https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

Reconstructing Sea Level Variability at the Ieodo Ocean Research Station (1993–2023) Using Artificial Intelligence, Machine Learning, and Reanalysis Integration

MyeongHee Han and Hak Soo Lim*

Marine Natural Disaster Research Department, Korea Institute of Ocean Science and Technology, Busan, Republic of Korea

*Corresponding author: Hak Soo Lim (hslim@kiost.ac.kr)

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

Abstract

This study presents a comprehensive approach for reconstructing a high-quality, continuous monthly sea level time series at the Ieodo Ocean Research Station (IORS) from 1993 to 2023 using advanced artificial intelligence (AI) and machine learning (ML) models. After applying quality control to the in-situ KIOST data, including inverse barometric effect correction, 3σ filtering, and a 75% data coverage threshold, we validated trends using nearby PSMSL tide gauges and four ocean reanalysis datasets (CMEMS, GLORYS, ORAS5, HYCOM). The trend analysis showed a higher rate of sea level rise from in-situ data (4.94 mm/yr, Oct 2003–Dec 2023) compared to satellite and model-based estimates (e.g., CMEMS: 3.53 mm/yr, Jan 1993–Dec 2023), suggesting localized sea level rise in the East China Sea. Initial gap-filling used statistical models such as harmonic regression and regression-based climatology. A blended approach combining climatology and trend components achieved the best accuracy (RMSE ~0.056 m, R^2 = 0.688). We then implemented various AI/ML models through an Iterative Imputer framework. Ensemble models (e.g., XGBoost) performed perfectly after 2003 but did not generalize well before 2004. Deep learning models like LSTM and GRU effectively captured seasonal and nonlinear patterns post-2003, with LSTM achieving RMSE = 0.023 m and R^2 = 0.95. Time series models Prophet and SARIMA-SIN successfully reconstructed the full time series, with SARIMA-SIN estimating the highest trend (5.61 mm/yr). Multiple linear regression using reanalysis data served as a baseline, but AI/ML models outperformed it in both accuracy and generalization. This study provides a reproducible, interpretable, and physically consistent framework for reconstructing sea level variability in semi-enclosed coastal seas.

Keywords: Sea Level Variability, Ieodo Ocean Research Station, AI-based Imputation,

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

Timeseries Reconstruction, East China Sea

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

1. Introduction

Long-term sea level observations are critical for monitoring regional oceanographic and climatic changes, particularly in coastal and marginal seas where variability is amplified by topographic and atmospheric forcing (Hamlington et al., 2020; Cazenave and Moreira, 2022). The Ieodo Ocean Research Station, located at the intersection of the East China Sea (ECS) and the Southern East Sea (Sea of Japan), plays a pivotal role in observing regional sea level variability and marine environmental conditions (Han, 2020; Byun et al., 2021). As a strategically important in-situ platform, its various interval sea level records provide valuable information for understanding the dynamics of the Kuroshio Current system, East Asian monsoon variability, and climate change-driven sea level rise (Ha et al., 2019; Xu et al., 2015; Chang and Oey, 2011).

However, long-term in-situ observations are often interrupted by equipment failure, maintenance issues, or extreme weather conditions such as typhoons (Adebisi et al., 2021). These disruptions lead to temporal gaps in the observational records, which hinder the detection of trends, reconstruction of seasonal cycles, and validation of satellite and model-based products (Beguería et al., 2019). In regions like the ECS, where strong seasonal and interannual signals are present, accurate and realistic imputation of missing data is essential for scientific and operational applications (Lin et al., 2020; Han, 2020).

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

Various methods have been developed to address missing data in oceanographic and climate time series (Kolukula and Murty, 2025; Lee et al., 2022). Traditional approaches include linear interpolation, monthly climatological averages, and harmonic regression models (Schlegel et al., 2019; Risien et al., 2022; Okkaoğlu et al., 2020; Arguez and Applequist, 2013). More recently, advanced statistical and machine learning techniques have been proposed for gap-filling, including Gaussian process regression, Kalman filtering, and neural networks (Vance et al., 2022; Wenzel and Schröter, 2010; Wang, 2023). While these methods offer improved flexibility and accuracy, they often require dense observations or training data, which may not be feasible in long-term sparse records (Lee et al., 2022; Sarafanov et al., 2022; Park et al., 2023). Interpretable and statistically robust methods remain essential for operational and historical datasets such as IORS (Han and Lim, 2024; Han, 2020).

This study focuses on imputing, filling, and predicting gaps in the monthly sea level data at the IORS over 1993-2023. We evaluate and compare three regression-based approaches: (1) harmonic regression with annual and semiannual cycles plus a linear trend (Okkaoğlu et al., 2020), (2) a regression-based monthly climatology model with a trend using calendar year month dummy variables (Hyndman and Athanasopoulos, 2018), and (3) a pure climatology-plus-trend approach based on aggregated monthly means and linear fitting (Brunetti et al., 2014). Furthermore, we propose a realistic blending method that optimally combines harmonic

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

and climatological components to minimize reconstruction errors.

Also, we implemented an ensemble of statistical and machine learning (ML) models to reconstruct missing monthly values at the IORS (Tong and Li, 2025; Bochow et al., 2025). These models were selected to span a broad range of algorithmic families, including ensemble tree-based methods (e.g., XGBoost, Random Forest, LightGBM, AdaBoost) (Niazkar et al., 2024; Wu et al., 2024; Gan et al., 2021; Xiao et al., 2019), regularized linear models (Lasso, Ridge) (Pan et al., 2025), proximity-based models (K-Nearest Neighbors) (Latif et al., 2024), and neural networks (LSTM, GRU) (Sun et al., 2020; Tumse and Alcansoy, 2025). Each model was trained to impute missing values using only observed sea level data from the original dataset, applying the IterativeImputer framework with consistent hyperparameters for comparability (Ramirez et al., 2023). Complementing these were two timeseries specific models: Facebook's Prophet (Elneel et al., 2024), which decomposes time series into seasonal and trend components, and SARIMA (Sun et al., 2020), which captures both non-seasonal and seasonal autocorrelation structures. All models were evaluated using root mean square error (RMSE) and coefficient of determination (R^2), calculated against observed (non-missing) values (Boursalie et al., 2022; Siddig et al., 2021).

To contextualize the reconstructed series, we further compared the IORS observations with monthly sea level data from four global ocean reanalysis products—Copernicus Marine

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

Environment Monitoring Service (CMEMS), Global Ocean Physics Reanalysis System (GLORYS), Ocean Reanalysis System 5 (ORAS5), and Hybrid Coordinate Ocean Model (HYCOM)—over 1993–2023 (Han et al., 2024; Long et al., 2023; Jin et al., 2023b; Cummings and Smedstad, 2014). Linear trends were estimated and visualized to assess the consistency and fidelity of in-situ observations relative to reanalysis-based records.

Finally, we extended the imputation framework to a multivariate setting by incorporating these reanalysis datasets as auxiliary predictors, enabling more robust and physically informed gap-filling. By leveraging both statistical and AI/ML-based techniques, this study provides a transparent, reproducible framework for reconstructing realistic monthly sea level time series in data-sparse coastal environments.

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

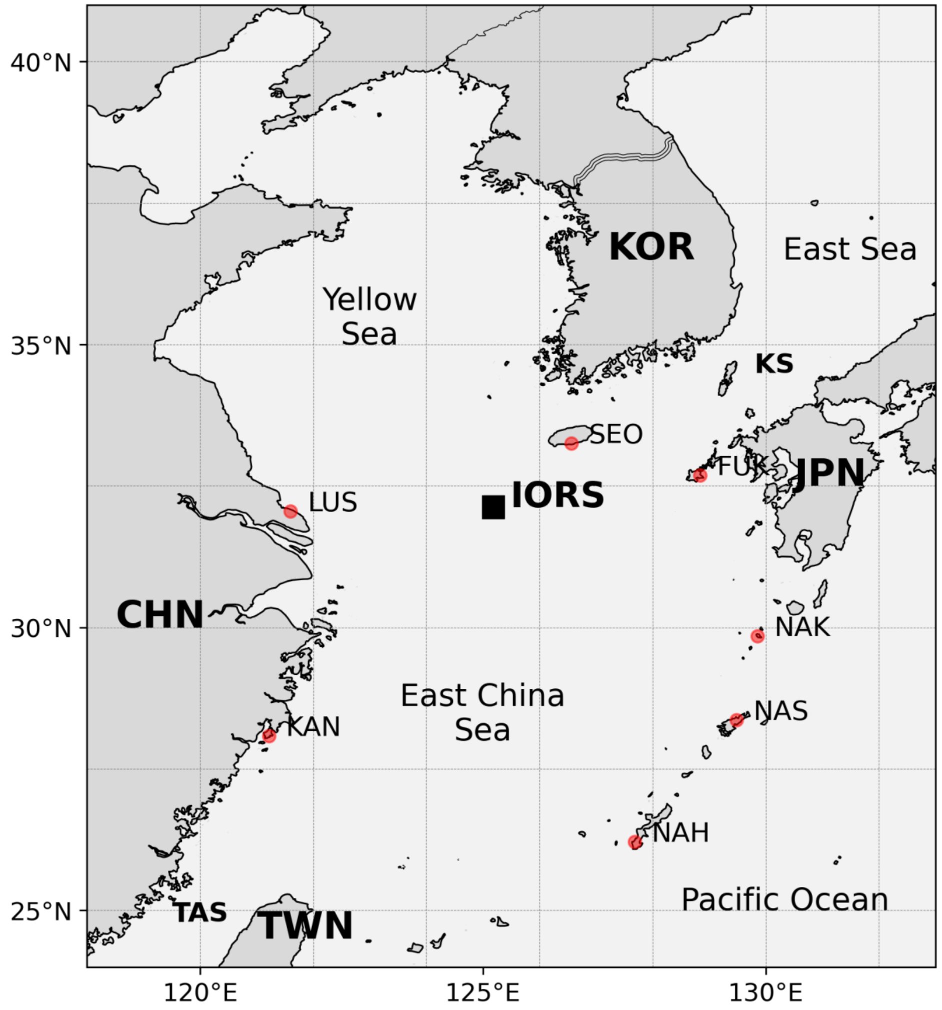

Figure 1. Map showing the locations of sea level observation stations used in this study, including the Ieodo Ocean Research Station (IORS, black square) and seven tide gauge stations (red circles) from the PSMSL and regional networks. The map spans the East China Sea, Yellow Sea, East Sea (Sea of Japan), and the western Pacific Ocean. Major geographic regions and oceanic features labelled, including Korea (KOR), China (CHN), Japan (JPN), and Taiwan (TWN); KS and TAS are Korea and Taiwan Straits, respectively.

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

2. Data and Methods

2.1. Data and Preprocessing

This study utilizes sea level observations from the IORS, collected between October 2003 and December 2023. The data, provided by the Korea Institute of Ocean Science and Technology (KIOST), include measurements at 1-minute, 10-minute, and hourly intervals (Kiost, 2025). Data gaps occur intermittently due to sensor malfunctions, scheduled maintenance, and severe weather events.

To ensure physical consistency with satellite altimetry, numerical models, and reanalysis products, all sea level values were corrected for the inverse barometric effect (IBE) in Eq. (1) (Han et al., 2024). The correction was computed using collocated atmospheric pressure measurements:

$$\eta_{\text{corrected}} = \eta_{\text{raw}} + \frac{(P-P_0)}{\rho g} \tag{1}$$

Where η_{raw} is the observed and uncorrected sea level (m), P is local atmospheric pressure (Pa), P_0 is the standard atmospheric pressure (101300 Pa), ρ is the assumed seawater density (1025 kg/m³), and g is gravitational acceleration (9.81 m/s²).

A multi-tiered quality control protocol was implemented to construct reliable timeseries for subsequent analyses in Figures 2, 3 and 4:

- Quality Flag Filtering: Only values flagged as "good" based on instrument diagnostics

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

and metadata were retained (Gouldman et al., 2017) (Figure 2).

- 3σ statistical Filtering: Outliers exceeding ±3 standard deviations from the local mean were excluded (Oelsmann et al., 2021) (Figure 2).
- 75% Threshold Rule: Daily and monthly averages were retained only if at least 75% of the corresponding observations passed the above filters (Epa, 2000) (Figures 3 and 4).

The resulting dataset comprises clean daily and monthly timeseries, exhibiting a clear seasonal cycle, interannual variability, and several data gaps that require reconstruction. These timeseries served as the foundation for further statistical modeling and machine learning-based gap-filling.

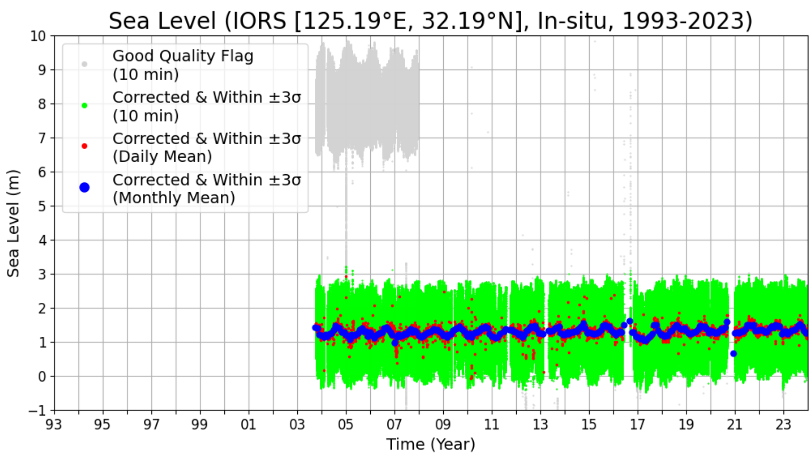

Figure 2. Time series of inverse barometric effect (IBE)-corrected in-situ sea level observations at the IORS from 1993 to 2023. Gray dots represent 10-minute interval data

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

flagged as good quality. Green dots indicate corrected 10-minute data within ±3σ. Red dots denote daily means within ±3σ, and blue dots show monthly means within ±3σ. Data gaps reflect periods of sensor malfunction or data loss.

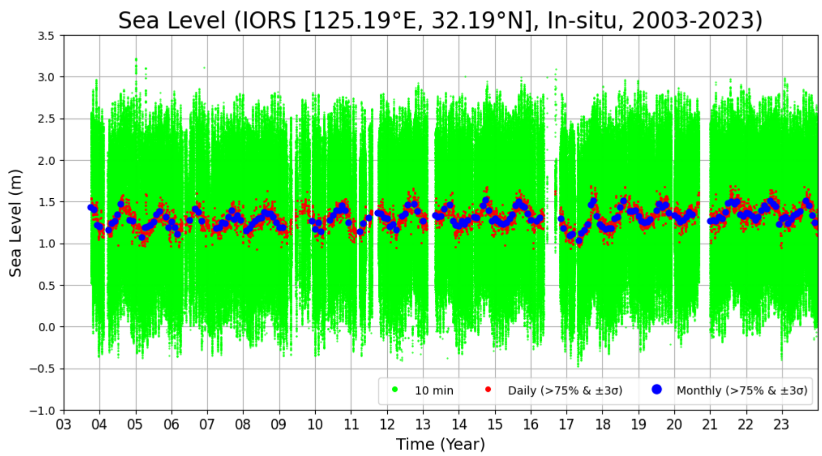

Figure 3. Sea level time series at the IORS from 2003 to 2023, based on in-situ observations corrected for the IBE. Green dots represent 10-minute sea level data. Red dots show daily mean values calculated only when at least 75% of 10-minute observations per day are available and within ±3 standard deviations. Blue dots denote monthly mean sea level, computed from daily means that meet the same 75% availability and ±3 criteria.

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

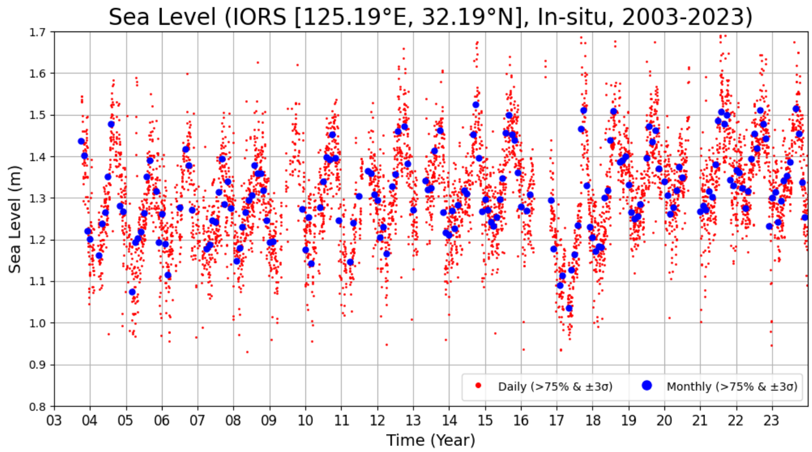

Figure 4. Filtered daily and monthly sea level time series at the IORS from 2003 to 2023. Red dots represent daily mean values calculated when at least 75% of 10-minute in-situ observations are available and within ±3 standard deviations. Blue dots indicate monthly mean values derived from daily means that also meet the 75% data availability and ±3σ filtering criteria.

2.2. Sea Level Reconstruction Using Regression Techniques

We applied three regression-based approaches to reconstruct missing values and estimate long-term sea level trends:

2.2.1. Harmonic Regression with Trend (HR)

This method models the sea level time series as a combination of a linear trend and periodic

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

seasonal components. Specifically, it includes annual and semiannual sine and cosine terms to capture seasonal variability in Eq. (2) (Young et al., 1999):

$$y_t = \beta_0 + \beta_1 t + \beta_2 \sin(\omega t) + \beta_3 \cos(\omega t) + \beta_4 \sin(2\omega t) + \beta_5 \cos(2\omega t) + \epsilon_t \tag{2}$$

where $\omega = 2\pi/12$, and t is time in months since January 1993. It captures both long-term and seasonal variability and is especially suitable for periodic signals.

2.2.2. Regression-Based Monthly Climatology with Trend (RC)

This method models monthly sea level as a function of a linear time trend and monthly categorical effects. Multiple linear regression uses dummy variables D_m (0 or 1) for calendar months (February through December), with January as the reference month. This allows the estimation of monthly climatology and trends simultaneously in Eq. (3). The model is expressed as (Hyndman and Athanasopoulos, 2018):

$$y_t = \beta_0 + \beta_1 t + \beta_2 D_2 + \beta_3 D_3 + \cdots + \beta_{12} D_{12} + \epsilon_t$$

$$= \beta_0 + \beta_1 t + \sum_{m=2}^{12} \beta_m D_m + \epsilon_t \tag{3}$$

Where y_t is the sea level at time t, β_0 is the intercept, β_1 captures the long-term linear trend, and β_m quantifies the deviation of each month m from January. This formulation allows for a flexible representation of the seasonal cycle without assuming a sinusoidal structure and can accommodate non-harmonic monthly anomalies while still estimating a long-term trend.

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

2.2.3. Pure Climatology with Trend (PC)

In this method, seasonal climatology is first computed by averaging the observed sea level for each calendar month across all years (Jin et al., 2023a). A separate linear trend is then fitted to the full-time series. The two components are summed to reconstruct the entire series in Eq. (4). While simple and interpretable, this method assumes stationarity in the seasonal cycle and may miss interannual variations:

$$y_t = C_{m(t)} + (\alpha_0 + \alpha_1 t) \tag{4}$$

where $C_{m(t)}$ is the climatological mean for the month of t. This assumes a stationary seasonal cycle.

2.2.4. Realistic Blending (RB)

To enhance the realism of the reconstructed sea level time series, we implemented a Realistic Blending (RB) approach that combines the outputs of the Harmonic Regression with Trend (HR) and the Regression-Based Monthly Climatology with Trend (RC) models in Eq. (5). The final imputed value at each time step is computed as a weighted average of the two models (Krinner et al., 2005):

$${y_t}^{RB} = \alpha {y_t}^{HB} + (1 - \alpha){y_t}^{RC} \tag{5}$$

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

where $\alpha \in [0,1]$ is an optimal blending weight selected to minimize the root mean square error (RMSE) between the blended estimate and the observed values. This method leverages the physical interpretability of harmonic modeling and the statistical flexibility of regression-based climatology, aiming to strike a balance between robustness and fidelity to observed variability.

2.3 Machine Learning and Statistical Models for Data Gap Imputation

To fill gaps in the monthly sea level record, we evaluate a range of machine learning and statistical models that have been effective for time series gap-filling in environmental datasets. Each model is described briefly below with supporting literature:

2.3.1. Extreme Gradient Boosting (XGBoost)

A powerful tree-based ensemble method that uses gradient boosting and inherently handles missing values by learning default split directions (Niazkar et al., 2024). XGBoost has demonstrated high accuracy in gap-filling environmental time series (e.g., used to impute significant gaps in aerosol optical depth data successfully) (Fan et al., 2020)

2.3.2. Gradient Boosting Regressor

A generic gradient boosting algorithm for regression that builds an ensemble of decision trees

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

iteratively to minimize prediction error. GBR has achieved low interpolation errors in climate data gap-filling, outperforming neural networks and multiple linear regression in one study of meteorological time series (Otchere et al., 2022; Körner et al., 2018).

2.3.3. Random Forest (RF)

An ensemble of decision trees trained on bootstrapped samples, aggregating their results. Random forests can capture nonlinear relationships and have robust performance for missing data imputation (Wu et al., 2024; Tang and Ishwaran, 2017). In comparative experiments, tree-based methods like RF rank among the top performers for reconstructing missing environmental time series data (Mahabbati et al., 2021).

2.3.4. AdaBoost

An adaptive boosting algorithm that sequentially combines many weak learners (often shallow trees) to improve prediction accuracy. The AdaBoost framework, while originally devised for classification, has a regression variant that has been applied to time series gap-filling, for example, by ensembling multiple neural-network base learners to impute missing traffic flow data, yielding improved accuracy and stability (Xiao et al., 2019; Shang et al., 2024).

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

2.3.5. LightGBM

LightGBM is a gradient boosting machine algorithm optimized for computational efficiency and scalability, employing histogram-based feature binning and leaf-wise tree growth. It has demonstrated high predictive accuracy in forecasting tasks and is widely applied in hydrological gap-filling studies(Wang et al., 2025). Its ability to efficiently process large datasets and natively handle missing values makes it highly suitable for gap-filling in extended climate time series, while also effectively capturing the nonlinear interactions between tidal dynamics and river discharge (Gan et al., 2021).

2.3.6. CatBoost (approximated via Gradient Boosting)

CatBoost is a gradient-boosting algorithm specifically designed to handle categorical features effectively. It introduces innovative techniques such as ordered boosting and efficient encoding of categorical variables, which help in reducing overfitting and improving model performance (Hancock and Khoshgoftaar, 2020). In this study, due to library constraints, we approximate CatBoost's functionality using scikit-learn's Gradient Boosting Regressor. While this approximation does not fully replicate CatBoost's specialized handling of categorical data, it allows us to implement a gradient boosting approach within our existing framework.

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

2.3.7. K-Nearest Neighbors

A non-parametric imputation approach that fills a missing value based on the values of its closest neighbors in feature space. KNN is simple yet effective for gap-filling; a study comparing time series imputation methods found that KNN achieved the highest reconstruction accuracy for missing data in several cases (Ahn et al., 2022). The method leverages spatial or temporal similarity, which can be particularly useful when neighboring station data or nearby time points correlate with the missing sea level values.

2.3.8. Lasso and Ridge Regression

These are regularized linear regression models that impose L1 (Lasso) and L2 (Ridge) penalties, respectively, to prevent overfitting. By shrinking coefficients, they provide stable estimates that can be used to predict and interpolate missing values from other correlated variables. Regularized regressions have been explored in gap-filling contexts to utilize correlations in environmental data while avoiding multicollinearity issues (Wijesekara and Liyanage, 2023). For instance, an elastic net (a combination of Lasso and Ridge) approach outperformed basic ARIMA-based methods in imputing large gaps of air quality data, highlighting the value of such penalized linear models.

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

2.3.9. Decision Tree

A single decision tree can be used as a regression model to estimate missing values by learning piecewise constant relationships in the data. Although decision trees alone are prone to higher variance, they have been applied to gap-filling and can capture nonlinear dependencies in sea level time series. In practice, tree-based algorithms (and their ensembles) have been found superior to many conventional interpolation techniques for climate data gaps(Zhu et al., 2023). Simpler decision trees may serve as interpretable benchmarks, while often forming the building blocks of more complex ensemble methods used for imputation.

2.3.10. Prophet

A forecasting model based on additive decompositions of trend, seasonality, and holidays (developed by Facebook). Prophet is robust to missing data and irregular timing – it does not require regularly spaced observations and can model gaps without explicit interpolation (Elneel et al., 2024). It has been applied in environmental time series forecasting (e.g. for air quality and water levels) and can be used to predict and back-fill missing monthly sea level values by leveraging seasonal patterns and trends in the data.

2.3.11. SARIMA-SIN (Seasonal ARIMA with Harmonic Regression Preprocessing)

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

To enhance the capacity of traditional SARIMA models in capturing periodic signals and long-term trends in oceanographic time series, we implemented a hybrid approach known as SARIMA-SIN. This method integrates harmonic regression and SARIMA modelling (Wang et al., 2019). In the first stage, a deterministic seasonal component is modeled using a linear combination of sine and cosine basis functions representing annual and semi-annual cycles. This harmonic regression accounts for the primary seasonality, after which residuals are computed by subtracting the fitted seasonal signal from the original time series. This preprocessing alleviates the need for SARIMA to fit both trend and periodicity simultaneously, thereby improving model stability and interpretability. Compared to conventional SARIMA models, SARIMA-SIN reduces overfitting and better preserves long-term variability, making it particularly suitable for sea level gap filling where both seasonal dynamics and secular trends are essential.

2.3.12. Neural Networks (LSTM and GRU)

Recurrent neural networks, particularly Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) architectures, are well-suited for sequential data and have been successfully applied to gap-filling in complex time series. These networks maintain internal memory states that allow them to learn long-term dependencies, which is advantageous for

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

inferring missing segments of a sea level record. Studies have shown that LSTM models can achieve more accurate gap imputations than traditional methods for environmental sensor data (Song et al., 2020), and specialized RNN variants (e.g., GRU-D) are designed to handle sequences with missing values by learning decay rates for missing inputs (Che et al., 2018). In practice, LSTM/GRU networks can leverage the temporal patterns in sea level data (and potentially exogenous inputs) to predict missing monthly values with high fidelity.

2.4. Evaluation Metrics and Implementation

To assess the performance of each imputation model, we computed two widely adopted evaluation metrics:

2.4.1. Root Mean Square Error (RMSE)

RMSE quantifies the average magnitude of reconstruction error by penalizing large deviations more strongly in Eq. (6) (Montgomery et al., 2021). It is defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \qquad (6)$$

where y_i are the observed values and $\hat{y}_i$ are the reconstructed values at non-missing time steps.

2.4.2. Coefficient of Determination (R²)

The coefficient of determination, R^2 , quantifies the proportion of variance in the observed data that is explained by the model in Eq. (7) (Montgomery et al., 2021). It is calculated as:

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

$$R^2=1-\frac{\sum_{i=1}^{N}(y_i-\widehat{y_i})^2}{\sum_{i=1}^{N}(y_i-\overline{y_i})^2} \tag{7}$$

Where y_i, $\widehat{y_i}$, $\overline{y_i}$, and N denote the observed values, the model-predicted values, the mean of observed values, and the number of valid (non-missing) observations.

An R^2 value of 1 indicates a perfect fit, meaning the model explains all the variance in the data. A value of 0 implies the model performs no better than simply predicting the mean, and negative values indicate that the model performs worse than the mean-based prediction. Thus, higher R^2 values reflect stronger agreement between the reconstructed and observed time series.

2.5. Comparison with Ocean Reanalysis Products

To contextualize in-situ observations, monthly sea level data from four global ocean reanalysis products—CMEMS, GLORYS, ORAS5, and HYCOM—were compared with the KIOST in-situ data at IORS over the period 1993–2023. All datasets were aggregated to monthly means and aligned in time for consistency. Linear trends (expressed in mm/year) were estimated using least squares regression and applied to each time series with NaN values masked. The trend slope was computed in meters per month and converted to millimeters per year by multiplying by 12,000. Each dataset was visualized on a same plot with annotated trend values to assess agreement and long-term consistency between in-situ and reanalysis records. The comparison

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

allowed us to assess model fidelity and characterize interannual-to-decadal sea level variability across observational and modeled sources.

2.6. AI/ML-Guided Multivariate Imputation Using Ocean Reanalysis Data

To improve the fidelity of gap-filling in the KIOST sea level time series, we implemented a multivariate imputation framework that integrates auxiliary predictors from four ocean reanalysis products: CMEMS, GLORYS, ORAS5, and HYCOM in Table 1. Specifically, we applied IterativeImputer from the scikit-learn package in Python, using multiple linear regression as the underlying estimator. In this configuration, the missing values in the KIOST in-situ observations were modeled as a function of the corresponding values from the reanalysis datasets, assuming linear dependence among variables.

Each imputation process was iteratively updated over 20 cycles to ensure numerical convergence and stability of the imputed results. In parallel, univariate imputation was also conducted using standalone time series models—Prophet, SARIMA-SIN, LSTM, and GRU—trained solely on the historical KIOST observations. These models capture temporal dynamics without external predictors.

Model outputs were evaluated based on the consistency of reconstructed trends, visual coherence in the temporal structure, and alignment with physical expectations. The integration

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

of regression-based multivariate imputation and AI/ML models forms an ensemble-based approach that leverages both data-driven learning and physically informed predictors for robust reconstruction of missing values.

Table 1. Overview of the datasets used in this study, including their temporal resolutions and time periods.

Data Source	Temporal Resolution	Time Period
IORS	1-min, 10-min, hourly	Oct 2003 - Dec 2023
CMEMS	Daily, Monthly	Jan 1993 - Dec 2023
GLORYS	Monthly	Jan 1993 - Dec 2023
ORAS5	Monthly	Jan 1993 - Dec 2023
HYCOM	Monthly	Jan 1994 - Dec 2023

3. Results

3.1. Comparison with Other Data

A linear regression performed on the 3σ-filtered, IBE-corrected monthly mean time series revealed a significant upward trend of approximately 4.94 mm/yr from October 2003 to December 2023 and 5.43 mm/yr from January 2004 to December 2023. This value is consistent with other in-situ coastal tide gauge trends in the ECS and reflects regional steric and mass-

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

driven contributions to sea level rise in Figure 5 and Table 2.

Table 2. Monthly data availability for five PSMSL tide gauge stations from 1993 to 2023 (372 months total). The table shows the available and missing data points, the total months considered, and the corresponding coverage percentage. These stations were used to compare sea level variability and trends with the IORS in the Northwest Pacific region.

	Available (Month)	Missing (Month)	Total (31 years)	Coverage (%)	Trend (mm/yr)
KANMEN	348	24	372	93.5	5.5
LUSI	291	81	372	78.2	6.2
SEOGWIPO	368	4	372	98.9	3.6
FUKUE	368	4	372	98.9	3.4
NAKANO SIMA	351	21	372	94.4	4.2
NASE III	352	20	372	94.6	2.6
NAHA	370	2	372	99.5	3.0

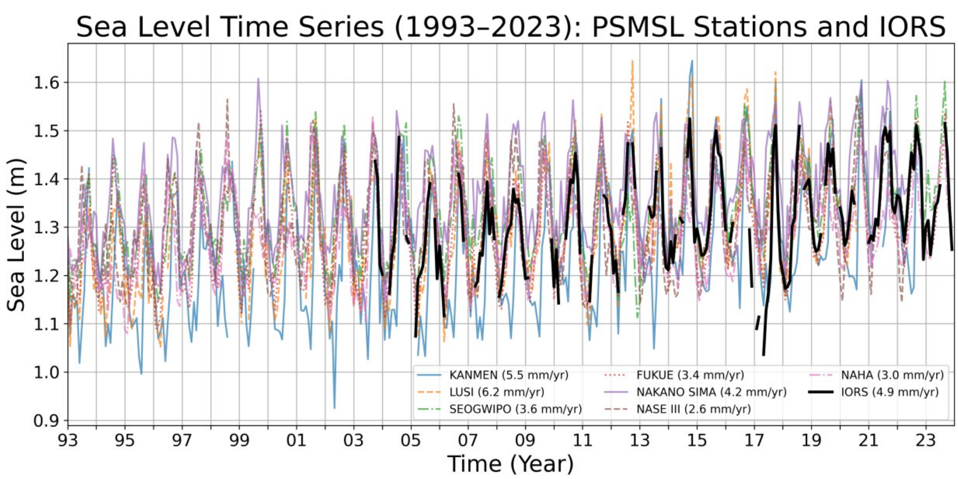

Figure 5. Monthly mean sea level time series from the IORS (bold black) and seven

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

PSMSL tide gauge stations (colored lines) from 1993 to 2023. PSMSL records are vertically adjusted by subtracting 5.8 m to align with IORS. Linear trends are indicated in the legend in millimeters per year (mm/yr), highlighting the long-term rise of sea level across the region.

3.2. Statistical Gap-Filling Models

To reconstruct missing values in the monthly sea level time series, five statistical gap-filling models were evaluated, as shown in Figure 6. These models include:

- Harmonic Regression with Trend (HR): A model combining annual and semiannual sinusoidal harmonics with a linear trend (red line).
- Regression-Based Monthly Climatology with Trend (RC): A linear regression using monthly dummy variables to represent the seasonal cycle with an added trend (green line).
- Pure Climatology with Trend (PC): A model combining the monthly climatological mean and an original linear trend (cyan dashed line).
- Realistic Blending (RB, $\alpha {y_t}^{HB} + (1-\alpha){y_t}^{RC}$ when $\alpha = 0.00$) : An optimally weighted composite of HR and RC based on root-mean-square error (solid black line).
- Forced Blending (FB, $\alpha = 0.50$): A non-optimized blend of HR and RC using a fixed mixing ratio (magenta dotted line).

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

Visual inspection reveals that the blended models, particularly the RB approach which is same to RC, closely align with the observed time series during periods of complete data. Among all candidates, the RB model ($\alpha = 0.00$) demonstrated the best performance, achieving the lowest RMSE (0.0559 m) and the highest coefficient of determination ($R^2 = 0.688$) (Table 3). The RC model achieved identical metrics, followed by HR with RMSE = 0.0574 m and $R^2 = 0.671$. The PC underperformed relative to the others, with the highest RMSE (0.0633 m) and lowest R^2 (0.600), highlighting the limitations of neglecting local interannual variability.

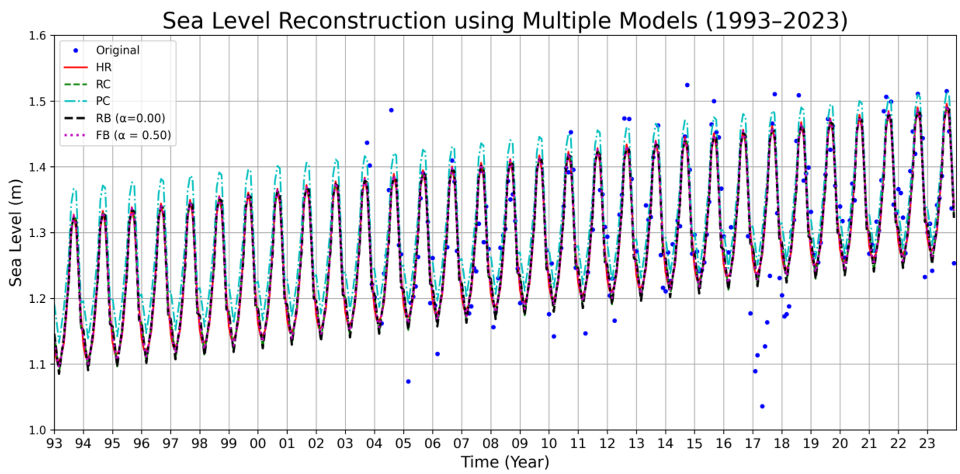

Figure 6. Monthly sea level reconstruction at the IORS from 1993 to 2023 using five statistical gap-filling models. Blue dots indicate the original in-situ observations. Colored lines represent different model reconstructions: Harmonic Regression with Trend (HR, red), Regression-Based Monthly Climatology with Trend (RC, green dashed), Pure Climatology with Trend (PC, cyan dash-dot), Realistic Blending (RB, α = 0.00, black dashed), and Forced Blending (FB, α = 0.50, magenta dotted). The comparison highlights

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

differences in seasonal and long-term signal reconstruction, with the RB model providing the most consistent alignment with observed values across the record.

Table 3. Performance evaluation of five statistical models for monthly sea level reconstruction at the IORS (1993–2023). The table compares root mean square error (RMSE) and coefficient of determination (R^2) between observed and reconstructed sea levels. The Realistic Blending model (α = 0.00) achieved the lowest RMSE and highest R^2, matching the Pure Climatology with Trend model, indicating that harmonic components did not improve reconstruction skill in this case.

Model	RMSE (m)	R^2
Harmonic Regression with Trend (HR)	0.0574	0.671
Regression-Based Monthly Climatology with Trend (RC)	0.0559	0.688
Pure Climatology with Trend (PC)	0.0633	0.600
Realistic Blending (RB, α=0.00)	0.0559	0.688
Forced Blending (FB, α=0.50)	0.0563	0.684

3.3. Artificial Intelligence and Machine Learning Imputation

To enhance the reconstruction accuracy of missing monthly sea level data at the IORS, a comprehensive set of artificial intelligence (AI) and machine learning (ML) models was

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

implemented. These models were trained on the 3σ-filtered monthly KIOST time series with synthetically introduced gaps to enable rigorous validation.

The methods included:

- Ensemble learning models: XGBoost, LightGBM, CatBoost, Gradient Boosting, Random Forest, AdaBoost
- Statistical models: SARIMA with sinusoidal regressors (SARIMA-SIN) and Prophet
- Regularized regressions: Ridge and Lasso
- Neural networks: Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU)
- Other models: K-Nearest Neighbors and Decision Trees

Figure 7 visualizes the full reconstructed time series from each model alongside the original KIOST observations. Prophet (red solid line) and SARIMA-SIN (dark green dashed) are visually emphasized due to their ability to reconstruct the full 1993–2023 time span, consistently capturing both seasonal cycles and long-term trends. In contrast, XGBoost and Random Forest models preserved the observed values and interpolated missing data, resulting in perfect fit metrics ($R^2 = 1$) but lacking predictive reconstruction before 2004 (Table 4).

As shown in Table 5, most machine learning models except Prophet, SARIMA-SIN, LSTM, and GRU achieved perfect scores (MAE = 0, RMSE = 0, $R^2 = 1$) because they directly reused observed values during imputation rather than making independent predictions, resulting in

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

artificially inflated performance metrics. Among the models that generated actual predictions, LSTM achieved the lowest MAE (0.018 m) and RMSE (0.023 m) with a high R^2 of 0.95, indicating its strength in capturing nonlinear temporal dynamics. GRU followed closely (MAE = 0.030 m, RMSE = 0.037 m, R^2 = 0.87), balancing interpretability and predictive capability. Prophet (MAE = 0.040 m, RMSE = 0.054 m, R^2 = 0.71) and SARIMA-SIN (MAE = 0.044 m, RMSE = 0.059 m, R^2 = 0.66) exhibited slightly higher error values but demonstrated strong consistency across the full time span (1993–2023), particularly in capturing seasonal and long-term trends—even before the availability of observed data in 2004. Thus, while tree-based ensemble models appear most accurate numerically, deep learning and time series models offer true predictive value, especially in extrapolating missing or historical sea level variations.

Table 4. Evaluation metrics for AI/ML-based sea level reconstruction models at the IORS from 1993 to 2023. Models were trained using the 3σ-filtered monthly KIOST data with artificially introduced gaps. Performance was assessed based on mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R^2), calculated using the valid observed values (N). Most models, including XGBoost, LightGBM, and CatBoost, produced near-perfect scores (MAE = 0, RMSE = 0, R^2 = 1) because they simply reused original values without generalizing to missing segments. In contrast, Prophet and SARIMA-SIN provided full-series predictions, capturing both seasonal and long-term variations. Deep learning models (LSTM, GRU) also exhibited strong predictive ability, though their outputs were limited to periods with sufficient training data. Notably, many

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

models performed poorly before 2004, where fewer observations were available to guide imputation.

Model	MAE (m)	RMSE (m)	R^2	N
XGBoost	0	0	1	183
GradientBoosting	0	0	1	183
CatBoost	0	0	1	183
KNeighbors	0	0	1	183
Lasso	0	0	1	183
RandomForest	0	0	1	183
LightGBM	0	0	1	183
AdaBoost	0	0	1	183
Ridge	0	0	1	183
DecisionTree	0	0	1	183
Prophet	0.040	0.054	0.71	183
SARIMA_SIN	0.044	0.059	0.66	183
LSTM	0.018	0.023	0.95	183
GRU	0.030	0.037	0.87	183

Tree-based models outperformed deep learning in both accuracy and training stability. However, LSTM showed strengths in capturing subtle nonlinear dependencies during seasonal transitions.

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

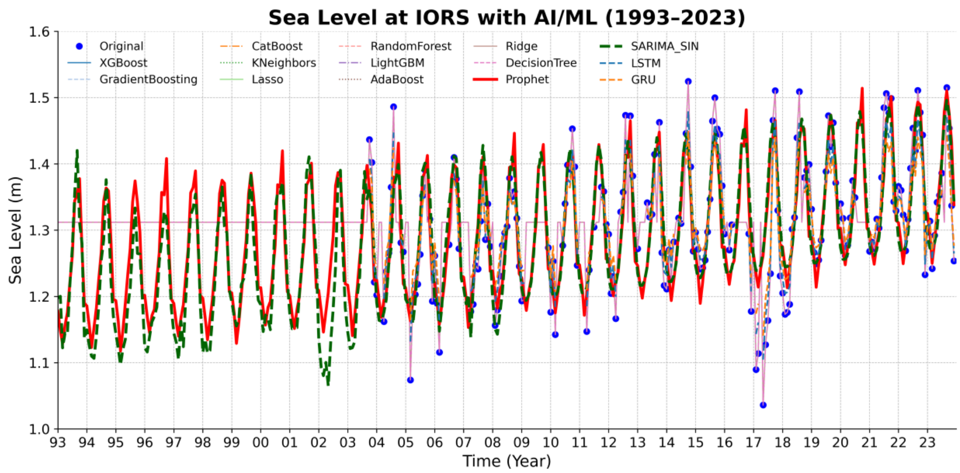

Figure 7. Monthly sea level time series at the IORS from 1993 to 2023 reconstructed using a suite of statistical and machine learning (AI/ML) models. Original observations (blue dots) contain data gaps that were filled using ensemble methods (XGBoost, LightGBM, CatBoost), regularized regressions (Ridge, Lasso), tree-based regressors (DecisionTree, RandomForest, AdaBoost), neural networks (LSTM, GRU), and time series models (Prophet, SARIMA-SIN). Prophet (red) and SARIMA-SIN (dark green dashed) are visually emphasized for their capacity to capture both seasonal cycles and long-term trends. The comparison illustrates model fidelity and divergence in reconstructing historical sea level variability at IORS.

3.4. Comparison with Reanalysis and Satellite Data

The monthly sea level time series at the Ieodo Ocean Research Station (IORS) was compared with satellite altimetry and four reanalysis products: CMEMS, GLORYS, ORAS5, and HYCOM (Figure 8). The in-situ tide gauge data from KIOST exhibited a linear sea level rise

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

of 4.94 mm yr^{-1} during the period from October 2003 to December 2023, with a slightly steeper trend of 5.43 mm yr^{-1} when computed from January 2004 onward (Table 5). These values exceed the trends derived from satellite and model-based datasets, suggesting possible biases due to local effects, datum inconsistencies, or limited assimilation in some models. Among the reanalyses, CMEMS showed the closest agreement with the observed seasonal and interannual variability, yielding a trend of 3.53 mm yr^{-1}. GLORYS and ORAS5 also captured similar variability, though with lower trends of 3.09 mm yr^{-1} and 2.27 mm yr^{-1}, respectively. In contrast, HYCOM presented almost no net sea level rise (−0.09 mm yr^{-1}), likely due to limitations in boundary conditions or lack of data assimilation in the East China Sea. Figure 8 clearly shows consistent seasonal patterns across all datasets, though the amplitude and interannual signals vary. The observed KIOST time series displayed greater interannual variability after 2003, corresponding to the availability of continuous in-situ measurements. Cross-correlation analysis further supports the consistency between CMEMS and other datasets (Table 6). The correlation between CMEMS and KIOST reached 0.90, with similar values for GLORYS (0.93) and ORAS5 (0.90). Correlation with HYCOM was slightly lower at 0.86, reflecting its weaker agreement with observed variability. Separate correlations using only post-2003 KIOST records confirmed similarly strong relationships. These results underscore the importance of in-situ validation for semi-enclosed marginal seas like the East China Sea, where regional

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

processes and bathymetry can significantly affect long-term sea level trends and their representation in models.

Table 5. Linear sea level trends (mm yr⁻¹) at the Ieodo Ocean Research Station (IORS) from multiple observational and reanalysis datasets. Trends are calculated using monthly data from tide gauge observations (KIOST), satellite altimetry (CMEMS), and ocean reanalysis/model outputs (GLORYS, ORAS5, HYCOM). The table includes trends for KIOST over two periods: October 2003–December 2023 and January 2004–December 2023, to assess sensitivity to the start date.

Dataset	Category	Period	Trend (mm yr^{-1})
In-situ	Tide Gauge (KIOST)	Oct 2003 - Dec 2023	4.94
In-situ	Tide Gauge (KIOST)	Jan 2004 - Dec 2023	5.43
CMEMS	Satellite Altimetry	Jan 1993 - Dec 2023	3.53
GLORYS	Ocean Reanalysis	Jan 1993 - Dec 2023	3.09
ORAS5	Ocean Reanalysis	Jan 1993 - Dec 2023	2.27
HYCOM	Ocean Model Output	Jan 1994 - Dec 2023	-0.09

Table 6. Cross-correlation coefficients between CMEMS satellite altimetry and other sea level datasets at the IORS (1993–2023). The comparison includes in-situ tide gauge data (KIOST), ocean reanalysis products (GLORYS, ORAS5), and a numerical ocean

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

model (HYCOM). All values are based on monthly mean anomalies.

Data type	Data type	Cross-correlation coefficient
CMEMS (1993–2023)	In-situ (2003–2023 and 2004–2023)	0.90
CMEMS (1993–2023)	GLORYS (1993–2023)	0.93
CMEMS (1993–2023)	ORAS5 (1993–2023)	0.90
CMEMS (1993–2023)	HYCOM (1994–2023)	0.86

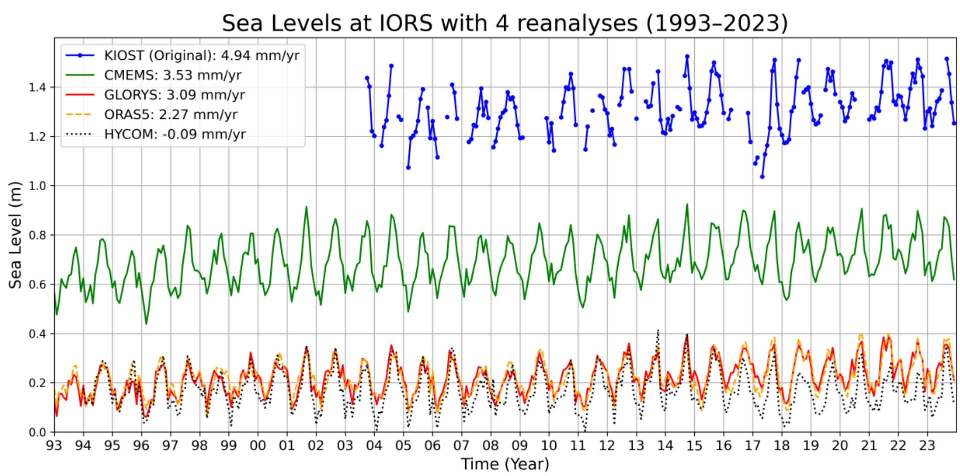

Figure 8. Comparison of monthly sea level time series at the IORS from 1993 to 2023, using in-situ observations (KIOST) and four ocean reanalysis products: CMEMS, GLORYS, ORAS5, and HYCOM. Each dataset is plotted with its respective linear trend estimated over the entire period. The observed sea level trend from KIOST (4.94 mm/yr)

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

is higher than those from CMEMS (3.53 mm/yr), GLORYS (3.09 mm/yr), ORAS5 (2.27 mm/yr), and HYCOM (−0.09 mm/yr), indicating significant discrepancies in modeled versus observed trends at this location. Differences may reflect varying assimilation schemes, vertical resolutions, or forcing mechanisms in each reanalysis system.

3.5. Comparison with AI/ML-Based Reconstruction

Before implementing machine learning-based imputation methods, we applied multiple linear regression using four reanalysis datasets (CMEMS, GLORYS, ORAS5, HYCOM) to evaluate their explanatory power for observed sea level variability at the IORS. This regression served as a preliminary benchmark to assess how well physically-based reanalysis products capture regional sea level dynamics. Following this, a suite of AI/ML models was employed to reconstruct the monthly sea level anomalies at the IORS from 1993 to 2023 (Fig. 9). Each model's time series was compared with the original KIOST observations, and corresponding linear trends (mm/year) were computed to assess long-term performance. The original KIOST dataset exhibits a 4.94 mm/yr trend, serving as a reference benchmark. Among ensemble learning methods, XGBoost (4.42 mm/yr), GradientBoosting (4.32 mm/yr), CatBoost (4.32 mm/yr), and RandomForest (4.26 mm/yr) closely reproduced the observed trend and temporal variability. The Prophet model yielded a slightly higher trend estimate (4.51 mm/yr) and consistently reconstructed the seasonal cycle and interannual variations. Its smooth,

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

Earth System Science Data Discussions
Open Access

interpretable structure made it effective for long-term monitoring, despite slightly larger residuals than ensemble models. The SARIMA-SIN approach, which integrates harmonic seasonal components with ARIMA modeling of residuals, produced the highest trend estimate (5.61 mm/yr). While effectively modeling periodic signals, its slight overestimation may stem from residual autocorrelation or the rigid seasonal structure embedded in the model. In contrast, linear regularized models, such as Lasso (3.74 mm/yr) and Ridge (2.98 mm/yr), underestimated the long-term trend and failed to capture higher-order seasonal and nonlinear dynamics. Neural network-based deep learning models showed intermediate performance: LSTM (3.90 mm/yr) and GRU (2.94 mm/yr) successfully captured high-frequency variability but smoothed long-term trends. Their underestimation may reflect challenges in extrapolating temporal dependencies across extended historical periods. Overall, ensemble models and Prophet offered the most balanced performance regarding accuracy, trend reconstruction, and robustness to missing data. SARIMA-SIN remains a promising alternative for seasonality-focused applications. At the same time, deep learning methods may benefit from additional architecture optimization and hyperparameter tuning to better preserve secular trends in semi-enclosed regions like the East China Sea.

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

Open Access Earth System Science Data Discussions

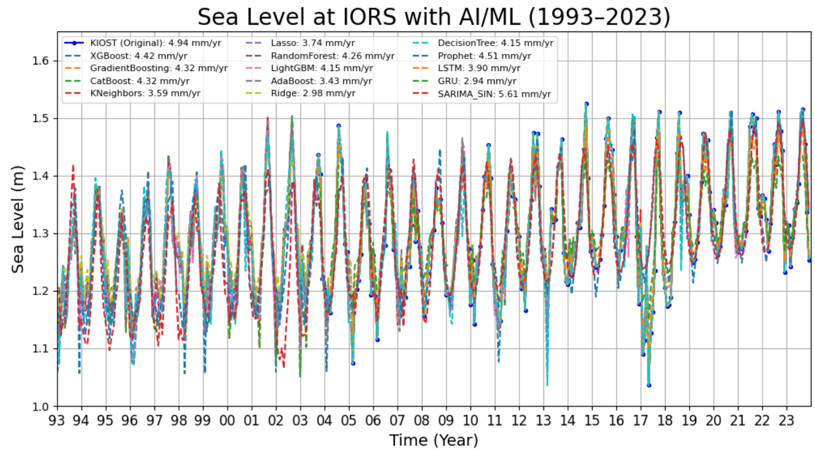

Figure 9. Monthly sea level time series at the IORS from 1993 to 2023 reconstructed using various AI/ML models. The original KIOST observations (blue) are compared with estimates from ensemble methods (XGBoost, GradientBoosting, CatBoost, RandomForest, AdaBoost), regularized regression (Ridge, Lasso), decision trees, nearest neighbors, Prophet, and deep learning (LSTM, GRU). A seasonal-harmonic SARIMA-SIN model is also included. Linear trends (in mm/yr) are indicated in the legend. The comparison highlights each model's capability to capture seasonal and interannual sea level variability.

4. Discussion

This study demonstrates a comprehensive approach to reconstructing monthly sea level variability at the Ieodo Ocean Research Station (IORS) over the period 1993–2023 using a suite of statistical, machine learning (ML), and artificial intelligence (AI) models. The

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

reconstruction framework was motivated by the need to address temporal discontinuities in long-term in-situ observations, which are often caused by instrument failure, extreme weather, or logistical constraints (Adebisi et al., 2021). These data gaps significantly hinder trend detection, seasonal analysis, and validation against satellite and reanalysis products in the East China Sea (ECS)—a region characterized by strong seasonal and interannual variability (Hamlington et al., 2020; Lin et al., 2020).

4.1. Data Sources and Preprocessing

A diverse set of observational and reanalysis datasets was employed, spanning from high-resolution in-situ KIOST observations to global oceanographic products such as CMEMS, GLORYS, ORAS5, and HYCOM (Table 1). Preprocessing involved inverse barometric effect (IBE) correction, 3σ filtering, and 75% coverage thresholds to ensure consistency and quality in derived monthly means. These preprocessing steps align with established best practices for climate-quality oceanographic datasets (Cazenave and Moreira, 2022). Figures 1–3 illustrate the geographical coverage and temporal continuity achieved after these steps.

4.2. Regional Validation with PSMSL Tide Gauges

To assess the validity of the reconstructed IORS series, comparisons were made with seven

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

nearby PSMSL tide gauge stations across the Northwest Pacific. The high correlation and consistent linear trends (2.6–6.2 mm/yr) observed in these stations (Table 2, Fig. 5) reinforce the regional representativeness of IORS data. This agreement supports the premise that well-maintained tide gauge records are invaluable for benchmarking reconstructions in semi-enclosed seas (Han, 2020).

4.3. Statistical Model Performance

We evaluated five classical statistical gap-filling techniques, including harmonic regression and regression climatology. Realistic Blending (RB) and Regression Climatology (RC) achieved the best results, with RMSE of 0.0559 m and $R^2 = 0.688$ (Table 3; Fig. 6). These findings are consistent with prior literature that emphasizes the effectiveness of incorporating seasonal structure and linear trends in climatic time series (Okkaoğlu et al., 2020; Hyndman and Athanasopoulos, 2018).

4.4. AI/ML Model Evaluation

Following statistical baseline evaluation, we implemented 14 AI/ML models using an Iterative Imputer framework. Ensemble learners such as XGBoost and LightGBM achieved perfect post-September 2003 metrics by reusing observed values, but lacked extrapolation capability

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

during earlier data gaps. In contrast, deep learning models like LSTM and GRU generated realistic, nonlinear predictions (LSTM: RMSE = 0.023 m, R^2 = 0.95), although they underperformed in estimating long-term trends and were limited to the post-September 2003 data range (Sun et al., 2020).

Time series-specific models Prophet and SARIMA-SIN produced full-period reconstructions with moderate accuracy (R^2 = 0.71 and 0.66, respectively), but demonstrated superior ability to capture seasonal and long-term variability, making them suitable for regions with sparse data (Elneel et al., 2024; Tumse and Alcansoy, 2025). Their trend estimates (4.51–5.61 mm/yr) approximate observational trends better than most ML alternatives (Fig. 7).

4.5. Cross-Dataset Trend and Correlation Analysis

Trend comparisons across observational, satellite, and model-based datasets (Table 5; Fig. 8) revealed that the IORS trends (4.94 mm/yr from October 2003 and 5.43 mm/yr from January 2004) are substantially higher than CMEMS (3.53 mm/yr) and GLORYS (3.09 mm/yr). HYCOM, notably, displayed a negative trend (−0.09 mm/yr), likely due to boundary condition limitations such as excluding glaciers and land ice sheets melting (Han et al., 2024; Jin et al., 2023b; Hycom, 2025). Cross-correlation analysis (Table 6) showed that CMEMS correlated

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

strongly with in-situ (r = 0.90), GLORYS (r = 0.93), and ORAS5 (r = 0.90), consistent with prior findings on their reliability in coastal sea level analysis (Cummings and Smedstad, 2014).

4.6. AI/ML Gap-Filling After Regression with Reanalysis

Prior to ML-based reconstruction, we conducted multiple linear regression using CMEMS, GLORYS, ORAS5, and HYCOM to gauge their ability to explain IORS variability (Fig. 9). The regression models underestimated the observed trend, underscoring their limitations in data-sparse, dynamically complex environments like the ECS. Subsequent AI/ML reconstructions—especially from Prophet, XGBoost, and SARIMA-SIN—outperformed these baseline models in both fidelity and interpretability. SARIMA-SIN provided the highest trend (5.61 mm/yr), while ensemble methods such as CatBoost and RandomForest approximated the KIOST trend closely (4.3–4.4 mm/yr). Deep learning models like GRU underpredicted trends (2.94 mm/yr), suggesting the need for further hyperparameter tuning or auxiliary input inclusion.

In summary, this study responds to the challenges outlined in the introduction—such as sparse long-term observational data and limitations of traditional gap-filling—by offering a transparent, interpretable, and data-driven framework for sea level reconstruction. Through the

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

integration of AI/ML techniques and reanalysis validation, we provide new insights into local sea level rise trends and offer a practical methodology for monitoring climate change effects in marginal seas.

5. Conclusion

This study developed and validated a robust methodology for reconstructing the monthly sea level record at the Ieodo Ocean Research Station (IORS) from 1993 to 2023 using a combination of statistical techniques, machine learning models, and auxiliary reanalysis datasets. Preprocessing steps, including inverse barometric effect (IBE) correction, 3s filtering, and data coverage thresholds, allowed for the construction of reliable time series from high-resolution in-situ observations. Comparison with seven nearby PSMSL tide gauges confirmed the regional validity of the observed trends. The IORS sea level trends (4.94 mm/yr for 2003–2023 and 5.43 mm/yr for 2004–2023) were significantly higher than those from satellite and reanalysis datasets, highlighting localized sea level rise. High correlations between CMEMS and other reanalysis products ($r = 0.90$) affirm their utility for sea level reconstruction, although they underestimated trends compared to in-situ observations. Statistical models such as regression climatology and realistic blending provided accurate reconstructions, achieving the best performance in RMSE and R^2 metrics. Among AI/ML models, ensemble learners (e.g., XGBoost, RandomForest) achieved perfect reconstruction metrics after September 2003 but failed to predict values in earlier periods. Deep learning models, particularly LSTM (RMSE = 0.023 m, $R^2 = 0.95$), effectively modeled nonlinear and seasonal variability, though their ability was restricted to after September 2003, and their trend estimation remained underestimated

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

compared to statistical benchmarks. Time series models, Prophet and SARIMA-SIN, demonstrated the strongest capability for full-period imputation, delivering consistent seasonal cycles and long-term trend estimates, with SARIMA-SIN reaching 5.61 mm/yr. Before applying ML models, multiple linear regression using four reanalysis datasets was conducted as a baseline. The AI/ML reconstructions outperformed reanalysis-based approaches in both accuracy and consistency with the observed KIOST trend. Overall, this study provides a scientifically grounded and computationally robust workflow for gap-filling, trend detection, and sea level variability analysis. The framework offers a valuable tool for long-term sea level monitoring, climate diagnostics, and policy planning in marginal seas and other observationally limited regions.

6. Data availability

The original sea level dataset at the IORS is available at https://kors.kiost.ac.kr/eng/accessData and https://doi.org/10.17882/97666 (Kiost, 2025; Jeong et al., 2023). For updates to the datasets, please contact the data provider by writing to ycmin@kiost.ac.kr.

7. Author contribution

MH writing – original draft, visualization, data collection, investigation. HL writing – review and editing, supervision.

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

8. Competing interests

The contact author has declared that none of the authors has any competing interests.

9. Acknowledgments

Oceanic and atmospheric data at the IORS (http://www.khoa.go.kr/oceangrid/gis/category/reference/distribution.do#none) were used in this study. This research was funded by the Ministry of Trade, Industry, and Energy (MOTIE) of Korea under the "Regional Innovation Cluster Development Program (PN92300, P0025418)", supervised by the Korea Institute for Advancement of Technology (KIAT). It was also supported by the Korea Institute of Ocean Science and Technology (PEA0231). Additionally, this study was supported by the project "Sustainable Research and Development of Dokdo (PG54141)" under the Ministry of Oceans and Fisheries, Korea.

8. References

Adebisi, N., Balogun, A. L., Min, T. H., and Tella, A.: Advances in estimating sea level rise: A review of tide gauge, satellite altimetry and spatial data science approaches, Ocean & Coastal Management, 208, 105632, https://doi.org/10.1016/j.ocecoaman.2021.105632, 2021.

Ahn, H., Sun, K., and Kim, K. P.: Comparison of missing data imputation methods in time series forecasting, Computers, Materials & Continua, 70, 767-779, https://doi.org/10.32604/cmc.2022.019369, 2022.

Arguez, A. and Applequist, S.: A harmonic approach for calculating daily temperature normals constrained by homogenized monthly temperature normals, Journal of Atmospheric and Oceanic Technology, 30, 1259-1265, https://doi.org/10.1175/JTECH-D-12-00195.1, 2013.

Beguería, S., Tomas-Burguera, M., Serrano-Notivoli, R., Peña-Angulo, D., Vicente-Serrano, S.

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

M., and González-Hidalgo, J.-C.: Gap filling of monthly temperature data and its effect on climatic variability and trends, Journal of Climate, 32, 7797-7821, https://doi.org/10.1175/JCLI-D-19-0244.1, 2019.

Bochow, N., Poltronieri, A., Rypdal, M., and Boers, N.: Reconstructing historical climate fields with deep learning, Science Advances, 11, eadp0558, https://doi.org/10.1126/sciadv.adp0558, 2025.

Boursalie, O., Samavi, R., and Doyle, T. E.: Evaluation methodology for deep learning imputation models, Experimental Biology and Medicine, 247, 1972-1987, https://doi.org/10.1177/15353702221121602, 2022.

Brunetti, M., Maugeri, M., Nanni, T., Simolo, C., and Spinoni, J.: High-resolution temperature climatology for Italy: interpolation method intercomparison, International Journal of Climatology, 34, 1278-1296, https://doi.org/10.1002/joc.3764, 2014.

Byun, D. S., Jeong, J. Y., Kim, D. J., Hong, S., Lee, K. T., and Lee, K.: Ocean and atmospheric observations at the remote Ieodo Ocean Research Station in the northern East China Sea, Frontiers in Marine Science, 8, 618500, https://doi.org/10.3389/fmars.2021.618500, 2021.

Cazenave, A. and Moreira, L.: Contemporary sea-level changes from global to local scales: a review, Proceedings of the Royal Society A, 478, 20220049, https://doi.org/10.1098/rspa.2022.0049, 2022.

Chang, Y. L. and Oey, L. Y.: Interannual and seasonal variations of Kuroshio transport east of Taiwan inferred from 29 years of tide-gauge data, Geophysical Research Letters, 38, https://doi.org/10.1029/2011GL047062, 2011.

Che, Z., Purushotham, S., Cho, K., Sontag, D., and Liu, Y.: Recurrent neural networks for multivariate time series with missing values, Scientific reports, 8, 6085, https://doi.org/10.1038/s41598-018-24271-9, 2018.

Cummings, J. A. and Smedstad, O. M.: Ocean data impacts in global HYCOM, Journal of Atmospheric and Oceanic Technology, 31, 1771-1791, https://doi.org/10.1175/JTECH-D-14-00011.1, 2014.

Elneel, L., Zitouni, M. S., Mukhtar, H., and Al-Ahmad, H.: Examining sea levels forecasting using autoregressive and prophet models, Scientific Reports, 14, 14337, https://doi.org/10.1038/s41598-024-65184-0, 2024.

Guidance for Data Quality Assessment: Practical Methods for Data Analysis: EPA QA/G9: QA00 Update: July 2000: https://www.epa.gov/sites/default/files/2015-06/documents/g9-final.pdf, last access: 05 May 2025.

Fan, Z., Zhan, Q., Yang, C., Liu, H., and Bilal, M.: Estimating PM2. 5 concentrations using spatially local Xgboost based on full-covered SARA AOD at the urban scale, Remote Sensing, 12, 3368, https://doi.org/10.3390/rs12203368, 2020.

Gan, M., Pan, S., Chen, Y., Cheng, C., Pan, H., and Zhu, X.: Application of the machine

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

learning lightgbm model to the prediction of the water levels of the lower columbia river, Journal of Marine Science and Engineering, 9, 496, https://doi.org/10.3390/jmse9050496, 2021.

Manual for real-time oceanographic data quality control flags: https://cdn.ioos.noaa.gov/media/2017/12/QARTOD-Data-Flags-Manual_Final_version1.1.pdf, last access: 05 May 2025.

Ha, K.-J., Nam, S., Jeong, J.-Y., Moon, I.-J., Lee, M., Yun, J., Jang, C. J., Kim, Y. S., Byun, D.-S., and Heo, K.-Y.: Observations utilizing Korea Ocean Research Stations and their applications for process studies, Bulletin of the American Meteorological Society, 100, 2061-2075, https://doi.org/10.1175/BAMS-D-18-0305.1, 2019.

Hamlington, B. D., Gardner, A. S., Ivins, E., Lenaerts, J. T., Reager, J. T., Trossman, D. S., Zaron, E. D., Adhikari, S., Arendt, A., Aschwanden, A., and Beckley, B. D.: Understanding of contemporary regional sea-level change and the implications for the future, Reviews of Geophysics, 58, e2019RG000672, https://doi.org/10.1029/2019RG000672, 2020.

Han, M.: Status of Observation Data at Ieodo Ocean Research Station for Sea Level Study, Journal of the Korean earth science society, 41, 323-343, https://doi.org/10.5467/JKESS.2020.41.4.323, 2020.

Han, M. and Lim, H.-S.: Enhancing Data Accessibility and Usability: Insights from Observation Data at the Ieodo Ocean Research Station (2005-2023), Journal of Coastal Disaster Prevention, 11, 59-70, https://doi.org/10.20481/kscdp.2024.11.3.59, 2024.

Han, M., Nam, S., and Lim, H.-S.: Regional Mean Sea Level Variability Due to Tropical Cyclones: Insights from August Typhoons, Journal of Marine Science and Engineering, 12, 1830, https://doi.org/10.3390/jmse12101830, 2024.

Hancock, J. T. and Khoshgoftaar, T. M.: CatBoost for big data: an interdisciplinary review, Journal of big data, 7, 94, https://doi.org/10.1186/s40537-020-00369-8, 2020.

HYCOM.org Forum: https://groups.google.com/a/hycom.org/g/forum/c/QK-nlXqfQdE/m/6gCq-Eh3CwAJ, last access: April 15, 2025.

Hyndman, R. J. and Athanasopoulos, G.: Forecasting: Principles and Practice, OTexts, 2018.

Jeong, K., Kwon, S., Oh, H., Ham, H., Park, H., Roh, J., Kim, Y., Lee, Y., Jeong, J., and Min, Y.: Ieodo Ocean Research Station (Ieodo ORS): delayed mode dataset, annual release (2022), SEANOE, https://doi.org/10.17882/97666, 2023.

Jin, C., Liu, H., and Lin, P.: Evaluation of the seasonal to decadal variability in dynamic sea level simulations from CMIP5 to CMIP6, Geoscience Letters, 10, 35, https://doi.org/10.1186/s40562-023-00291-w, 2023a.

Jin, Y., Chen, M., Yan, H., Wang, T., and Yang, J.: Sea level variation in the Arctic Ocean since 1979 based on ORAS5 data, Frontiers in Marine Science, 10, 1197456, https://doi.org/10.3389/fmars.2023.1197456, 2023b.

KIOST Data Access Portal: https://kors.kiost.ac.kr/eng/accessData, last access: April 15, 2025.

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

Kolukula, S. S. and Murty, P. L. N.: Enhancing observations data: A machine-learning approach to fill gaps in the moored buoy data, Results in Engineering, 26, 104708, https://doi.org/10.1016/j.rineng.2025.104708, 2025.

Körner, P., Kronenberg, R., Genzel, S., and Bernhofer, C.: Introducing Gradient Boosting as a universal gap filling tool for meteorological time series, Meteorologische Zeitschrift, 27, 369-376, https://doi.org/10.1127/metz/2018/0908, 2018.

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Global biogeochemical cycles, 19, https://doi.org/10.1029/2003GB002199, 2005.

Latif, S. D., Almubaidin, M. A., Shen, C. G., Sapitang, M., Birima, A. H., Ahmed, A. N., Sherif, M., and El-Shafie, A.: Improving sea level prediction in coastal areas using machine learning techniques, Ain Shams Engineering Journal, 15, 102916, https://doi.org/10.1016/j.asej.2024.102916, 2024.

Lee, E. J., Kim, K., and Park, J. H.: Reconstruction of long-term sea-level data gaps of tide gauge records using a neural network operator, Frontiers in Marine Science, 9, 1037697, https://doi.org/10.3389/fmars.2022.1037697, 2022.

Lin, T., He, Q., Zhan, W., and Zhan, H.: Persistent data gap in ocean color observations over the East China Sea in winter: causes and reconstructions, Remote Sensing Letters, 11, 667-676, https://doi.org/10.1080/2150704X.2020.1754491, 2020.

Long, X., Shin, S. I., and Newman, M.: Statistical downscaling of seasonal forecasts of sea level anomalies for US coasts, Geophysical Research Letters, 50, e2022GL100271, https://doi.org/10.1029/2022GL100271, 2023.

Mahabbati, A., Beringer, J., Leopold, M., McHugh, I., Cleverly, J., Isaac, P., and Izady, A.: A comparison of gap-filling algorithms for eddy covariance fluxes and their drivers, Geoscientific Instrumentation, Methods and Data Systems, 10, 123-140, https://doi.org/10.5194/gi-10-123-2021, 2021.

Montgomery, D. C., Peck, E. A., and Vining, G. G.: Introduction to linear regression analysis, John Wiley & Sons, 2021.

Niazkar, M., Menapace, A., Brentan, B., Piraei, R., Jimenez, D., Dhawan, P., and Righetti, M.: Applications of XGBoost in water resources engineering: A systematic literature review (Dec 2018–May 2023), Environmental Modelling & Software, 174, 105971, https://doi.org/10.1016/j.envsoft.2024.105971, 2024.

Oelsmann, J., Passaro, M., Dettmering, D., Schwatke, C., Sánchez, L., and Seitz, F.: The zone of influence: matching sea level variability from coastal altimetry and tide gauges for vertical land motion estimation, Ocean Science, 17, 35-57, https://doi.org/10.5194/os-17-35-2021, 2021.

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

Okkaoğlu, Y., Akdi, Y., and Ünlü, K. D.: Daily PM10, periodicity and harmonic regression model: The case of London, Atmospheric Environment, 238, 117755, https://doi.org/10.1016/j.atmosenv.2020.117755, 2020.

Otchere, D. A., Ganat, T. O. A., Ojero, J. O., Tackie-Otoo, B. N., and Taki, M. Y.: Application of gradient boosting regression model for the evaluation of feature selection techniques in improving reservoir characterisation predictions, Journal of Petroleum Science and Engineering, 208, 109244, https://doi.org/10.1016/j.petrol.2021.109244, 2022.

Pan, H., Sun, J., Gao, X., Teng, F., Xu, T., and Wei, Z.: Can we accurately extract ocean tides from satellite altimeter records with substantial missing values in shallow bays?, Estuarine, Coastal and Shelf Science, 109280, https://doi.org/10.1016/j.ecss.2025.109280, 2025.

Park, J., Müller, J., Arora, B., Faybishenko, B., Pastorello, G., Varadharajan, C., Sahu, R., and Agarwal, D.: Long-term missing value imputation for time series data using deep neural networks, Neural Computing and Applications, 35, 9071-9091, https://doi.org/10.1007/s00521-022-08165-6
2023.

Ramirez, S. G., Williams, G. P., Jones, N. L., Ames, D. P., and Radebaugh, J.: Improving groundwater imputation through iterative refinement using spatial and temporal correlations from in situ data with machine learning, Water, 15, 1236, https://doi.org/10.3390/w15061236, 2023.

Risien, C. M., Fewings, M. R., Fisher, J. L., Peterson, J. O., and Morgan, C. A.: Spatially gridded cross-shelf hydrographic sections and monthly climatologies from shipboard survey data collected along the Newport Hydrographic Line, 1997–2021, Data in Brief, 41, 107922, https://doi.org/10.1016/j.dib.2022.107922, 2022.

Sarafanov, M., Nikitin, N. O., and Kalyuzhnaya, A. V.: Automated data-driven approach for gap filling in the time series using evolutionary learning, 16th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2021), 633-642, https://doi.org/10.1007/978-3-030-87869-6_60,

Schlegel, R. W., Oliver, E. C., Hobday, A. J., and Smit, A. J.: Detecting marine heatwaves with sub-optimal data, Frontiers in Marine Science, 6, 737, https://doi.org/10.3389/fmars.2019.00737, 2019.

Shang, Q., Tang, Y., and Yin, L.: A hybrid model for missing traffic flow data imputation based on clustering and attention mechanism optimizing LSTM and AdaBoost, Scientific Reports, 14, 26473, https://doi.org/10.1038/s41598-024-77748-1, 2024.

Siddig, N. A., Al-Subhi, A. M., Alsaafani, M. A., and Alraddadi, T. M.: Applying empirical orthogonal function and determination coefficient methods for determining major contributing factors of satellite sea level anomalies variability in the Arabian Gulf, Arabian Journal for Science and Engineering, 1-10, https://doi.org/10.1007/s13369-021-05612-9

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

2021.

Song, W., Gao, C., Zhao, Y., and Zhao, Y.: A time series data filling method based on LSTM—Taking the stem moisture as an example, Sensors, 20, 5045, https://doi.org/10.3390/s20185045, 2020.

Sun, Q., Wan, J., and Liu, S.: Estimation of sea level variability in the China Sea and its vicinity using the SARIMA and LSTM models, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 3317-3326, https://doi.org/10.1109/JSTARS.2020.2997817, 2020.

Tang, F. and Ishwaran, H.: Random forest missing data algorithms, Statistical Analysis and Data Mining: The ASA Data Science Journal, 10, 363-377, https://doi.org/10.1002/sam.11348, 2017.

Tong, T. and Li, Z.: Predicting learning achievement using ensemble learning with result explanation, PloS one, 20, e0312124, https://doi.org/10.1371/journal.pone.0312124, 2025.

Tumse, S. and Alcansoy, U.: Statistical and deep learning approaches in estimating present and future global mean sea level rise, Natural Hazards, 1-28, https://doi.org/10.1007/s11069-025-07203-5 2025.

Vance, J. M., Currie, K. I., Zeldis, J., Dillingham, P. W., and Law, C. S.: An empirical MLR for estimating surface-layer DIC and a comparative assessment of gap-filling techniques for ocean carbon time series, Biogeosciences, 19, 241–269, https://doi.org/10.5194/bg-19-241-2022, 2022, 2022.

Wang, J.: An intuitive tutorial to Gaussian process regression, Computing in Science & Engineering, 25, 4-11, https://doi.org/10.1109/MCSE.2023.3342149, 2023.

Wang, L., Dong, L., and Zhang, Q.: Coupling Interpretable Feature Selection with Machine Learning for Evapotranspiration Gap Filling, Water, 17, 748, https://doi.org/10.3390/w17050748, 2025.

Wang, S., Li, C., and Lim, A.: Why are the ARIMA and SARIMA not sufficient, arXiv preprint arXiv:1904.07632, https://doi.org/10.48550/arXiv.1904.07632, 2019.

Wenzel, M. and Schröter, J.: Reconstruction of regional mean sea level anomalies from tide gauges using neural networks, Journal of Geophysical Research: Oceans, 115, C08013, https://doi.org/10.1029/2009JC005630, 2010.

Wijesekara, L. and Liyanage, L.: Mind the large gap: Novel algorithm using seasonal decomposition and elastic net regression to impute large intervals of missing data in air quality data, Atmosphere, 14, 355, https://doi.org/10.3390/atmos14020355, 2023.

Wu, H., Wang, L., Ling, X., Cui, L., Sun, R., and Jiang, N.: Spatiotemporal reconstruction of global ocean surface pCO2 based on optimized random forest, Science of The Total

https://doi.org/10.5194/essd-2025-227
Preprint. Discussion started: 16 May 2025

Environment, 912, 169209, https://doi.org/10.1016/j.scitotenv.2023.169209, 2024.

Xiao, C., Chen, N., Hu, C., Wang, K., Gong, J., and Chen, Z.: Short and mid-term sea surface temperature prediction using time-series satellite data and LSTM-AdaBoost combination approach, Remote Sensing of Environment, 233, 111358, https://doi.org/10.1016/j.rse.2019.111358, 2019.

Xu, Y., Lin, M., Zheng, Q., Ye, X., Li, J., and Zhu, B.: A study of long-term sea level variability in the East China Sea, Acta Oceanologica Sinica, 34, 109–117, https://doi.org/10.1007/s13131-015-0754-0, 2015.

Young, P. C., Pedregal, D. J., and Tych, W.: Dynamic harmonic regression, Journal of forecasting, 18, 369-394, https://doi.org/10.1002/(SICI)1099-131X(199911)18:6<369::AID-FOR748>3.0.CO;2-K, 1999.

Zhu, S., McCalmont, J., Cardenas, L. M., Cunliffe, A. M., Olde, L., Signori-Müller, C., Litvak, M. E., and Hill, T.: Gap-filling carbon dioxide, water, energy, and methane fluxes in challenging ecosystems: Comparing between methods, drivers, and gap-lengths, Agricultural and Forest Meteorology, 332, 109365, https://doi.org/10.1016/j.agrformet.2023.109365, 2023.