Applied Computing and Geosciences 23 (2024) 100173

Contents lists available at ScienceDirect

Applied Computing and Geosciences

journal homepage: www.sciencedirect.com/journal/applied-computing-and-geosciences

# Enhancing estuary salinity prediction: A Machine Learning and Deep Learning based approach

Leonardo Saccotelli [a], Giorgia Verri [a], Alessandro De Lorenzis [a], Carla Cherubini [c,a,b], Rocco Caccioppoli [a], Giovanni Coppini [a], Rosalia Maglietta [b,a,*]

[a] CMCC Foundation - Euro-Mediterranean Center on Climate Change, Italy, Lecce
[b] Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing, CNR-STIIMA, Italy, Bari
[c] Department of Computer Science, University of Bari, Italy, Bari

ARTICLE INFO

Dataset link: https://github.com/CMCC-Foundation/CMCC-Hybrid-EstuaryBoxModel/tree/ML-EBM-PoGoro

*Keywords:*
Machine Learning
Least-Squares Boosting
Long Short-Term Memory
Estuary salinity estimation
Salt wedge intrusion

ABSTRACT

As critical transitional ecosystems, estuaries are facing the increasingly urgent threat of salt wedge intrusion, which impacts their ecological balance as well as human-dependent activities. Accurately predicting estuary salinity is essential for water resource management, ecosystem preservation, and for ensuring sustainable development along coastlines. In this study, we investigated the application of different machine learning and deep learning models to predict salinity levels within estuarine environments. Leveraging different techniques, including Random Forest, Least-Squares Boosting, Artificial Neural Network and Long Short-Term Memory networks, the aim was to enhance the predictive accuracy in order to better understand the complex interplay of factors influencing estuarine salinity dynamics. The Po River estuary (Po di Goro), which is one of the main hotspots of salt wedge intrusion, was selected as the study area. Comparative analyses of machine learning models with the state-of-the-art physics-based Estuary box model (EBM) and Hybrid-EBM models were conducted to assess model performances. The results highlighted an improvement in the machine learning performance, with a reduction in the RMSE (from 4.22 psu obtained by physics-based EBM to 2.80 psu obtained by LSBoost-Season) and an increase in the $R^2$ score (from 0.67 obtained by physics-based EBM to 0.85 by LSBoost-Season), computed on the test set. We also explored the impact of different variables and their contributions to the predictive capabilities of the models. Overall, this study demonstrates the feasibility and effectiveness of ML-based approaches for estimating salinity levels due to salt wedge intrusion within estuaries. The insights obtained from this study could significantly support smart management strategies, not only in the Po River estuary, but also in other location.

## 1. Introduction

Estuaries are transitional systems that modulate freshwater inputs into the sea, with ocean saltwater entering the river mouth and merging with the zero-salinity river streamflow. Estuarine systems are highly productive environments that provide a multitude of crucial ecosystem services including the availability of raw materials, cultural services, coastal erosion protection and the maintenance of fish stocks (Keyes et al., 2021; Barbier et al., 2011; Townsend et al., 2013; Boerema and Meire, 2017). Among the environmental threats that challenge the functionality and biodiversity of estuaries, saltwater intrusion is one of the most serious (Herbert et al., 2015). It increases the salt content in waters and expands the salt wedge intrusion throughout the river channels (Wong et al., 2014). Monitoring the environmental health status of estuaries is essential for water resource management and ecosystem preservation, and for ensuring sustainability (Ghalambor et al., 2021). Since estuaries are poorly monitored, there is a rising demand for the development of modeling tools that can be easily applied across diverse estuaries for accurately estimating the extent and patterns of salt wedge intrusion. To address this issue, the CMCC Estuary Box Model (CMCC EBM model) (Verri et al., 2020) has been specifically developed and calibrated for the Po di Goro estuary (Verri et al., 2021; Kurdistani et al., 2022). This physics-based model comprehensively represents all the fundamental physical processes governing estuarine dynamics and at the same time it is orography agnostic, thus reducing the source of uncertainty which may arise when this information is

* Corresponding author at: Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing, CNR-STIIMA, Italy, Bari.
*E-mail addresses:* leonardo.saccotelli@cmcc.it (L. Saccotelli), giorgia.verri@cmcc.it (G. Verri), alessandro.delorenzis@cmcc.it (A. De Lorenzis), c.cherubini@phd.poliba.it (C. Cherubini), rocco.caccioppoli@cmcc.it (R. Caccioppoli), giovanni.coppini@cmcc.it (G. Coppini), rosalia.maglietta@cnr.it (R. Maglietta).

https://doi.org/10.1016/j.acags.2024.100173
Received 25 March 2024; Received in revised form 7 June 2024; Accepted 10 June 2024
Available online 18 June 2024

unavailable or not well represented. Overall, it stands out as a flexible and powerful tool linking hydrology and ocean modeling for both forecast and climate purposes by estimating the net river release at river mouths and salt-wedge intrusion. The very low computational cost and the limited data storage required make its application competitive with the advanced 3D finite element modeling approach. Indeed, the EBM has recently been exploited to perform centennial climate projections of the salt wedge-intrusion along the Po di Goro estuary and to judge a site-specific nature-based solution to counteract the salinization issue by means of long-term sensitivity experiments (Verri et al., 2024a). Despite the interesting characteristics of EBM, there is a clear need to improve the performance of salinity estimation, which has ample room for improvement. With this objective in mind, it is evident that Machine Learning (ML) offers enormous opportunities, even by resorting to traditional models that maintain a reduced computational cost compared to Deep Learning (DL) strategies.

A recent advancement in the fully physics-based EBM model was presented in Maglietta et al. (2023c), named Hybrid-EBM. This innovative model integrates ML techniques into the traditional physics-based model by replacing its parameterization equations that include several tunable coefficients, resulting in a significant enhancement of the overall model performance. However, in recent years, several very effective methodologies in the field of ML and DL have been developed and implemented across a range of application domains (Maglietta et al., 2023b,a; Dimauro et al., 2023b,a; Maglietta et al., 2022). With the continuous advancements in ML technology, novel solutions can be used to address the issue of salinization at river mouths. Several studies have estimated the salinity of estuaries using ML algorithms. For example, in Saccotelli et al. (2023) the authors presented different SVM models to predict the estuary salinity at the Po Goro river. In Lu et al. (2021) they used the Bayesian model averaging method to create an integrated forecast model to predict the monthly saltwater intrusion at the Pearl River Delta. In Liu et al. (2015) the authors presented a Random Forest model to estimate sea surface salinity in the Hong Kong Sea, China, by integrating in situ and remotely sensed data. Lastly, similar applications can also be found in Fang et al. (2017), Tran et al. (2022), Qiu and Wan (2013), Guillou et al. (2023), Qi et al. (2022a), Nguyen et al. (2021), Rath et al. (2017), Hoai et al. (2022), Tran et al. (2021), Ye et al. (2020), Mohamad et al. (2018), Lal and Datta (2018), Qi et al. (2022b). In this study, the salinity conservation equation of the physically-based EBM was replaced by ML and DL algorithms, namely Random Forest, Least-Squares Boosting, Artificial Neural Network and Long Short-Term Memory network algorithms. The ML algorithms thus benefit from the principles of the salinity conservation equation to identify the key variables conditioning the target field, but without needing to explicitly use the conservation equation itself. ML techniques can also produce more effective and realistic estimates. This is an important finding because the estuarine salinity is a non-linear combination of multiple coastal forcings.

DL strategies were also adopted to highlight temporal dependencies in a forecasting framework. Long Short-Term Memory (LSTM) network was selected to model sequential data, since it is well suited for time-series data because of its ability to learn long-term dependencies. LSTM can predict estuary salinity levels capturing temporal dependencies, such as daily fluctuations, tidal patterns, and other variations. The goal was to forecast estuarine salinity levels up to seven days beyond the input data.

An initial statistical analysis was conducted to gain insights into yearly and seasonal salinity patterns. Two sets of experiments were then carried out: firstly, Random Forest, Least-Squares Boosting, Artificial Neural Network were trained to predict estuary salinity levels, where the temporal dimension was not explicitly considered. A separate set of experiments was then undertaken, focusing exclusively on Long Short-Term Memory models. This strategy used a time-series methodology to develop a weekly prediction model. Our study also explores the impact of different variables and their contributions to the predictive capabilities of the models. Starting from data collected between 2016 and 2019 at the Po di Goro river mouth, one of the main global hotspots of salt wedge intrusion (Tarolli et al., 2023), ML and DL models were trained and validated, and compared with the performances of the physics-based EBM and Hybrid-EBM.

## 2. Material and methods

### 2.1. Study area

The Po River is the Mediterranean Sea's second-largest freshwater inflow, coursing through northern Italy before ending in the northern Adriatic Sea. It maintains an annual flow rate averaging approximately 1500 $m^3\ s^{-1}$ (Boldrin et al., 2005) and accounts for approximately 30% of the total freshwater input into the Adriatic (Cushman-Roisin et al., 2002). The Po River delta has nine separate branches; from south to north: Po di Goro, Po di Gnocca, Po di Tolle, Po di Bastimento, Po di Scirocco, Po di Bonifazi, Po di Dritta, Po di Tramontana, and Po di Maistra. The mouths of these branches are denoted by the yellow markers in Fig. 1. The Po di Goro, one of the branches within the Po River delta system, is responsible for transporting approximately 13% of the overall runoff from the Po River. It is a salt wedge estuary (Valle-Levinson, 2010). In fact, it functions as a river-dominated estuary, leading into a micro-tidal sea, with a well-defined salt wedge under conditions of moderate to low river runoff. In contrast, during exceptionally high runoff events, salinity levels decrease, and the salt wedge may recede.

### 2.2. Data collection

Estuarine salinity observations ($S_{ul}$) were collected for the study area of Po di Goro from the Agenzia Regionale per la Prevenzione, l'Ambiente e l'Energia dell'Emilia Romagna (Arpae). Arpae provided a comprehensive dataset of 1201 hourly salinity data collected at the Manufatto gauge station, located near the river mouth, over the timeframe selected for the experiments, i.e. 2016 to 2019. These observations therefore include high frequency sub-tidal processes, e.g. the semi-diurnal tidal cycle of about 12 h and the "blocking effects" which may occur within a few hours. For experimental purposes, hourly observations were averaged daily for comparison with the EBM results which were averaged over the diurnal tidal cycle. The salinity data are readily accessible online (Arpae, 2024b). Details of the dataset can be found in Table 1.

The study uses the following input features: $Q_{river}$, $Q_{ll}$, $S_{ll}$, $S_{ocean}$ and $Q_{tidef}$ (see Table 2). At the estuary's head, the input feature is the river volume flux ($Q_{river}$). This is derived from data collected by Arpae at the Pontelagoscuro gauge station, located upstream of the river delta, considering 13% of the data recorded from this station (Arpae, 2024a; Verri et al., 2021). The input features at the estuary mouth related to the ocean inflow are the volume flux ($Q_{ll}$) and salinity ($S_{ll}$) entering from the bottom and corresponding to the lower layer of EBM (Verri et al., 2020, 2021). In addition, the depth averaged ocean salinity ($S_{ocean}$) is also considered, which entails computing the average of the ocean salinity measured at different depths along the water column. $Q_{ll}$, $S_{ll}$ and $S_{ocean}$ were derived from a Copernicus Marine Product, the CMEMS Mediterranean Sea Physical Reanalysis (Escudier et al., 2020, 2021; Nigam et al., 2021). CMEMS data were organized on a grid with a horizontal resolution of 1/24° (ca. 4.5 km), with 141 unevenly spaced vertical levels up to a depth of about 6000 m and a daily frequency. Lastly, the tidal impact on the net volume flux at the river mouth ($Q_{tidef}$) was estimated using the OTPS (Ocean Tidal Prediction Software) barotropic model (Egbert and Erofeeva, 2002). This was necessary because the CMEMS product under consideration does not include tidal signals.


**Fig. 1.** The map of the Po delta system (Google Earth image) along with its the river mouths, namely, Po di Maistra, Po della Pila (which splits into Po di Tramontana, Po di Dritta, Po di Scirocco, Po di Bonifazi and Po di Bastimento), Po di Tolle, Po di Gnocca, and Po di Goro (yellow markers). The red path is the Po di Goro branch selected as the test case. The location of the Manufatto gauge station, which is close to the Goro mouth but in a secondary channel, is shown in the inset. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
*Source:* Image from Verri et al. (2021).

**Table 1**
Basic description and statistical information of the in situ observations of river mouth salinity ($S_{ul}$). Min, Max, Avg, and Std dev represent the minimum, maximum, average, standard deviation values, respectively.

| In situ observations | Num. of observations | Period | Frequency | Unit | Min | Max | Avg | Std dev |
|---|---|---|---|---|---|---|---|---|
| $S_{ul}$ | 1201 | 2016–2019 | hourly | psu | 0.1 | 29.5 | 13.89 | 7.17 |

**Table 2**
Sources of input features used for the river mouth salinity ($S_{ul}$) regression model.

| Forcing | Source | Input | Unit |
|---|---|---|---|
| Runoff forcing | Arpae (Arpae, 2024a) | $Q_{river}$ | $m^3\ s^{-1}$ |
| Ocean forcing | CMEMS Reanalyses (Escudier et al., 2020, 2021; Nigam et al., 2021) | $Q_{ll}$ | $m^3\ s^{-1}$ |
| | | $S_{ll}$, $S_{ocean}$ | psu |
| Tidal forcing | OTPS model (Egbert and Erofeeva, 2002) | $Q_{tidef}$ | $m^3\ s^{-1}$ |

### 2.3. Machine learning algorithms

#### 2.3.1. Random Forest (RF)

Random Forest (RF) (Breiman, 2001; Maglietta et al., 2023b, 2018) is an ensemble learning method which, when used for regression tasks, combines multiple decorrelated decision trees to make predictions for continuous numerical values. It works by growing multiple decision trees, each trained on a randomly selected subset of the data and features. Predictions from each tree are averaged and provide the final output for the model. To build this model the fitrensemble function was used (The MathWorks Inc., 2023a). The optimizable hyperparameters for RF are the number of ensemble learning cycles (NumLearningCycles[1]), the minimum leaf size (MinLeafSize[1]), the maximum number of splits (MaxNumSplits[1]) and number of predictors to sample (NumVariablesToSample[1]).

#### 2.3.2. Least-Squares Boosting (LSBoost)

The Least-Squares Boosting (LSBoost) algorithm (Friedman, 2001; Breiman, 2001; Hastie et al., 2009; Maglietta et al., 2023b) is a variant of the generic Gradient Boosting algorithm (Friedman, 2001), in which the least-squares (LS) strategy is applied. Through a weighted combination of the outputs produced by a set of weak learners, LSBoost defines a function that is able to estimate the target feature. More specifically, at each step, the algorithm fits a new learner to the difference between the observed response and the aggregated prediction of all learners grown previously. LSBoost minimizes the squared error loss through gradient descent: $\frac{(y_i - f(x_i))^2}{2}$. All the new learners are fitted to $y_i - \eta f(x_i)$, where $i = 1, \ldots, m$ is the $i$th observation, with $m$ equal to the total number of observations, $y_i$ is the observed response, $\eta f(x_i)$ is the aggregated prediction from all weak learners grown thus far for observation $x_i$ and $\eta$ is the learning rate. This model was built using the fitrensemble function (The MathWorks Inc., 2023a). The optimizable hyperparameters for LSBoost are the number of ensemble learning cycles (NumLearningCycles[1]), the minimum leaf size (MinLeafSize[1]), the maximum number of splits (MaxNumSplits[1]), the number of predictors to sample (NumVariablesToSample[1]) and the learning rate (LearnRate[1]).

#### 2.3.3. Artificial Neural Networks (NNs)

Artificial Neural Network (NNs) (Goodfellow et al., 2016) are a class of networks consisting of three basic layers (input, hidden and output) of computational units, usually interconnected in a feed-forward way.

[1] See MATLAB documentation for details.

This means that each neuron in one layer has direct connections only to the neurons of the subsequent layer; each neuron in the hidden and output layers, connected by weights and biases, uses a nonlinear activation function. Two main processes characterize a neural network: forward propagation of the input and back propagation of the errors. During the forward propagation phase, the inputs are propagated from the input layer, throughout all the hidden layers, to the output layer. Then during the back propagation phase, the errors computed on the outputs of the neural network are back propagated, layer by layer, to the input layer, so that each of the neural network's weights is updated. To build this model the fitrnet function was used (The MathWorks Inc., 2023b). In this function a limited-memory Broyden–Fletcher–Goldfarb–Shanno quasi-Newton algorithm (LBFGS) (Nocedal and Wright, 2006) is used as loss function minimization technique to minimize the mean squared error (MSE). The optimizable hyperparameters for NN are the number and size of each hidden layer (LayerSizes[1]), the activation function (Activations[1]), the regularization term strength (Lambda[1]), the initialization strategy for the layer biases (LayerBiasesInitializer[1]) and the initialization strategy for the layer weights (LayerWeightsInitializer[1]).

#### 2.3.4. Long Short-Term Memory Networks (LSTMs)

LSTMs (Hochreiter and Schmidhuber, 1997) are a special type of Recurrent Neural Network (RNN) designed to overcome the vanishing gradient problem and to learn long-term dependencies between time steps in time-series and sequence data. The core idea in LSTM is to introduce a cell state $C_t$, which contains information learned from the previous time steps. Using a structure called gate, the LSTM has the ability to remove or add information from the cell state. Firstly, the forget gate $f_t$ determines the information to be discarded from the cell state. Then, the input gate $i_t$, implemented through a sigmoid layer, identifies which values to update in the cell state. A tanh layer generates a vector of new candidate values that might be incorporated into the cell state. The cell state will thus be updated. The output gate $o_t$, which is also a sigmoid layer, determines the information to be emitted from the cell state. Finally, the hidden state $h_t$ is updated by processing the cell state through a tanh layer and multiplying it by the output gate. This updated hidden state carries information forward to subsequent steps and contributes to the final prediction.

### 2.4. Experimental setup

#### 2.4.1. Experimental setup for ML algorithms

RF, LSBoost and NN were implemented in MATLAB (The MathWorks Inc., 2023c) using the Statistics and Machine Learning Toolbox (The MathWorks Inc., 2023d). For the ML algorithms, automated hyperparameter optimization and the cross-validation technique were employed. The schema of the methodology adopted to train the ML models is shown in Fig. 2. Firstly, the input dataset is divided into a training set, comprising 80% of the input examples, and a test set, with the remaining 20% examples. The training set is used for the automatic optimization of hyperparameters, performed by the Bayesian optimization method (Bull, 2011; Gelbart et al., 2014; Snoek et al., 2012). This method employs a surrogate for the objective function, which is considerably easier to optimize than the objective function itself. The process involves identifying the next set of hyperparameters for evaluating the objective function, and it does so by selecting hyperparameters that result in superior performance for the surrogate function. The k-fold cross validation technique was also used, which splits training data into k subsets, or folds, of equal size (Refaeilzadeh et al., 2016; Rodriguez et al., 2010). During each iteration, k-1 folds serve as the training dataset, with the remaining fold used for validation. This process is repeated k times to ensure that each instance undergoes validation exactly once. The estimated errors from each iteration are then averaged across the k folds, resulting in a comprehensive measure of the model's performance. In this study, we used a 5-fold cross-validation strategy. We thus identified the hyperparameter settings that produced the best model performance. Finally, we applied the model, trained with these optimal hyperparameters, to the test dataset and presented the associated performance metrics.

#### 2.4.2. Experimental setup for LSTM

The LSTM network was employed to forecast salinity at river mouth since it is designed to model sequential data. By training the model using data from previous days, LSTM strategy developed here forecasts estuarine salinity for up to seven subsequent days. The forecasted values are not used as inputs for the subsequent forecasting, in order to simplify the model architecture and training process. In fact, models only rely on past observations to make predictions. This approach may be more stable and easier to train, especially for simpler forecasting tasks or when the relationship between past and future values is relatively straightforward. The sequence of daily estuarine salinity was converted into a sequence of input/output pairs, in order to train the network with a supervised learning procedure. A total of 20% of the available data was reserved for testing the trained network. Of the remaining 80%, an additional 20% was set aside as the validation set, while the remaining data were used to train the network. A Bayesian optimization method (Bull, 2011; Gelbart et al., 2014; Snoek et al., 2012) was employed to fine-tune network hyperparameters including the number of LSTM layers and the number of units in the LSTM layer, the dropout factor and the learning rate. A batch size of 16 was used. The network was trained by minimizing the mean squared error (MSE) for 100 epochs with the Adam optimizer (Kingma and Ba, 2017). The LSTM architecture was designed using the Keras module (Chollet et al., 2015) with a Tensorflow backend (Abadi et al., 2016).

### 2.5. Evaluation metrics

Performances were assessed using a combination of quantitative and visual metrics. Quantitative performance was measured through the Root Mean Square Error ($RMSE$), the Mean Absolute Error ($MAE$) and the coefficient of determination ($R^2$) (Witten et al., 2011). The mathematical expressions for these metrics can be found in Eqs. (1)–(3):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{m} (\hat{y}_i - y_i)^2}{m}} \tag{1}$$

$$MAE = \frac{\sum_{i=1}^{m} |\hat{y}_i - y_i|}{m} \tag{2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{m} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{m} (y_i - \bar{y})^2} \tag{3}$$

where $y_i$ is the $i$th observed value, $\hat{y}_i$ is the $i$th predicted value, $\bar{y}$ is the mean of the observed values and $m$ is the dataset size. Another evaluation metric employed was the Percentage Within Bound (PWB) (Chitturi et al., 2021), which represents the percentage of test examples whose predicted values are within a given percentage error bound. The error bound can vary (for example it can be equal to 10%, 20% or 30%), and is selected by the researcher, i.e., PWB10 refers to the percentage of test observations whose predicted values are within 10% of their observed values. Model performances are also evaluated using visual performance metrics, such as scatter plots of observed and predicted values.

## 3. Results

### 3.1. Annual and seasonal salinity patterns

A preliminary statistical analysis was conducted to gain deeper insights into the annual and seasonal patterns of the observed salinity. The Kruskal–Wallis test (Kruskal and Wallis, 1952) was used to determine whether there were statistically significant differences between the medians of different groups. The Kruskal–Wallis test uses ranks of the data, rather than numerical values, to compute the test statistics. The null hypothesis ($H_0$) states that there is no significant difference between the medians of the groups being compared. In other words, all groups come from the same population. The alternative hypothesis

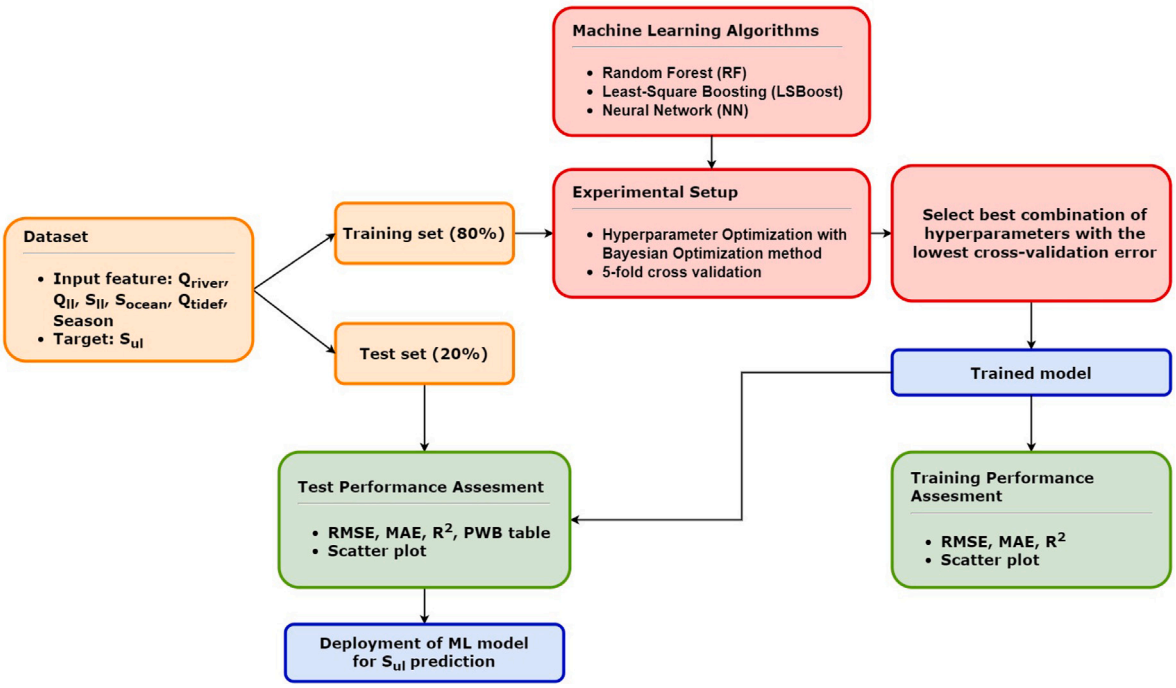


**Fig. 2.** Flow chart of the experimental process used to train and test the ML models to predict the estuarine salinity ($S_{ul}$).

($H_1$) suggests that at least one group's median is significantly different from the others. In addition, Dunn's post-hoc test (Dunn, 1964), was also performed to make pairwise comparisons between groups and to determine which groups were significantly different from each other. The null hypothesis ($H_0$) for a pairwise comparison between group mean ranks $i$ and $j$ states that there is no significant difference between the means: $m_i = m_j$, while, the alternative hypothesis ($H_1$) states that there is a difference between the means: $m_i \neq m_j$. The significance level of both tests was set to 0.01.

#### 3.1.1. Annual salinity patterns

Salinity data were grouped by year and the number of observations for each group is reported in Supplementary Table 1. The boxplots in Fig. 3-a show summary statistics for the observed salinity $S_{ul}$, grouped by year. Data collected in 2016, 2017 and 2019 seem to have very similar pattern, while a very different pattern is shown by the data collected in 2018. To confirm this hypothesis a Kruskal–Wallis test was performed to determine whether there were any statistically significant differences between the medians of different groups. Supplementary Table 2 shows the results of this statistical test. The returned $p$-value ($< 0.01$) indicates that the test rejects the null hypothesis at the 1% significance level. Thus, at least one group's median was significantly different from the others. Dunn's test was thus performed to determine which groups were significantly different from each other. The results provided in Supplementary Table 3 and Fig. 4-a confirm that the data collected in 2016 and 2017 were not significantly different from each other, while all the other groups were significantly different from each other.

#### 3.1.2. Seasonal salinity patterns

The salinity data were grouped by season and the number of observations for each group and are reported in Supplementary Table 1. Boxplots in Fig. 3-b shows summary statistics for the observed salinity $S_{ul}$, grouped by season. Data collected in each season seem to have a very different pattern. To confirm this hypothesis a Kruskal–Wallis test has been performed to determine whether there were any statistically significant differences between the medians of different groups. Results of this statistical test is provided in Supplementary Table 4. The returned $p$-value ($< 0.01$) indicates that the test rejects the null hypothesis at the 1% significance level. Thus, at least one group's median is significantly different from the others Dunn's test was then used to determine which groups were significantly different from each other. The results reported in Supplementary Table 5 and Fig. 4-b confirm that the data collected in each season are significantly different from each other.

### 3.2. Estuary salinity evaluation with machine learning and deep learning algorithms

#### 3.2.1. Machine learning algorithms for estuary salinity estimate combining multiple forcings

ML models, namely RF, LSBoost, and NN, were trained using the following input features, as detailed in Section 2.2: $Q_{river}$, $Q_{tidef}$, $Q_{ll}$, $S_{ocean}$, $S_{ll}$. Additionally, based on the insights provided in Section 3.1.2, the feature $Season$ was included among the other input features, and the models were retrained to assess the predictive power of $Season$. The performances of the ML models were therefore compared to the fully-physics EBM and the Hybrid-EBM model performance. A total of 80% of the data were used to train the models (963 observations), while the remaining 20% (238 observations) were employed to test the models. Supplementary Table 6 presents the optimal hyperparameters used to build the ML models, both with and without the embedded feature $Season$. Training performances are detailed in Supplementary Table 7. The test performances of the ML models (both with and without the $Season$ embedded), compared to those of the EBM and Hybrid-EBM models, are reported in Table 3 and Fig. 5. All three ML models trained without the $Season$ as the input feature show very similar results and perform better than the fully-physics and hybrid models, in terms of each metric. The top-performing ML model without the $Season$ embedding was RF, with an RMSE equals to 3.07 psu, compared with the 4.22 psu and 3.41 psu obtained by the EBM and Hybrid-EBM-LSBoost models, respectively. The MAE decreased from 3.27 psu to 2.34 psu, while the $R^2$ score increased from 0.67 to 0.82. However, all three ML models featuring the $Season$ embedding outperformed those without it and significantly outperformed both the fully-physics EBM and Hybrid-EBM models. In more detail, when $Season$ embedding was

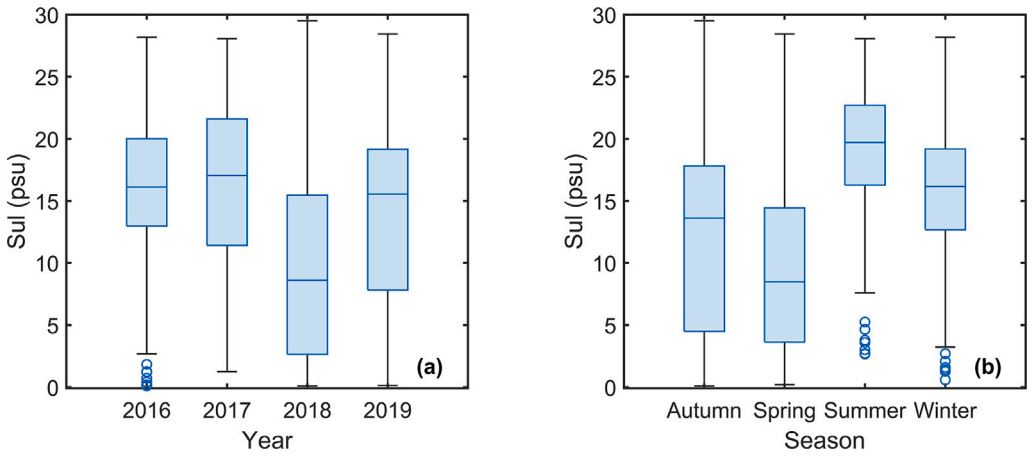


**Fig. 3.** Boxplot of the observed salinity $S_{ul}$: (a) group by year, (b) group by season.

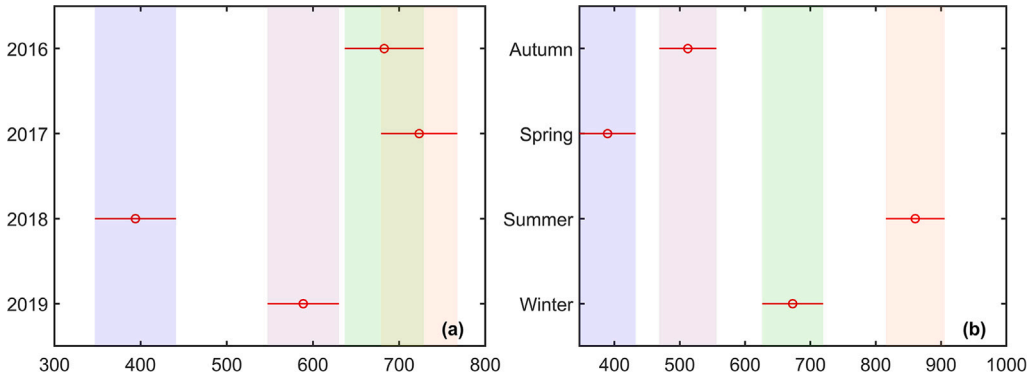


**Fig. 4.** Dunn's test results. Dunn's test performs pairwise comparison between different groups. Each circle represents the group mean ranks, while the comparison interval is represented by a line. Two groups are significantly different if the comparison intervals are disjoint; two groups are not significantly different if the comparison intervals overlap. Salinity observations grouped by: (a) year, (b) season.

**Table 3**
Test performances of ML models for the salinity $S_{ul}$ prediction, trained with and without the feature *Season* embedded. Three ML algorithms are used: RF, LSBoost and NN. The performance metrics consist of Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) in psu and coefficient of determination ($R^2$). Bold characters are used to highlight the best model performance.

| Test performance | | | | | | |
|---|---|---|---|---|---|---|
| Model | No-Season | | | Season | | |
| | RMSE (psu) | MAE (psu) | $R^2$ | RMSE (psu) | MAE (psu) | $R^2$ |
| EBM | 4.22 | 3.27 | 0.67 | – | – | – |
| Hybrid-EBM-RF | 3.46 | 2.64 | 0.78 | – | – | – |
| Hybrid-EBM-LSBoost | 3.41 | 2.63 | 0.78 | – | – | – |
| RF | 3.07 | 2.34 | 0.82 | 2.85 | 2.17 | 0.85 |
| LSBoost | 3.15 | 2.38 | 0.81 | **2.80** | **2.15** | **0.85** |
| NN | 3.17 | 2.41 | 0.81 | 2.88 | 2.23 | 0.84 |

employed, LSBoost demonstrated the best performance, achieving an RMSE of 2.80 psu, MAE of 2.15 psu, and an $R^2$ score of 0.85.

These results were also confirmed by looking at the PWB table (see Table 4), in which, for most of the fixed error bound, ML model performances, with and without the *Season* embedding, outperformed the results obtained with the fully-physics and hybrid models. In particular, with a fixed threshold equal to 10%, RF with *Season* embedding obtained 39.92% of the test data with a lower prediction error than this threshold, compared with the 25.63% obtained with the EBM model and 34.45% obtained with the Hybrid-EBM-LSBoost. Then, with a fixed threshold of 20%, RF with *Season* embedding shows that 65.97% of the test data show a lower prediction error than this threshold, compared to 47.90% and 55.04% with EBM and Hybrid-EBM-LSBoost, respectively. Lastly, with a fixed threshold of 30%, LSBoost with *Season* embedding shows that 76.89% of the test data show a lower prediction error than 30%; for EBM 64.29% of the test data have a prediction error of less than 30%, while for the hybrid models about 69% of the test data have a prediction error lower than 30%. Therefore, ML models, particularly those with a *Season* embedding, ensured a remarkable increase in the overall performance.

Further conclusions can be drawn by assessing the ability of the models to manage both low ($S_{ul} \leq 5$ psu) and high ($S_{ul} \geq 20$ psu) salinity events, considering the following: EBM, Hybrid-EBM-LSBoost, RF-No-Season and LSBoost-Season (see Table 5). For low salinity events, LSBoost-Season demonstrated significant improvements across nearly all fixed thresholds, particularly for PWB10, PWB20 and PWB30. Concerning high salinity events, moderate improvements were observed for PWB10, PWB20, and PWB30 with the inclusion of the *Season* embedding, while more significant improvements were obtained in PWB1 and

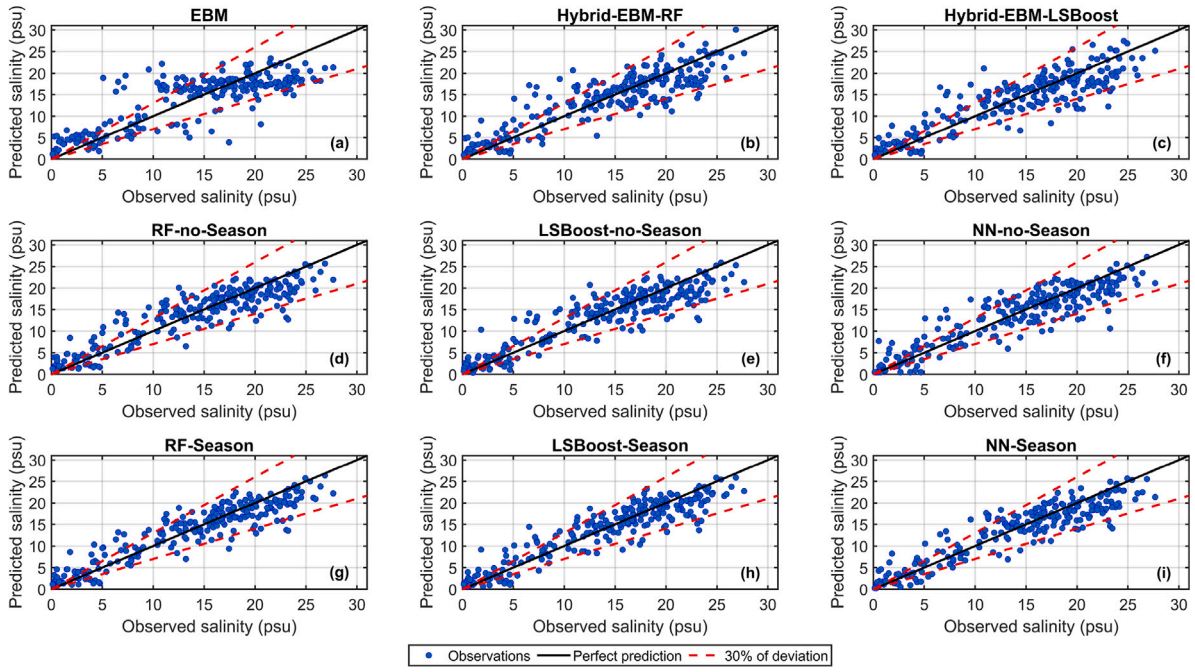


**Fig. 5.** Plots of the observed salinity $S_{ul}$ versus the predicted $S_{ul}$ values for the test dataset. The physics and hybrid models are: (a) EBM, (b) Hybrid-EBM-RF, (c) Hybrid-EBM-LSBoost; ML models trained without the Season embedded as input feature are: (d) RF-no-Season, (e) LSBoost-no-Season, and (f) NN-no-Season. The ML models trained with the Season embedded as input feature are: (g) RF-Season, (h) LSBoost -Season, and (i) NN -Season. The black solid line shows the perfect prediction; the red dashed lines indicate 30% deviation compared to the perfect prediction. Units in psu.

**Table 4**
PWB table reports the percentage of test examples for which the predicted salinity ($S_{ul}$) falls within a specified percentage error bound. Bold characters are used to highlight the best model performance.

| % of test samples | | | | | | |
|---|---|---|---|---|---|---|
| Model | | PWB1 | PWB5 | PWB10 | PWB20 | PWB30 |
| No-Season | EBM | 3.78 | 12.61 | 25.63 | 47.90 | 64.29 |
| | Hybrid-EBM-RF | 2.10 | **21.43** | 34.45 | 54.62 | 69.33 |
| | Hybrid-EBM-LSBoost | 3.78 | 17.23 | 31.93 | 55.04 | 69.75 |
| | RF | **6.72** | 18.49 | 37.39 | 60.08 | 73.11 |
| | LSBoost | 4.20 | 20.17 | 36.97 | 58.82 | 73.95 |
| | NN | 4.62 | 18.91 | 34.87 | 57.14 | 75.63 |
| Season | RF | 3.78 | 20.59 | **39.92** | **65.97** | 76.47 |
| | LSBoost | 4.62 | 18.07 | 38.66 | 65.55 | **76.89** |
| | NN | 3.36 | 19.75 | 36.16 | 64.29 | 76.47 |

PWB5 using RF-No-Season. However, both ML models with and without the *Season* embedding exhibited superior predictive capabilities for both low and high salinity events compared to those of the fully-physics and hybrid models.

### 3.2.2. Evaluation of feature importance

Feature importance was computed for RF and LSBoost models, both with and without the *Season* embedding, to understand what variables significantly influence the predictive capabilities of the models and how they contribute to salinity prediction accuracy (see Supplementary Figure 1 and Fig. 6). Table 6 shows the performance of LSBoost trained with different combinations of the input features. The most significant feature for salinity prediction was identified in $Q_{river}$. In fact, its exclusion during training (LSBoost#1) led to the worst performance. The second most important feature is *Season*. As discussed in Section 3.2.1 LSBoost#6 (trained including *Season*) ensured the best performance of the model, confirming the importance of *Season* embedding. Seasonal changes significantly affect estuarine systems through variations in rainfall and evaporation, which could lead to an increase or decrease in the freshwater discharge. Specific seasons are often associated with more abundant precipitation levels, leading to more freshwater input into estuaries. This process increases the river flow and either pushes back or dilutes the saltwater wedge, thereby limiting its intrusion into the estuary. In contrast, higher temperatures, land-ice contraction and

**Table 5**
PWB table reports the percentage of test examples for which the predicted salinity falls within a specified percentage error bound. Bold characters are used to highlight the best model performance.

| % test samples $S_{ul} \leq 5$ psu | | | | | |
|---|---|---|---|---|---|
| Model | PWB1 | PWB5 | PWB10 | PWB20 | PWB30 |
| EBM | 0 | 0 | 2.33 | 9.30 | 16.28 |
| Hybrid-EBM-LSBoost | 0 | **6.98** | 9.30 | 16.28 | 20.93 |
| RF-No-Season | 0 | 2.33 | 4.65 | 11.63 | 16.28 |
| LSBoost-Season | **2.33** | 4.65 | **11.63** | **27.91** | **35.56** |
| % test samples $S_{ul} \geq 20$ psu | | | | | |
| EBM | 3.64 | 9.09 | 14.55 | 50.91 | 87.27 |
| Hybrid-EBM-LSBoost | 5.45 | 14.55 | 36.36 | 72.73 | 89.09 |
| RF-No-Season | **12.73** | **29.09** | 45.45 | 78.18 | 90.91 |
| LSBoost-Season | 7.27 | 23.64 | **50.91** | **80.00** | **92.73** |

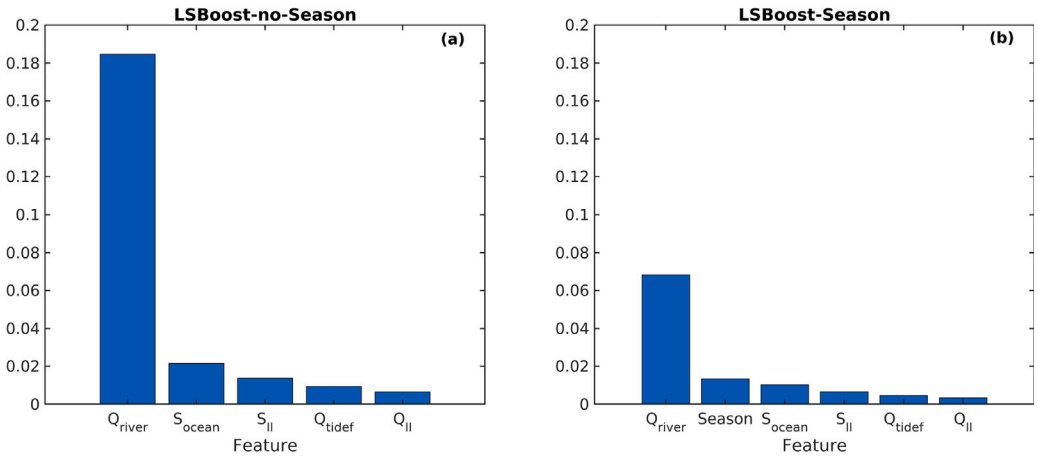


**Fig. 6.** Feature Importance. ML models trained with and without the Season embedded as input feature are: (a) LSBoost-no-Season, (b) LSBoost-Season.

**Table 6**
Test performances of LSBoost models for salinity $S_{ul}$ prediction, trained with different input features. The performance metrics consist of Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) in psu along with the coefficient of determination ($R^2$). Bold characters are used to highlight the best model performance.

| Test performance | | | | | |
|---|---|---|---|---|---|
| Model | Input features | Removed feature | RMSE (psu) | MAE (psu) | $R^2$ |
| LSBoost#1 | $Q_{ll}$, $S_{ll}$, $S_{ocean}$, $Q_{tidef}$, $Season$ | $Q_{river}$ | 5.97 | 4.84 | 0.33 |
| LSBoost#2 | $Q_{river}$, $Q_{ll}$, $S_{ll}$, $Q_{tidef}$, $Season$ | $S_{ocean}$ | 3.32 | 2.55 | 0.79 |
| LSBoost#3 | $Q_{river}$, $S_{ll}$, $S_{ocean}$, $Q_{tidef}$, $Season$ | $Q_{ll}$ | 3.15 | 2.44 | 0.81 |
| LSBoost#4 | $Q_{river}$, $Q_{ll}$, $S_{ocean}$, $Q_{tidef}$, $Season$ | $S_{ll}$ | 3.11 | 2.46 | 0.82 |
| LSBoost#5 | $Q_{river}$, $Q_{ll}$, $S_{ll}$, $S_{ocean}$, $Season$ | $Q_{tidef}$ | 2.99 | 2.33 | 0.83 |
| LSBoost#6 (LSBoost-Season) | $Q_{river}$, $Q_{ll}$, $S_{ll}$, $S_{ocean}$, $Q_{tidef}$, $Season$ | – | **2.80** | **2.15** | **0.85** |

rainfall scarcity can lead to an increase in evaporation and a decrease in river runoff (Verri et al., 2024b), thus favoring the intrusion of saltier seawater along rivers and shifting the river release toward higher salinity levels (Bellafiore et al., 2021; Conroy et al., 2020). The third most important feature was $S_{ocean}$. Thus, excluding it led to the second worst performance (LSBoost#2), since it is fundamental in regulating the dynamics of saltwater intrusion. Additionally, fluctuations in sub-seasonal tidal cycles (i.e. the spring and neap tides) may play a key role in the movement and mixing of water masses within estuaries. The exclusion of $Q_{ll}$, $S_{ll}$ and $Q_{tidef}$ appeared to provide very similar results, as reported by LSBoost#3, LSBoost#4 and LSBoost#5. In summary, as confirmed by LSBoost#6, the seasonal fluctuations in rainfall, river flow, temperature and evapoconcentration processes as well as the tidally driven patterns at daily to sub-seasonal scales, collectively shape the dynamics of salt wedge intrusion in estuaries. Understanding these seasonal variations holds immense importance in comprehending and accurately predicting changes in salinity levels.

### 3.2.3. Deep learning algorithms for estuary salinity prediction

LSTM was used to establish the most effective forecasting schema for predicting daily salinity ($S_{ul}$) from 1-step ahead to 7-steps ahead. Three distinct schemas were developed based on the employed features: one using only $S_{ul}$ (LSTM-Sul), another using only $Q_{river}$ (LSTM-Qriver), and finally, a schema incorporating all the following features (LSTM-All): $Q_{river}$, $Q_{tidef}$, $Q_{ll}$, $S_{ocean}$, and $S_{ll}$. For each schema, the analysis considered input time steps ranging from 1 (one day) to 7 (one week) previous days to establish the number of preceding daily input steps required for implementing 1-step ahead forecasting. After determining the optimal number of previous input time steps for each schema, the models and their best configurations are selected to evaluate their performance for 7-step ahead forecasting.

The training and test performances are presented in Supplementary Tables 8 and 9, and Supplementary Figures 2 and 7 for the 1-step ahead models, across the three different schemas and seven distinct configurations of input time steps. The test performances are summarized in Fig. 7. The LSTM-Sul model demonstrates the best performance for a 1-step ahead prediction using three previous days as input ($n_{in}$ = 3), resulting in an RMSE of 2.67 psu, an MAE of 2.01 psu, and an $R^2$ score of 0.87. At the same time, the LSTM-Qriver achieved the best performance using four previous days as input ($n_{in}$ = 4), resulting in an RMSE of 3.07 psu, MAE of 2.31 psu, and an $R^2$ score of 0.83. The LSTM-All model obtained the best performance using only one previous day as input ($n_{in}$ = 1), with an RMSE of 3.19 psu, an MAE of 2.39, and an $R^2$ score of 0.82.

Three input time steps were therefore fixed ($n_{in}$ = 3) for LSTM-Sul, four input time steps ($n_{in}$ = 4) for LSTM-Qriver, and one input time step ($n_{in}$ = 1) for LSTM-All. Based on these configurations, three LSTM models were selected to carry out 7-step ahead predictions of daily salinity. Supplementary Table 10 presents the optimal hyperparameters used to build the three LSTM models for 7-step ahead predictions. The training performances are detailed in Supplementary Table 11. The results on the test set are summarized in Fig. 8, Fig. 9, and detailed in Supplementary Table 12, and Supplementary Figures 8–10. As expected, all three LSTM models presented the same trend, showing a decrease in model performance with each consecutive step ahead. In terms of overall performance, all three models showed similar results. However, LSTM-Qriver slightly outperformed the other models, with an average RMSE of 4.19 psu, MAE of 3.11 psu, and an $R^2$ score of 0.69. Similarly, LSTM-All exhibited the second-best performance, with an average RMSE of 4.25 psu, MAE of 3.27 psu, and an $R^2$ score of 0.68. LSTM-Sul demonstrated a slightly weaker performance, with an average RMSE of 4.67 psu, MAE of 3.44 psu, and an $R^2$ score of 0.61.

For the 1-step ahead predictions, the best performances were achieved by LSTM-Sul with an RMSE and MAE of 2.85 psu and 2.19 psu, respectively, along with an $R^2$ score of 0.86. It outperformed both LSTM-Qriver and LSTM-All at the same step ahead. However, for all subsequent steps ahead, LSTM-Sul performed worse than LSTM-Qriver and LSTM-All, as indicated by its higher RMSE and MAE, and lower $R^2$ scores. When comparing LSTM-Qriver and LSTM-All, the

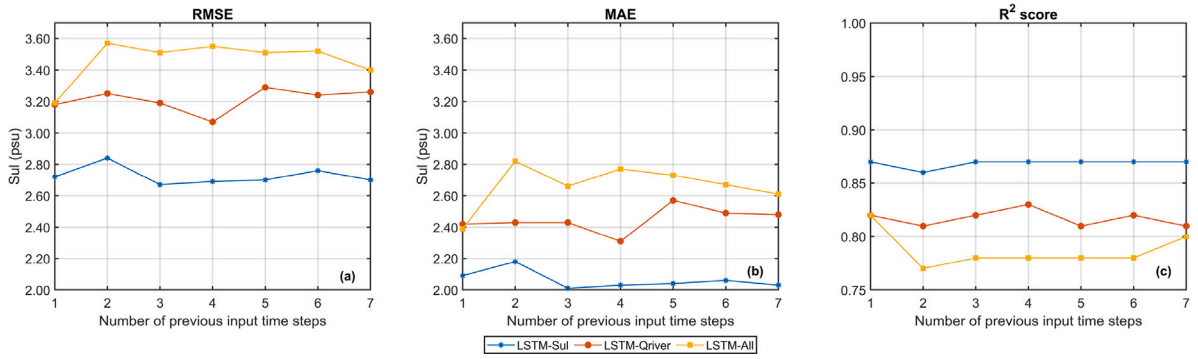


**Fig. 7.** Performance reached on the test set for 1-step ahead LSTM. Three different schemas are used: LSTM-ul means that the models are trained using only the $S_{ul}$ as input; LSTM-Qriver using only the $Q_{river}$ as input; and LSTM-All using the $Q_{river}$, $Q_{tidef}$, $Q_{ll}$, $S_{ocean}$, $S_{ll}$ as inputs. The performance metrics consist of (a) Root Mean Square Error (RMSE) and (b)Mean Absolute Error (MAE) in psu, and (c) coefficient of determination ($R^2$).

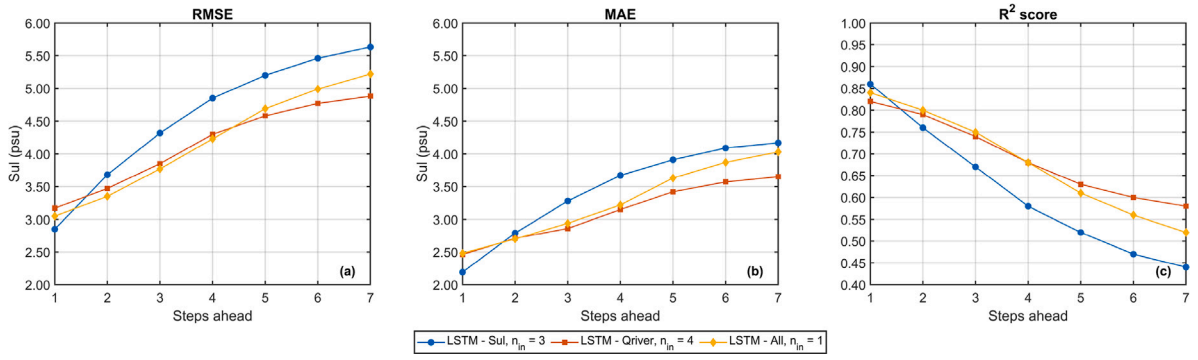


**Fig. 8.** Performance achieved on the test set for the 7-step-ahead LSTM. Three different schemas are used: LSTM-Sul means that the models are trained using only the $S_{ul}$ as input; LSTM-Qriver using only the $Q_{river}$ as input; LSTM-All using the $Q_{river}$, $Q_{tidef}$, $Q_{ll}$, $S_{ocean}$, $S_{ll}$ as input. Performance metrics consist of (a) Root Mean Square Error (RMSE) and (b)Mean Absolute Error (MAE) in psu, and (c) coefficient of determination ($R^2$).

results showed no significant differences in terms of performance. They exhibited very similar behaviors in predicting from 1-step ahead up to 5-steps ahead, with an average RMSE ranging from approximately 3 psu to 4.50 psu, an MAE ranging from 2.50 psu to 3.50 psu, and $R^2$ scores ranging from approximately 0.85 to 0.60. Lastly, LSTM-Qriver performed better for 6-steps ahead and 7-steps ahead, with the lowest RMSE and MAE, and highest $R^2$ scores, compared to LSTM-All. These results are also confirmed in Fig. 9, which shows scatter diagrams of the results obtained from 1-step ahead to 7-step ahead predictions. The results, show that, in general, the LSTM-Qriver and LSTM-All models exhibit the best performances. The best results were obtained in the 1-step ahead prediction by LSTM-Sul, with an $R^2$ score higher than 0.85. LSTM-Qriver and LSTM-All also performed well, with $R^2$ scores higher than 0.60 for predictions from 2-steps ahead up to 5-steps ahead.

## 4. Discussion and conclusions

This study presents the design of a tool that predicts salinity levels at the estuarine river mouth, employing ML and DL techniques. All the ML algorithms demonstrated good predictive capabilities for estuarine salinity, of which the best performance was shown by LSBoost. The comprehensive statistical analysis, which was conducted to extract information regarding yearly and seasonal salinity patterns, confirms statistically significant differences between the salinity levels collected in the four seasons, and highlights that the significantly highest salinity levels were recorded in the summer. This is in line with other studies in the literature and it is mainly due to the evapoconcentration processes during the dry summer season, which has also been observed in other coastal-lagoon low-inflow estuaries (Cardoso-Mohedano et al., 2018; Nascimento et al., 2021; Mastrocicco et al., 2020). The results also suggest including seasonal information among the variable inputs used to train ML models, which showed better performances than physics-based and hybrid models. The most significant features, evaluated by the LSBoost algorithm for salinity prediction, thus include river flow and seasonal period, confirming results discussed in recent papers (Verri et al., 2024b; Bellafiore et al., 2021; Conroy et al., 2020). Finally, the study moves from the idea of predicting salinity levels without considering temporal dependencies in the data, to the concept of salinity forecasting by using DL, which considers temporal dependencies. LSTM models trained with different feature combinations showed different performances, with the LSTM-Qriver model achieving the best overall result, thus confirming the significant role that $Q_{river}$ plays in the salt wedge intrusion phenomenon. All LSTM models exhibited a decrease in performance with an increase in the number of prediction steps, indicating a greater difficulty in long-term prediction. The results show that both ML and DL models can be used to improve estuarine salinity predictions, providing crucial information for water resource management and coastal ecosystem conservation.

There are numerous opportunities offered by ML and DL models for predicting and forecasting salinity levels, albeit accompanied by certain limitations. In fact, ML-based approaches have been shown to be promising in recognizing complex patterns in data and offer high flexibility, making them easily adaptable to different data types and problem domains. They can scale efficiently to handle large datasets, capture non-linear relationships between variables, and generalize well to unseen data. In general, DL models are black box models because of their lack both of transparency and explicit interpretability. On the other hand, physics-based approaches rely on equations, laws and principles that govern the behavior of physical systems, providing a deep understanding of the underlying mechanisms. They can also be easily interpretable and understood based on physical principles. Physics-based models are generally based on certain assumptions and

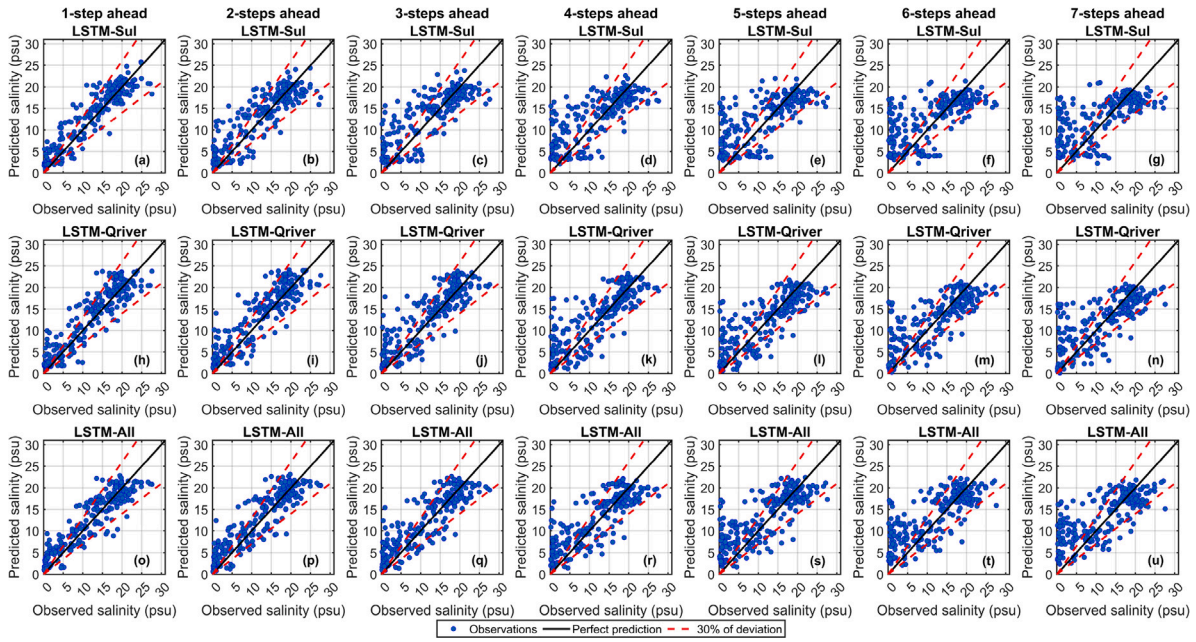


**Fig. 9.** Plots of the observed salinity $S_{ul}$ versus the predicted $S_{ul}$ values for the test dataset. The models are: (a–g) LSTM-Sul $n_{in}$ = 3 for 7-step ahead predictions; (h–n) LSTM-Qriver $n_{in}$ = 4 for 7-step ahead predictions; and (o-u) LSTM-All $n_{in}$ = 1 for 7-step ahead predictions. The black solid line shows the perfect prediction; the red dashed lines indicate 30% deviation compared to the perfect prediction. Units in psu.

may fail to capture complex nonlinearities or unknown phenomena if they are not explicitly included in the model. The physics-based model considered as a baseline in this study suffers from a low generalization ability and requires customization to specific study areas. Moreover, it requires different equations, which can be challenging to learn due to the limited number of observations. ML-based approaches offer a solution by learning the mapping between variables and salinity values directly, bypassing the need for intricate equations and potentially improving prediction accuracy.

An important limitation of this type of study lies in the relatively limited availability of data. In particular, this study relies on data collected between 2016 and 2019 at the Po di Goro River mouth. Incorporating a longer historical dataset to capture long-term trends and rare events would be beneficial. Moreover, collecting and integrating data from multiple estuaries with diverse characteristics would enhance the models' generalization, helping to develop models that are robust across different environmental settings. In line with the suggestions presented in the recent literature (Kratzert et al., 2024), the longer-term strategic objective of the study is to build an artificial intelligence system trained on a large number of basins in the Mediterranean region and hopefully beyond. Further directions should involve the development of specific strategies for handling missing values in input data, which could be used to compensate for the scarcity of available salinity data.

In addition, the spatial resolution of the Copernicus Marine Environment Monitoring Service (CMEMS) data, here used, is approximately 4.5 km. Utilizing higher spatial resolution datasets would better capture the fine-scale variations and high-frequency dynamics affecting estuarine salinity.

The present study forecasts estuarine salinity levels up to seven days beyond the input data and longer forecasting horizons were not explored, which might be a future challenge.

Finally, the variables here considered can be further expanded to encompass a more detailed description of the environment. Integrating external factors related to anthropogenic influences and meteorological-oceanographic features into the modeling framework to account for their impact on estuarine salinity could present a significant opportunity. For example, future plans will further develop EBM integrated with ML and DL based strategies, including equations and/or algorithms for estuarine nutrient budgets to support water quality assessments near river mouths.

The operative purpose of our work is a preliminary investigation and presentation of the landscape of ML and DL techniques on the Po River. The aim of this paper is therefore to provide an initial overview of the problem we are addressing, laying the groundwork for a broader discussion in the hopefully not too distant future.

## Computer Code Availability

Name of the code/library: CMCC-ML-Estuary-Salinity-Estimation

Contact: leonardo.saccotelli@cmcc.it

Year first available: 2024

Hardware requirements: Any Intel or AMD x86-64 processor with two or more cores, RAM 8 GB (Minimum)

Program language: MATLAB, Python

Software required: MATLAB 2023b, Python 3.10.12, and Jupyter Server

Program size: 145 MB

The source codes are available for downloading here: https://github.com/CMCC-Foundation/CMCC-Hybrid-EstuaryBoxModel/tree/ML-EBM-PoGoro

## CRediT authorship contribution statement

**Leonardo Saccotelli:** Methodology, Software, Validation, Investigation, Writing – original draft, Writing – review & editing. **Giorgia Verri:** Conceptualization, Resources, Data curation, Writing – original draft. **Alessandro De Lorenzis:** Resources, Data curation, Validation, Writing – original draft. **Carla Cherubini:** Writing – original draft, Writing – review & editing. **Rocco Caccioppoli:** Software. **Giovanni Coppini:** Supervision. **Rosalia Maglietta:** Supervision, Conceptualization, Methodology, Validation, Writing – original draft, Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

https://github.com/CMCC-Foundation/CMCC-Hybrid-EstuaryBoxModel/tree/ML-EBM-PoGoro.

## Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT 3.5 to improve language and readability. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication. This paper was also edited by a professional scientific editing agency.

## Acknowledgments

The work was supported by the EstuarIO project, funded by Mercator Ocean under the Copernicus Marine - Innovation Science Evolution projects LOT N° 2 - 21036L02-COP-INNO SCI-9100 - MARINE COASTAL ENVIRONMENT. The authors would like to thank Michele Attolico for his technical assistance.

## Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.acags.2024.100173.

## References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mane, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viegas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X., 2016. TensorFlow: Large-scale machine learning on heterogeneous distributed systems. arXiv:1603.04467.

Arpae, 2024a. Observations of river discharge performed at Pontelagoscuro station, Agenzia Regionale per la Prevenzione, l'Ambiente e l'Energia dell'Emilia Romagna, URL: https://simc.arpae.it/dext3r/.

Arpae, 2024b. Observations of salinity performed at manufatto station, Agenzia Regionale per la Prevenzione, l'Ambiente e l'Energia dell'Emilia Romagna, URL: https://simc.arpae.it/dext3r/.

Barbier, E.B., Hacker, S.D., Kennedy, C., Koch, E.W., Stier, A.C., Silliman, B.R., 2011. The value of estuarine and coastal ecosystem services. Ecol. Monogr. 81 (2), 169–193. http://dx.doi.org/10.1890/10-1510.1, URL: https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/10-1510.1.

Bellafiore, D., Ferrarin, C., Maicu, F., Manfè, G., Lorenzetti, G., Umgiesser, G., Zaggia, L., Levinson, A.V., 2021. Saltwater intrusion in a Mediterranean Delta under a changing climate. J. Geophys. Res.: Oceans 126 (2), http://dx.doi.org/10.1029/2020JC016437, arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020JC016437, URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020JC016437, e2020JC016437 2020JC016437.

Boerema, A., Meire, P., 2017. Management for estuarine ecosystem services: A review. Ecol. Eng. 98, 172–182. http://dx.doi.org/10.1016/j.ecoleng.2016.10.051, URL: https://www.sciencedirect.com/science/article/pii/S0925857416305791.

Boldrin, A., Langone, L., Miserocchi, S., Turchetto, M., Acri, F., 2005. Po River plume on the Adriatic continental shelf: Dispersion and sedimentation of dissolved and suspended matter during different river discharge rates. Mar. Geol. 222, 135–158. http://dx.doi.org/10.1107/S1600576721010840.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Bull, A.D., 2011. Convergence rates of efficient global optimization algorithms. arXiv:1101.3501.

Cardoso-Mohedano, J.-G., Lima-Rego, J., Sanchez-Cabeza, J.-A., Ruiz-Fernández, A.-C., Canales-Delgadillo, J., Sánchez-Flores, E.-I., Páez-Osuna, F., 2018. Sub-tropical coastal lagoon salinization associated to shrimp ponds effluents. Estuar. Coast. Shelf Sci. 203, 72–79. http://dx.doi.org/10.1016/j.ecss.2018.01.022, URL: https://www.sciencedirect.com/science/article/pii/S0272771417306492.

Chitturi, S.R., Ratner, D., Walroth, R.C., Thampy, V., Reed, E.J., Dunne, M., Tassone, C.J., Stone, K.H., 2021. Automated prediction of lattice parameters from X-ray powder diffraction patterns. J. Appl. Crystallogr. 54 (6), 1799–1810. http://dx.doi.org/10.1107/S1600576721010840.

Chollet, F., et al., 2015. Keras. GitHub, https://github.com/fchollet/keras.

Conroy, T., Sutherland, D.A., Ralston, D.K., 2020. Estuarine exchange flow variability in a seasonal, Segmented Estuary. J. Phys. Oceanogr. 50, 595–613, URL: https://api.semanticscholar.org/CorpusID:213005984.

Cushman-Roisin, B., Gacić, M., Poulain, P.-M., Artegiani, A., 2002. Physical Oceanography of the Adriatic Sea: Past, Present and Future. Kluwer Academic Publishers.

Dimauro, G., Camporeale, M.G., Dipalma, A., Guarini, A., Maglietta, R., 2023a. Anaemia detection based on sclera and blood vessel colour estimation. Biomed. Signal Process. Control 81, 104489. http://dx.doi.org/10.1016/j.bspc.2022.104489, URL: https://www.sciencedirect.com/science/article/pii/S1746809422009430.

Dimauro, G., Griseta, M.E., Camporeale, M.G., Clemente, F., Guarini, A., Maglietta, R., 2023b. An intelligent non-invasive system for automated diagnosis of anemia exploiting a novel dataset. Artif. Intell. Med. 136, 102477. http://dx.doi.org/10.1016/j.artmed.2022.102477, URL: https://www.sciencedirect.com/science/article/pii/S0933365722002299.

Dunn, O.J., 1964. Multiple comparisons using rank sums. Technometrics 6 (3), 241–252, URL: http://www.jstor.org/stable/1266041.

Egbert, G.D., Erofeeva, S.Y., 2002. Efficient inverse modeling of barotropic ocean tides. J. Atmos. Ocean. Technol. 19 (2), 183–204. http://dx.doi.org/10.1175/1520-0426(2002)019<0183:EIMOBO>2.0.CO;2, URL: https://journals.ametsoc.org/view/journals/atot/19/2/1520-0426_2002_019_0183_eimobo_2_0_co_2.xml.

Escudier, R., Clementi, E., Cipollone, A., Pistoia, J., Drudi, M., Grandi, A., Lyubartsev, V., Lecci, R., Aydogdu, A., Delrosso, D., Omar, M., Masina, S., G., C., Pinardi, N., 2021. Mediterranean Sea Physical Reanalysis INTERIM (CMEMS MED-Currents, E3R1i system) (Version 1) Data set. Copernicus Monitoring Environment Marine Service (CMEMS). http://dx.doi.org/10.25423/CMCC/MEDSEA_MULTIYEAR_PHY_006_004_E3R1I.

Escudier, R., Clementi, E., Omar, M., Cipollone, A., Pistoia, J., Aydogdu, A., Drudi, M., Grandi, A., Lyubartsev, V., Lecci, R., Cretí, S., Masina, S., Coppini, G., Pinardi, N., 2020. Mediterranean Sea Physical Reanalysis (CMEMS MED-Currents) (Version 1) Data set. Copernicus Monitoring Environment Marine Service (CMEMS). http://dx.doi.org/10.25423/CMCC/MEDSEA_MULTIYEAR_PHY_006_004_E3R1.

Fang, Y., wei Chen, X., Cheng, N.-S., 2017. Estuary salinity prediction using a coupled GA-SVM model: A case study of the Min River Estuary, China. Water Sci. Technol.: Water Supply 17, 52–60.

Friedman, J.H., 2001. Greedy function approximation: A gradient boosting machine. Ann. Statist. 29 (5), 1189–1232, URL: http://www.jstor.org/stable/2699986.

Gelbart, M.A., Snoek, J., Adams, R.P., 2014. Bayesian optimization with unknown constraints. arXiv:1403.5607.

Ghalambor, C.K., Gross, E.S., Grosholz, E.D., Jeffries, K.M., Largier, J.L., McCormick, S.D., Sommer, T., Velotta, J.P., Whitehead, A., 2021. Ecological effects of climate-driven salinity variation in the San Francisco Estuary: Can we anticipate and manage the coming changes? San Francisco Estuary Watershed Sci. 19 (2), 1–30. http://dx.doi.org/10.15447/SFEWS.2021V19ISS2ART3.

Goodfellow, I., Bengio, Y., Courvillei, A., 2016. Deep Learning. MIT Press.

Guillou, N., Chapalain, G., Petton, S., 2023. Predicting sea surface salinity in a tidal estuary with machine learning. Oceanologia 65 (2), 318–332. http://dx.doi.org/10.1016/j.oceano.2022.07.007, URL: https://www.sciencedirect.com/science/article/pii/S0078323422000835.

Hastie, T., Tibshirani, R., Friedman, J.H., Friedman, J.H., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, vol. 2, Springer.

Herbert, E.R., Boon, P., Burgin, A.J., Neubauer, S.C., Franklin, R.B., Ardón, M., Hopfensperger, K.N., Lamers, L.P., Gell, P., 2015. A global perspective on wetland salinization: Ecological consequences of a growing threat to freshwater wetlands. Ecosphere 6 (10), http://dx.doi.org/10.1890/ES14-00534.1, URL: https://par.nsf.gov/biblio/10084337.

Hoai, P., Quoc, P., Thai, T.T., 2022. Apply machine learning to predict saltwater intrusion in the Ham Luong River, Ben Tre Province. VNU J. Sci.: Earth Environ. Sci. 38 (3), http://dx.doi.org/10.25073/2588-1094/vnuees.4852, URL: https://js.vnu.edu.vn/EES/article/view/4852.

Hochreiter, S., Schmidhuber, J., 1997. Long Short-Term Memory. Neural Comput. 9 (8), 1735–1780. http://dx.doi.org/10.1162/neco.1997.9.8.1735, arXiv:https://direct.mit.edu/neco/article-pdf/9/8/1735/813796/neco.1997.9.8.1735.pdf.

Keyes, A.A., McLaughlin, J.P., Barner, A.K., Dee, L.E., 2021. An ecological network approach to predict ecosystem service vulnerability to species losses. Nature Commun. 12 (1), 1586. http://dx.doi.org/10.1038/s41467-021-21824-x.

Kingma, D.P., Ba, J., 2017. Adam: A method for stochastic optimization. arXiv:1412.6980.

Kratzert, F., Gauch, M., Klotz, D., Nearing, G., 2024. HESS opinions: Never train an LSTM on a single basin. Hydrol. Earth Syst. Sci. Discuss. 2024, 1–19. http://dx.doi.org/10.5194/hess-2023-275, URL: https://hess.copernicus.org/preprints/hess-2023-275/.

Kruskal, W.H., Wallis, W.A., 1952. Use of ranks in one-criterion variance analysis. J. Amer. Statist. Assoc. 47 (260), 583–621. http://dx.doi.org/10.1080/01621459.1952.10483441.

Kurdistani, S.M., Verri, G., Pinardi, N., Coppini, G., 2022. Climate projections of salt-wedge intrusions in a Po river branch (Northern Adriatic Sea). Fondazione Centro Euro-Mediterraneo Sui Cambiamenti Climatici.

Lal, A., Datta, B., 2018. Development and implementation of support vector machine regression surrogate models for predicting groundwater pumping-induced saltwater intrusion into coastal aquifers. Water Resour. Manag. 32 (7), 2405–2419. http://dx.doi.org/10.1007/s11269-018-1936-2.

Liu, M., Liu, X., Liu, D., Ding, C., Jiang, J., 2015. Multivariable integration method for estimating sea surface salinity in coastal waters from in situ data and remotely sensed data using random forest algorithm. Comput. Geosci. 75, 44–56. http://dx.doi.org/10.1016/j.cageo.2014.10.016, URL: https://www.sciencedirect.com/science/article/pii/S0098300414002490.

Lu, P., Lin, K., Xu, C.-Y., Lan, T., Liu, Z., He, Y., 2021. An integrated framework of input determination for ensemble forecasts of monthly estuarine saltwater intrusion. J. Hydrol. 598, 126225. http://dx.doi.org/10.1016/j.jhydrol.2021.126225, URL: https://www.sciencedirect.com/science/article/pii/S0022169421002729.

Maglietta, R., Bussola, A., Carlucci, R., Fanizza, C., Dimauro, G., 2023a. ARIANNA: A novel deep learning-based system for fin contours analysis in individual recognition of dolphins. Intell. Syst. Appl. 18, 200207. http://dx.doi.org/10.1016/j.iswa.2023.200207, URL: https://www.sciencedirect.com/science/article/pii/S2667305323000327.

Maglietta, R., Carlucci, R., Fanizza, C., Dimauro, G., 2022. Machine learning and image processing methods for cetacean photo identification: A systematic review. IEEE Access 10, 80195–80207. http://dx.doi.org/10.1109/ACCESS.2022.3195218.

Maglietta, R., Milella, R., Caccia, M., Bruzzone, G., 2018. A vision-based system for robotic inspection of marine vessels. SIViP 12, 471–478.

Maglietta, R., Saccotelli, L., Fanizza, C., Telesca, V., Dimauro, G., Causio, S., Lecci, R., Federico, I., Coppini, G., Cipriano, G., Carlucci, R., 2023b. Environmental variables and machine learning models to predict cetacean abundance in the Central-eastern Mediterranean Sea. Sci. Rep. 13 (1), 2600.

Maglietta, R., Verri, G., Saccotelli, L., Delorenzis, A., Cherubini, C., Caccioppoli, R., Dimauro, G., Coppini, G., 2023c. Advancing estuarine box modeling: A novel hybrid machine learning and physics-based approach. (submitted for publication).

Mastrocicco, M., Busico, G., Colombani, N., Usai, A., Ruberti, D., 2020. Seasonal salinity variations in a coastal wetland induced by complex interactions between sea, river and evapoconcentration processes. In: Nguyen, K.D., Guillou, S., Gourbesville, P., Thiébot, J. (Eds.), Estuaries and Coastal Zones in Times of Global Change. Springer Singapore, Singapore, pp. 77–88.

Mohamad, J.A., Mohamad, R.K., Malihe, D., Jason, A., Shahabbodin, S., Kwok-Wing, C., 2018. Effect of river flow on the quality of estuarine and coastal waters using machine learning models. Eng. Appl. Comput. Fluid Mech. 12 (1), 810–823. http://dx.doi.org/10.1080/19942060.2018.1528480.

Nascimento, Â., Biguino, B., Borges, C., Cereja, R., Cruz, J.P.C., Sousa, F., Dias, J., Brotas, V., Palma, C., Brito, A.C., 2021. Tidal variability of water quality parameters in a mesotidal estuary (Sado Estuary, Portugal). Sci. Rep. 11 (1), 23112. http://dx.doi.org/10.1038/s41598-021-02603-6, URL: https://europepmc.org/articles/PMC8633344.

Nguyen, T.G., Tran, N.A., Vu, P.L., Nguyen, Q.-H., Nguyen, H.D., Bui, Q.-T., 2021. Salinity intrusion prediction using remote sensing and machine learning in data-limited regions: A case study in Vietnam's Mekong Delta. Geoderma Reg. 27, e00424. http://dx.doi.org/10.1016/j.geodrs.2021.e00424, URL: https://www.sciencedirect.com/science/article/pii/S2352009421000699.

Nigam, T., Escudier, R., Pistoia, J., Aydogdu, A., Omar, M., Clementi, E., Cipollone, A., Drudi, M., Grandi, A., Mariani, A., Lyubartsev, V., Lecci, R., Cretí, S., Masina, S., Coppini, G., Pinardi, N., 2021. Mediterranean Sea Physical Reanalysis (CMEMS MED-Currents) (Version 1) Data set. Copernicus Monitoring Environment Marine Service (CMEMS). http://dx.doi.org/10.25423/CMCC/MEDSEA_MULTIYEAR_PHY_006_004_E3R1.

Nocedal, J., Wright, S.J., 2006. Numerical Optimization, 2e ed Springer, New York, NY, USA.

Qi, S., He, M., Bai, Z., Ding, Z., Sandhu, P., Chung, F., Namadi, P., Zhou, Y., Hoang, R., Tom, B., Anderson, J., Roh, D.M., 2022a. Novel salinity modeling using deep learning for the Sacramento-San Joaquin Delta of California. Water 14 (22), http://dx.doi.org/10.3390/w14223628, URL: https://www.mdpi.com/2073-4441/14/22/3628.

Qi, S., He, M., Bai, Z., Ding, Z., Sandhu, P., Zhou, Y., Namadi, P., Tom, B., Hoang, R., Anderson, J., 2022b. Multi-location emulation of a process-based salinity model using machine learning. Water 14 (13), http://dx.doi.org/10.3390/w14132030, URL: https://www.mdpi.com/2073-4441/14/13/2030.

Qiu, C., Wan, Y., 2013. Time series modeling and prediction of salinity in the Caloosahatchee River Estuary. Water Resour. Res. 49 (9), 5804–5816. http://dx.doi.org/10.1002/wrcr.20415, arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/wrcr.20415, URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/wrcr.20415.

Rath, J.S., Hutton, P.H., Chen, L., Roy, S.B., 2017. A hybrid empirical-Bayesian artificial neural network model of salinity in the San Francisco Bay-Delta estuary. Environ. Model. Softw. 93, 193–208. http://dx.doi.org/10.1016/j.envsoft.2017.03.022, URL: https://www.sciencedirect.com/science/article/pii/S1364815216305862.

Refaeilzadeh, P., Tang, L., Liu, H., 2016. Cross-validation. In: Encyclopedia of Database Systems. Springer New York, New York, NY, pp. 1–7. http://dx.doi.org/10.1007/978-1-4899-7993-3_565-2.

Rodriguez, J.D., Perez, A., Lozano, J.A., 2010. Sensitivity analysis of k-Fold cross validation in prediction error estimation. IEEE Trans. Pattern Anal. Mach. Intell. 32 (3), 569–575. http://dx.doi.org/10.1109/TPAMI.2009.187.

Saccotelli, L., Verri, G., De Lorenzis, A., Caccioppoli, R., Cherubini, C., Dimauro, G., Coppini, G., Maglietta, R., 2023. Estuary salinity prediction using a support vector machine based approach: A case study of the po di goro estuary. In: 2023 IEEE International Workshop on Metrology for the Sea; Learning to Measure Sea Health Parameters. MetroSea, pp. 294–298. http://dx.doi.org/10.1109/MetroSea58055.2023.10317103.

Snoek, J., Larochelle, H., Adams, R.P., 2012. Practical Bayesian optimization of machine learning algorithms. arXiv:1206.2944.

Tarolli, P., Luo, J., Straffelini, E., Liou, Y.-A., Nguyen, K.-A., Laurenti, R., Masin, R., D'Agostino, V., 2023. Saltwater intrusion and climate change impact on coastal agriculture. PLoS Water 2 (4), 1–5. http://dx.doi.org/10.1371/journal.pwat.0000121.

The MathWorks Inc., 2023a. Fitrensemble. The MathWorks Inc., Natick, Massachusetts, United States, URL: https://it.mathworks.com/help/stats/fitrensemble.html.

The MathWorks Inc., 2023b. Fitrnet. The MathWorks Inc., Natick, Massachusetts, United States, URL: https://it.mathworks.com/help/stats/fitrnet.html.

The MathWorks Inc., 2023c. MATLAB Version: 23.2.0 (R2023b). The MathWorks Inc., Natick, Massachusetts, United States, URL: https://www.mathworks.com.

The MathWorks Inc., 2023d. Statistics and Machine Learning Toolbox: 23.2 (R2023b). The MathWorks Inc., Natick, Massachusetts, United States, URL: https://it.mathworks.com/.

Townsend, M.J., Hewitt, J.E., Davies, K.K., Lohrer, A.M., Lundquist, C.J., Cartner, K., 2013. The many uses and values of estuarine ecosystems.

Tran, T.T., Pham, N.H., Pham, Q.B., Pham, T.L., Ngo, X.Q., Nguyen, D.L., Nguyen, P.N., Veettil, B.K., 2022. Performances of different machine learning algorithms for predicting saltwater intrusion in the Vietnamese Mekong Delta using limited input data: A Study from Ham Luong river. Water Resour. 49 (3), 391–401. http://dx.doi.org/10.1134/S0097807822030198.

Tran, D.A., Tsujimura, M., Ha, N.T., Nguyen, V.T., Binh, D.V., Dang, T.D., Doan, Q.-V., Bui, D.T., Anh Ngoc, T., Phu, L.V., Thuc, P.T.B., Pham, T.D., 2021. Evaluating the predictive power of different machine learning algorithms for groundwater salinity prediction of multi-layer coastal aquifers in the Mekong Delta, Vietnam. Ecol. Indic. 127, 107790. http://dx.doi.org/10.1016/j.ecolind.2021.107790, URL: https://www.sciencedirect.com/science/article/pii/S1470160X21004556.

Valle-Levinson, A., 2010. Definition and classification of estuaries. In: Contemporary Issues in Estuarine Physics. Cambridge University Press, pp. 1–11. http://dx.doi.org/10.1017/CBO9780511676567.002.

Verri, G., De Lorenzis, A., Santos da Costa, V., Sorolla, A., A., L., Ribot, M., et al., 2024a. Salt-wedge estuary's response to rising sea level, reduced discharge and nature based solutions. Front. Clim. (submitted for publication).

Verri, G., Furnari, L., Gunduz, M., Senatore, A., Santos da Costa, V., De Lorenzis, A., Fedele, G., Manco, I., Mentaschi, L., Clementi, E., Coppini, G., Mercogliano, P., Mendicino, G., Pinardi, N., 2024b. Climate projections of the Adriatic sea: Role of river release. Front. Clim. 6, http://dx.doi.org/10.3389/fclim.2024.1368413, URL: https://www.frontiersin.org/articles/10.3389/fclim.2024.1368413.

Verri, G., Mahmoudi Kurdistani, S., Coppini, G., Valentini, A., 2021. Recent advances of a box model to represent the estuarine dynamics: Time- Variable Estuary length and eddy diffusivity. J. Adv. Modelling Earth Syst. 13 (4), http://dx.doi.org/10.1029/2020MS002276, arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020MS002276, URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020MS002276, e2020MS002276 2020MS002276.

Verri, G., Pinardi, N., Bryan, F., heng Tseng, Y., Coppini, G., Clementi, E., 2020. A box model to represent estuarine dynamics in mesoscale resolution ocean models. Ocean Model. 148, 101587. http://dx.doi.org/10.1016/j.ocemod.2020.101587, URL: https://www.sciencedirect.com/science/article/pii/S1463500319300642.

Witten, I.H., Frank, E., Hall, M.A., 2011. Credibility: Evaluating what's been learned. In: Witten, I.H., Frank, E., Hall, M.A. (Eds.), Data Mining: Practical Machine Learning Tools and Techniques (Third Edition), third ed. In: The Morgan Kaufmann Series in Data Management Systems, Morgan Kaufmann, Boston, pp. 147–187. http://dx.doi.org/10.1016/B978-0-12-374856-0.00005-5, URL: https://www.sciencedirect.com/science/article/pii/B9780123748560000055.

Wong, P.P., Losada, I., Gattuso, J.-P., Hinkel, J., Khattabi, A., Mcinnes, K., Saito, Y., Sallenger, A., 2014. Coastal Systems and Low-Lying Areas. pp. 361–409. http://dx.doi.org/10.1017/CBO9781107415379.010.

Ye, R., Kong, J., Shen, C., Zhang, J., Zhang, W., 2020. An alternative statistical model for predicting salinity variations in estuaries. Sustainability 12 (24), http://dx.doi.org/10.3390/su122410677, URL: https://www.mdpi.com/2071-1050/12/24/10677.