typical situation in the context of data reconstruction. Hence, GRU may be the right solution to LSTM, as it needs less data to generalize and may train a bit faster. This being said, in our study, GRU vs. LSTM is unlikely to make a big difference, as we consider the long-term dependencies through the long record of our proxy data.

2.3. Development of a Novel Architecture

In this study, we first built a novel GRU-RNN architecture uniquely designed for each tide gauge location. The training procedure is as follows.

The tide gauge record to be reconstructed at a given location was split into two parts: the portion of data with all the missing points (data gaps) and the remaining (available) data points. The latter were used as input for training together with either the related regional proxy or the (reconstructed) time series of the neighboring stations to build the local model (as shown later). For this process, the training set was in turn split into train/validation sets. A ratio of 90:10 was considered as the best overall ratio for training and validating the models. Thus, 10% of the data was set aside for validation purposes.

The following step consists in tuning the relevant hyperparameters (the parameters that control the learning process specified with the MATLAB trainingOptions function), which are the number of hidden neurons or units (HUs) and the epochs. A HU is the function that applies weights to the inputs according to the knowledge “learned”, and its number was established by experiment runs using the adam solver (an iterative optimization algorithm). The results were not highly sensitive to the particular type of optimization method. The number of epochs (i.e., the number of times that the learning algorithm will run through the dataset) was automatically determined by the following MSE stopping criterion. The MSE-based feature scheme was set to stop training when the validation MSE loss is equal to (or larger than) the one at two previous epochs for several consecutive passes through the training set.

The net architecture was designed with the following 5 layers (see Figure 1): a sequence input layer, the reset/update gate layers (with 30–150 neurons to regulate the flow of information through the sequence chain), a dropout layer of 30% (that prevents overfitting by randomly dropping neurons from the neural network during training in each iteration), a fully connected layer (that connects every neuron in one layer to every neuron in another layer), and the output regression layer (Sowmya & Supriya, 2021).

These training options and multi-layers architecture were implemented (in the MATLAB trainNetwork function) with the training dataset to forecast the unseen data. This is, the predicted outcome for new data is obtained using (the MATLAB predict function with) the net configuration learned from training data through the MSE-based approach on the test dataset (the part of the regional proxy or near neighbor tide gauge estimates corresponding to the missing data). Thus, the total size of the test dataset is not prescribed by any standard split of data, but equivalent to the number of points to be reconstructed. The percentage of missing data points depends on the station, ranging from 13% to 34% of data for the cases included here.
3. Results and Discussion

3.1. ML for Past Sea Level Reconstruction Using Proxy-Climate Fingerprints

The simplest type of reconstruction involves using the existing tide gauge data along with the proxy data (as training dataset) to either reconstruct the entire record (see Figure 2) or to simply fill-in the gaps. Either way yields comparable results (not shown).

Figure 1. The proposed GRU-RNN architecture. Our model architecture contains one input and output layer and three hidden layers (the Gated Recurrent Unit or GRU, the Dropout or DO and the Fully-Connected or FC layer) made up of a set of neurons (see previous section). Each neuron (shown as a circle) is connected through weighted connections (lines). The weights are updated through error backpropagation (indicated by long backward red arrows).

Figure 2. Our proxy-based sea level reconstructions using a GRU-RNN approach compare well with the regional tide gauge (TG) observations. Time series (a) modeled (in blue) and observed (for reconstruction in black and for comparison in gray) for the North-East Pacific Ocean. (b and c) same as (a) except for the East Indian Ocean and North-East Atlantic Ocean. The TG stations considered for reconstruction and the nearby TG stations are shown on the left panel (in yellow and red, respectively). Sea level expressed in mm.
The best proxy estimate was chosen according to Nieves et al. (2021) (see Section 2). In their study, several temperature changes in the upper layers of open ocean regions—a proxy for climate variability conditions—were analyzed (for each depth layer, time-lag and across a range of timescales) and ranked based on their performance modeling coastal sea level variability using a GP approach. They concluded that, in many regions like the North-East Pacific Ocean (NE PO) and North-East Atlantic Ocean (NE AO), the vertically averaged temperature over the top 100 m is a good proxy indicator of coastal sea level variability. In contrast, in places like the East Indian Ocean (E IO), we need to reach deeper layers and consider the vertically integrated ocean heat content down to 700 m depth to capture the largest changes in coastal sea levels (Nieves et al., 2021). These three regions (NE PO, NE AO and E IO) were selected as case studies here.

Next, we applied the GRU-RNN-based approach to reconstruct the incomplete time series chosen for each region (highlighted with a yellow star symbol on the left panel in Figure 2). As indicated above, the GRU-RNN algorithm first learns the relevant information on past sea level variability with the training set using an error-back-propagation training (see Section 2). Once the model is trained, we can make predictions on the unseen data. We found that the optimum performance was obtained using at least 30 HUs for the NE PO and 150 HUs for both the E IO and NE AO regions.

Figure 2 shows the regional proxy-based models (in blue, right panel) built with this configuration. Since unfortunately there are no known independent measurements of the missing points, a certain degree of confidence can be obtained by comparison against data from nearby stations (gray symbol lines). Superimposed is also plotted the record being reconstructed (black solid line). We can see that the reconstruction is able to reproduce the variability (and even the magnitude at some locations) exhibited by the data. The model is remarkably good at identifying broad patterns and also excluding data points that are considered outliers (see, for example, 2005 in El Hierro, NE AO). The slightly different performance qualities are consistent with the different regional dynamics (Nieves et al., 2021).

For additional quality assessment, we compared the results from an alternative ML approach (further details are available in Supporting Information S1). We performed an experiment with GPs (Camps-Valls et al., 2016) applied to the same data, which provided equally valid models (Figures S1–S3 in Supporting Information S1). Yet, RNNs models exhibit slightly different patterns due to the complex architecture of the method. RNN-based models, by design, can help unveil and mimic complex behavior more efficiently (see Figure S4 in Supporting Information S1). Each data point can be influenced by every other data point in a way that adapts relationships that it sees. Also note that, although the GP construction is simpler and hence faster, the computational cost difference (between RNN and GP reconstructions) is insignificant here.

### 3.2. Regional Consistency of the ML Model Reconstructions

The sea level pattern consistency between stations can also be easily identified in each region (Figure 2) and reinforces the idea that there is a common large-scale regional factor (such as local climate) modulating historical changes in all basins. It also suggests the usefulness of using the information of the neighboring stations with long sea level records to predict regional variability at other nearby stations. Hence, in practice, rather than reconstructing a single-site tide gauge record (light blue lines in Figure 3), we can produce reconstructions of the surrounding stations within each region and compute its average (thick dark blue line, same figure), which provides the most representative approximation of the regional variations in sea level. The standard deviation resulting from the spread of reconstructions can be used as an error estimate of the interval of uncertainty. Results suggest general confidence in the model reliability, and higher similarity between reconstructions and tide gauge/satellite observed values (also in black) in the Indo-Pacific area (the region most influenced by natural internal variability and with many operational stations since the early 1900s) (see Figures 3a and 3b).

Considering all of the above, our ML-approach offers three reconstruction options designed from the proxy database (as schematized in Figure 4a). The first (aforementioned) option is to model the gaps in the sea level record of a given station using regional proxy-data, a good strategy when there is a limited number of (reliable) stations around the region. Otherwise, we can use the proxy-reconstruction of nearby stations as model inputs (the second option), which should give us much better-quality models than those with a single station. The same is concluded for the mean of all reconstructions (the last option). Of all options, feeding a network with information from other regional stations will always result in a more robust reconstruction, as it will enable the model to learn a range of
possible local patterns within a region (as opposed to using a particular one or the average of the available ones). Nevertheless, all methods, shown (in different shades of blue) in Figures 4b–4d, perform similarly to each other and to the data (in black).

4. Summary and Conclusions

To summarize, compared to current statistical methods of reconstructing past changes in sea levels, the proposed ML-based approach embeds the inherent complexity of machine learning methods while being physically consistent with the region-specific underlying ocean-climate dynamics, by applying this methodology with regional proxy data. These proxies, which are derived from temperature estimates within open ocean regions away from the coasts, have shown its usefulness in a recent study to model and predict regional coastal sea levels induced by offshore processes, including internal climate variability. Therefore, we exploited these regional proxies here to learn the weights in the machine learning reconstruction models. At the same time, we have taken advantage of the broad regional similarity of the sea level patterns seen in tide gauge records for a number of different types of reconstruction. Whether we use data from one location only or from nearby locations, our machine learning algorithms show comparable ability to successfully reconstruct past sea level variability. Looking forward, the potential of our approach introduces new possibilities of machine learning modeling in the context of local to regional sea levels to effectively improve representation of past and future sea level changes at specific locations.
Figure 4. Three ML-based reconstruction frameworks validated against each other’s models and TG data. (a) Overview of our multiple approach for sea level reconstructions applied to a particular station (shown in Figure 2, black line) in the North-East Pacific Ocean (b), East Indian Ocean (c) and North-East Atlantic Ocean (d). This approach includes analysis of regional proxy data (P, in cyan), proxy-reconstructions of the nearby stations (R(P)i, in steel blue), and the mean of the regional reconstructions (〈R(P)i〉, in blue) to reconstruct historical data of a single site (R*). The station under study and the rest of stations are identified by an asterisk and subindex i, respectively. Sea level expressed in mm.
Data Availability Statement

Data supporting the findings of this study and code is available at https://doi.org/10.5281/zenodo.5721614 (Radin & Nieves, 2021).

References

Ayzel, G., & Heistermann, M. (2021). The effect of calibration data length on the performance of a conceptual hydrological model versus LSTM and GRU: A case study for six basins from the CAMELS dataset. *Computers and Geosciences*, 149, 104708. https://doi.org/10.1016/j.cageo.2021.104708

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

Campos-Valls, G., Verrelst, J., Munoz-Mari, J., Laparra, V., Mateo-Jimenez, F., & Gomez-Dans, J. (2016). A survey on Gaussian processes for earth-observation data analysis: A comprehensive investigation. *IEEE Geoscience and Remote Sensing Magazine*, 4(2), 58–78. https://doi.org/10.1109/MGRS.2015.2510084

Carson, M., Köhl, A., Stammer, D., Meyssignac, B., Church, J., Schröter, J., et al. (2017). Regional sea level variability and trends, 1960–2007: A comparison of sea level reconstructions and ocean syntheses. *Journal of Geophysical Research: Oceans*, 122(11), 9068–9091. https://doi.org/10.1002/2017JC013299

Dangendorf, S., Marcos, M., Müller, A., Zorita, E., Riva, R., Berk, K., et al. (2015). Detecting anthropogenic footprints in sea level rise. *Nature Communications*, 6, 7849. https://doi.org/10.1038/ncomms8849

Dangendorf, S., Rybski, D., Mudersbach, C., Müller, A., Kaufmann, E., Zorita, E., et al. (2014). Evidence for long-term memory in sea level. *Geophysical Research Letters*, 41(15), 5530–5537. https://doi.org/10.1002/2014GL060538

Ha, S., Liu, D., & Mu, L. (2021). Prediction of Yangtze River streamflow based on deep learning neural network with El Niño–Southern Oscillation. *Scientific Reports*, 11(1), 1–23. https://doi.org/10.1038/s41598-021-90964-3

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

Levitus, S., Antonov, J. I., Boyer, T. P., Locarnini, R. A., Garcia, H. E., and Reagan, G. M. (2012). World ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010. *Geophysical Research Letters*, 39, L10603. https://doi.org/10.1002/2012GL051106

Liu, W., Xie, S. P., & Lu, J. (2016). Tracking ocean heat uptake during the surface warming hiatus. *Nature Communications*, 7, 1–9. https://doi.org/10.1038/ncomms10926

Martinez, E., Brini, A., Gorgues, T., Drumetz, L., Roussillon, J., Tandeo, P., et al. (2020). Neural network approaches to reconstruct phytoplankton time-series in the global ocean. *Remote Sensing*, 12(24), 1–13. https://doi.org/10.3390/rs12244156

Nieves, V., Marcos, M., & Willis, J. K. (2017). Upper-ocean contribution to short-term regional coastal sea level variability along the United States. *Journal of Climate*, 30(11), 4037–4045. https://doi.org/10.1175/JCLI-D-16-0896.1

Nieves, V., Radin, C., & Camps-Valls, G. (2021). Predicting regional coastal sea level changes with machine learning. *Scientific Reports*, 11(1), 1–6. https://doi.org/10.1038/s41598-021-87460-z

Nieves, V., Willis, J. K., & Patzert, W. C. (2015). Recent hiatus caused by decadal shift in Indo-Pacific heating. *Science*, 349, 532–535. https://doi.org/10.1126/science.aaa4521

Radin, C., & Nieves, V. (2021). Machine learning sea level reconstructions package. Version 1.0. Zenodo. https://doi.org/10.5281/zenodo.5721614

Roemmich, D., Fox, G., Gripaios, G., Kottmeier, C., Raymond, B., & Smith, R. (2005). Unabated planetary warming and its ocean structure since the 20th century. *Journal of Geophysical Research: Oceans*, 110(C8), 1–15. https://doi.org/10.1029/2004JC002513

Radin, C. (2013). Convolutional Neural Network (CNN) Fundamental Operational Survey. In M. N. Favorskaya, S. L. Peng, M. Simic, B. Alhaddad, & S. Pal (Eds.), *Intelligent computing paradigm and cutting-edge technologies. Learning and analytics in intelligent systems series* (Vol. 21, pp. 245–258). Springer. https://doi.org/10.1007/978-3-319-65407-8_21

Vieira, F., Cavalcante, G., Campos, E., & Taveira-Pinto, F. (2020). A methodology for data gap filling in wave records using Artificial Neural Networks. *Applied Ocean Research*, 98, 102109. https://doi.org/10.1016/j.apor.2020.102109

Wenzel, M., & Schröter, J. (2010). Reconstruction of regional mean sea level anomalies from tide gauges using neural networks. *Journal of Geophysical Research: Oceans*, 115(8), 1–15. https://doi.org/10.1029/2009JC005630

Wenzel, M., & Schröter, J. (2014). Global and regional sea level change during the 20th century. *Journal of Geophysical Research: Oceans*, 119, 7493–7508. https://doi.org/10.1002/2014JC009990