nature communications

Article

https://doi.org/10.1038/s41467-025-62901-9

AI-powered spatiotemporal imputation and prediction of chlorophyll-a concentration in coastal ecosystems

Received: 19 February 2025

Accepted: 4 August 2025

Published online: 18 August 2025

Check for updates

Fan Zhang[1,2], Hiusuet Kung[2,3], Fa Zhang[1], Can Yang[1,2] ✉ & Jianping Gan[1,2,3] ✉

Predicting spatiotemporal Chlorophyll-a (Chl_a) distributions is essential for diagnosing and analysing productivity and ecosystem health of coastal oceans. Yet, current tools remain inadequate for prognosing marine ecosystems through predicting spatiotemporal Chl_a distributions, particularly in the dynamic coastal ocean. Coupled physics-biogeochemical models struggle to resolve complex trophic interactions, while data-driven approaches are limited by incomplete satellite observations. We developed an advanced AI-powered spatiotemporal imputation and prediction (STIMP) model for predicting Chl_a in coastal ocean. STIMP adopts a novel paradigm that first imputes and subsequently predicts Chl_a across a broad spatiotemporal scale, resolving difficulties arising from incompletion, nonstationary temporal variations, and spatial heterogeneity of data through integrating specially designed modules. We demonstrated the STIMP's robust imputation and prediction of Chl_a in four representative global coastal oceans. STIMP reduced the imputation mean absolute error (MAE) by 45.90–81.39% compared with the data interpolating empirical orthogonal function method in geoscience and by 8.92–43.04% against leading AI methods. With accurate imputation, STIMP demonstrated superior predictive accuracy, achieving MAE reductions of 58.99% over biogeophysical models and 6.54–13.68% over AI benchmarks. STIMP offers a new approach for predicting oceans' Chl_a that typically have spatiotemporally limited data.

Coastal oceans are the world's most productive marine ecosystems[1] because terrestrial nutrient input and active hydrodynamics produce high biological productivity and biodiversity[2,3]. The ecosystems of coastal oceans, however, are vulnerable to frequent and severe eutrophication[4], biogeochemical extremes[5], and hypoxia[6] that substantially threaten the sustainability of these coastal environments[7,8] and the blue economies of the coastal zones[9,10].

The concentration of chlorophyll-a (Chl_a) is a key indicator of the overall health of marine environments[11,12]. An increasing number of studies have diagnosed and predicted Chl_a concentration in marine ecosystems. These studies can be broadly classified into two categories: coupled physics-biogeochemistry numerical simulations and data-driven predictions. Simulations predict the concentrations by integrating a suit of physical and biogeochemical equations[13–15]. Unfortunately, complex and unknown trophic transfers in ocean ecosystems pose qualitative and quantitative challenges[16] in these coupled physical and biogeochemical models so that factors such as energy fluxes and biomass are not represented well because of simplified

[1]Department of Mathematics, The Hong Kong University of Science and Technology, Hong Kong SAR, China. [2]Center for Ocean Research in Hong Kong and Macau, The Hong Kong University of Science and Technology, Hong Kong SAR, China. [3]Department of Ocean Science, The Hong Kong University of Science and Technology, Hong Kong SAR, China. ✉e-mail: macyang@ust.hk; magan@ust.hk

biogeochemical cycles. Unlike open ocean, Chl_a identification and prediction are even more challenge in the coastal waters under controls of multiple forcings from terrestrial, active coupled hydrodynamic-biogeochemical processes and influx from open ocean. These errors accumulate[17] continuously during the long-term integration of nonlinear systems, ultimately affecting the accuracy of the simulation results. Many works[18,19] intend to integrate observation of Chl_a and coupled physics-biogeochemistry numerical simulation model to mitigate error accumulations. By calibrating the state variables, these methods enhance the forecasting capability of ocean numerical simulations. However, error accumulations are inevitable in coupled physics-biogeochemistry numerical simulations[20,21].

In contrast, data-driven approaches that use historic observations allow discovery of underlying dynamic patterns by analysing spatiotemporal variations of Chl_a. Without requiring extensive oceanic knowledge, computational-friendly data-driven approaches, such as mechanistic[22], statistical[23,24] and deep learning methods[25–27], have gained more attention in marine ecosystem study by offering an alternative solution to the numerical modelling challenges. Moreover, remote sensing measurements provide precious data support for the developing data-driven spatiotemporal Chl_a prediction methods. Sea-viewing Wide Field-of-view Sensor (SeaWiFS), Moderate Resolution Imaging Spectroradiometers (MODIS), including Terra MODIS and Aqua MODIS, and other remotely sensing data provide near-daily coverage of the global surface ocean[28] and capture large-scale changes of phytoplankton biomass in the ocean[29].

However, there remain three challenges in developing large-scale spatiotemporal Chl_a data-driven prediction method. First, temporal variations are difficult to capture in existing data. Temporal variations of Chl_a are controlled by complex physical and biogeochemical stressors, such as water stratification, nutrient supply, and light conditions, whose variability exacerbate the unpredictability of Chl_a concentrations, even over short periods[30]. Second, spatial heterogeneity and relationships are difficult to model, and distinct temporal patterns in Chl_a distributions have been observed across different locations due to terrain, ocean currents, and other conditions. Third, high rates of missing observations render spatiotemporal variations more challenging to comprehend. Satellite remote sensing measurements are highly susceptible to environmental factors such as orbital position of satellite and cloud coverage[31]. In a typical coastal ocean such as the Pearl River Estuary, one of our study sites in the northern South China Sea (Fig. 1a), the average rate of missing 8-day Chl_a data is 50.29%. These kinds of spatiotemporally limited observation data hinder resolving the variations of Chl_a that are needed for accurate predictions.

In this study, we developed an advanced AI-powered spatiotemporal imputation and prediction (STIMP) model for predicting Chl_a in coastal ocean. The success of STIMP is attributed to the paradigm of first imputing and subsequently predicting. High rates of missing observations render spatiotemporal variations of Chl_a and pose challenges to comprehend and predict of the ecosystem. Most data-driven methods are able to make the prediction of Chl_a when the observation datasets are complete, but struggle to identify its temporal patterns and spatial relationships when there are only partial datasets of observations. STIMP decomposes the prediction of Chl_a into two sequential steps: 1) *the imputation process*, which reconstructs multiple potential complete spatiotemporal Chl_a distributions from partial observations, and 2) *the prediction process*, which accurately predicts Chl_a based on each reconstructed continuous and complete spatiotemporal Chl_a distribution. Using Rubin's rules[32], the final Chl_a prediction is obtained by averaging the outcomes of multiple imputation and prediction processes. In this way, our STIMP method not only improves the overall predictive performance through accurate imputation of missing data but also provides confidence intervals to quantify the prediction uncertainties.

Moreover, STIMP can address the three challenges of incompletion, nonstationary temporal variations, and spatial heterogeneity of data that we are facing by integrating three specially designed modules. We conducted benchmark studies of four representative global coastal regions: northern Gulf of Mexico, Chesapeake Bay, the Yangtze River Estuary, and the Pearl River Estuary. These four regions all suffer eutrophication and hypoxia due to explosively growths of algae with high concentrations of Chl_a. STIMP demonstrated superiority in large-scale spatiotemporal Chl_a imputation and prediction over baseline methods. Our results also indicated that the importance of imputation in the prediction. STIMP has proofed its capabilities to diagnose and prognose broad oceans that typically have spatiotemporally limited data.

Results

STIMP method

STIMP is a deep learning-based method that accurately imputes and predicts Chl_a in the coastal oceans (Fig. 1). The inputs for STIMP include observations of Chl_a from coastal oceans, denoted as $\mathbf{X}^{ob}$, and a spatial graph, $\mathbf{G}$, that contains the geographic coordinates of the observations. STIMP simultaneously outputs a complete Chl_a dataset, $\mathbf{X}$, and accurately predicts Chl_a, $\tilde{\mathbf{Y}}$, based on $\mathbf{X}$. STIMP and is formulated as:

$$p(\tilde{\mathbf{Y}}|\mathbf{X}^{ob}) = \int_{\mathbf{X}} p_{\Phi}(\tilde{\mathbf{Y}}|\mathbf{X}) p_{\theta}(\mathbf{X}|\mathbf{X}^{ob}) d\mathbf{X}. \quad (1)$$

Specifically, the primary challenge in learning the imputation function $p_{\theta}(\mathbf{X}|\mathbf{X}^{ob})$ lies in approximating the spatiotemporal distribution of Chl_a. We designed a Spatiotemporal Denoising Diffusion Model (STDDM) as the imputation function. Specifically, STDDM decompose the complicated task into L simple tasks, with the objective of increasing the signal-to-noise ratio. Finally, the spatiotemporal distribution of Chl_a is transformed from an easy-to-sample distribution $p(\mathbf{X}_L) = \mathcal{N}(\mathbf{X}_L; \mathbf{0}, \mathbf{I})$, using a joint distribution in the form of a Markov chain:

$$p_{\theta}(\mathbf{X}|\mathbf{X}^{ob}) = \int p(\mathbf{X}_L) p_{\theta}(\mathbf{X}|\mathbf{X}_1, \mathbf{X}^{ob}) \prod_{l=L}^{2} p_{\theta}(\mathbf{X}_{l-1}|\mathbf{X}_l, \mathbf{X}^{ob}) d\mathbf{X}_{1:L}. \quad (2)$$

Then the large scale spatiotemporal future Chl_a distribution is predicted using the function, $p_{\Phi}(\tilde{\mathbf{Y}}|\mathbf{X})$.

The network structures of the imputation and prediction functions are shown in the Fig. 1. For both imputation and prediction functions, nonstationary temporal variations and spatial heterogeneity of Chl_a are all difficult and crucial. For nonstationary temporal variations, we employ Temporal Linear Transformer (TLT). TLT preserves crucial information regarding the Chl_a variations, computing across all elements of the time series to facilitate understanding of the temporal patterns of Chl_a. For spatial heterogeneity, we designed a Heterogeneous Spatial Graph Neural Network (HSGNN). HSGNN learns location-specific parameters from parameter pools to represent spatial heterogeneity. By doing so, distinct temporal patterns across different locations can be well accommodated in spatial-specific parameters in a computationally friendly way. We reuse the same structure of TLT and HSGNN in both the imputation function and the prediction function, ensuring that the spatiotemporal relationships of Chl_a can be well utilized in both tasks.

Benchmarking spatiotemporal imputation and prediction

We evaluated how well STIMP separately imputed and predicted Chl_a concentrations. We used MODIS Aqua project data[33] for: (1) the Pearl River Estuary, 4325 positions within 20.52–22.98°N, 112.02–115.98°E, with rates of missing data for each position ranging from 32.06% to 89.56% with an average rate of missing data across all positions equal

https://doi.org/10.1038/s41467-025-62901-9

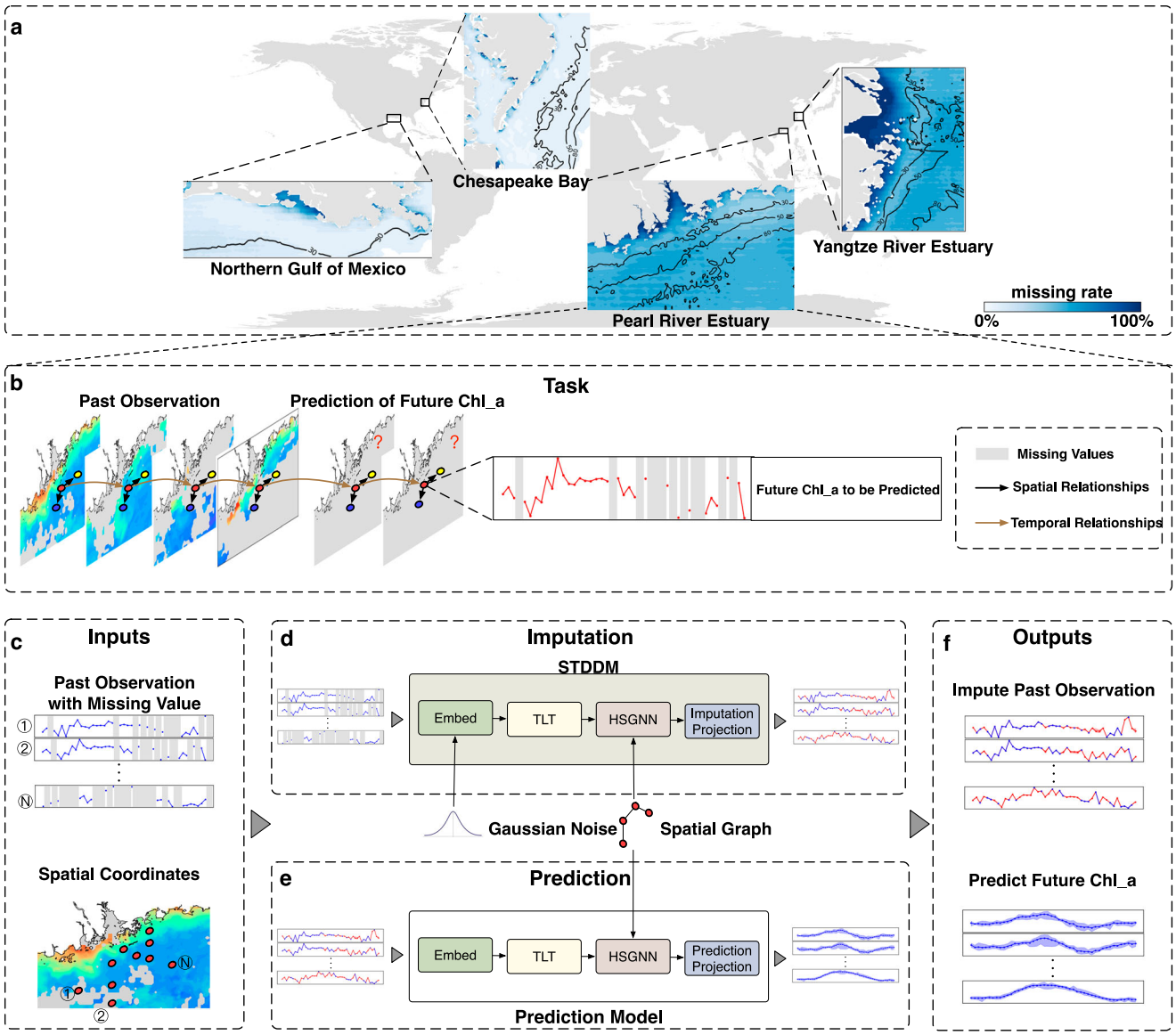

Fig. 1 | Overview of STIMP. a We applied STIMP to four representative global coastal regions: the Pearl River Estuary, northern Gulf of Mexico, Chesapeake Bay, and the Yangtze River Estuary. **b** STIMP is designed to predict future Chlorophyll-a (Chl_a) according to past partial observation. **c**–**f** The overall framework for STIMP involves two successive stages: Imputation and Prediction. **c** STIMP uses past partial observations of Chl_a, with spatial coordinates of the observations as inputs. **d** STIMP uses the Spatiotemporal Denoising Diffusion Model (STDDM) to reconstruct a completed Chl_a dataset with spatiotemporal relationships from the past partial observation dataset. **e** STIMP outputs the future Chl_a based on the completed Chl_a dataset. **f** The outputs of STIMP, including imputed past observations of Chl_a and predicted future Chl_a for the entire ocean, can be used for marine conservation initiatives and policymaking for coastal regions.

to 50.29%; (2) northern Gulf of Mexico, 2907 positions within 28.52° to 29.98°N, 89.02° to 93.98°W, with rates of missing data for each position ranging from 5.35% to 94.23% with an average rate of missing data across all positions equal to 18.19%; (3) Chesapeake Bay, 1893 positions within 36.52–38.98°N, 74.52–76.48°W, with rates of missing data for each position ranging from 4.09% to 88.05% with an average rate of missing data across all positions equal to 10.49%; and (4) the Yangtze River, 4572 positions within 28.02° to 31.98°N, 121.02° to 123.98°E, with rates of missing data for each position ranging from 41.93% to 89.94% with an average rate of missing data across all positions equal to 54.28%. All data are from 01 January 2003 to 22 September 2023. For default settings, training data span 1 January 2003 through 31 January 2015, with the independent test set encompassing observations from 7 February 2015 to 22 September 2023.

For the spatiotemporal imputation, we assessed the performance of STIMP based on four categories, comparing our results with different baseline methods in these categories: (1) a data interpolating empirical orthogonal function (DINEOF) method[34,35] from geoscience, note that, we selected $t = 1$ in our implementation after conducting a comprehensive comparison of various values of t (Supplementary Figs. S11 and S12); (2) classic statistical methods, including Slide Window and Lin-ITP; (3) spatiotemporal machine learning method, including TRMF[36]; (4) deep learning methods, including CSDI[37], ImputeFormer[38], Inpainter[39] and MaskedAE[40]. Because our collected data were not ground truthed for unobserved data, we randomly selected nine different rates of missing data, ranging from 10% to 90%, for choosing observed data as imputation targets. In so doing, we were able to verify the performance of all methods with different rates of missing data. Overall, we saw that STIMP outperformed the other baseline methods for different datasets in terms of the mean absolute error (MAE), demonstrating we could apply STIMP effectively to different regions. STIMP significantly reduced MAE by 45.90–81.39% compared to

https://doi.org/10.1038/s41467-025-62901-9

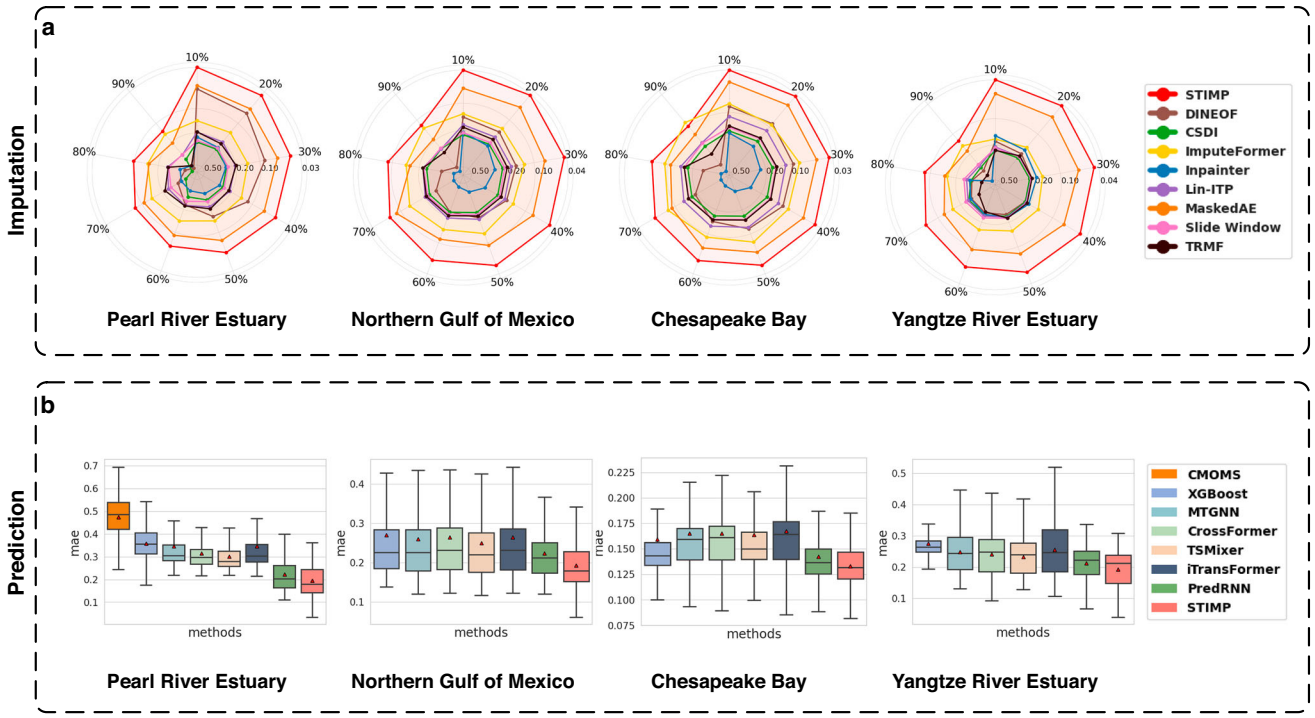

Fig. 2 | A benchmarking study of spatiotemporal imputation and prediction of Chl_a for four global representative coastal oceans. a Overall mean absolute error (MAE) between ground truth and imputed Chl_a for different interpolation methods. For each coastal ocean, we constructed datasets with nine different rates of missing data, ranging from 10% to 90%. Source data are provided as a Source Data file. **b** Overall MAE between ground truth and predicted Chl_a of different prediction methods. Source data are provided as a Source Data file.

DINEOF, and by 8.92–43.04% compared to the state-of-the-art AI methods. STIMP reduced mean square error (MSE) by 69.47–95.64% compared to DINEOF, and by 12.04–68.30% compared to the state-of-the-art AI methods (Supplementary Fig. S1). In addition, DINEOF performed even worse than linear interpolation Lin-ITP due to the complexity of the spatiotemporal Chl_a distribution. Specifically, the intricate relationships present in the large-scale, long-term observations challenged DINEOF in effectively capturing the spatiotemporal relationships using singular value decomposition. The performance of CSDI and ImputeFormer was inevitably unsatisfactory because neither considers spatial relationships during imputation. Our results indicate that the spatial and temporal relationships are essential for imputing data.

We comprehensively compared different methods of predicting Chl_a (Fig. 2b). For the spatiotemporal prediction, we compared results against baseline methods in three categories: (1) machine learning method, XGBoost[41]; (2) time series prediction methods, including CrossFormer[42], TSMixer[43] and iTransFormer[44]; (3) spatiotemporal prediction methods, including MTGNN[45] and PredRNN[46]. We also compared our STIMP results with results from the well-calibrated numerical model, CMOMS[15] for Pearl River Estuary. STIMP performed well for all four regions. Compared to the second best method, STIMP reduced MAE by 12.24%, 13.68%, 6.54%, and 9.72% (Fig. 2b) and reduced MSE by 22.28%, 22.81%, 15.37%, and 12.17% for all four regions (Supplementary Fig. S2). PredRNN performed the second best. PredRNN uses spatiotemporal recurrent neural networks with long short-term memory to memorize spatial appearances and temporal variations, performing better than time series prediction methods. The results demonstrate the importance of the spatiotemporal relationships for accurate prediction. However, due to the inherent lack of observations of Chl_a, PredRNN could not perform as well as STIMP.

STIMP enables spatiotemporal imputation

We used the Pearl River Estuary to demonstrate the effectiveness of STIMP spatiotemporal imputation. The number of missing observations from the Pearl River Estuary reflects the data deficiencies for most estuaries. The advantage of using the Pearl River Estuary as an example is that we have results from numerical simulations to compare to our STIMP results. There are several major cities in the Pearl River Estuary, including Shenzhen, Guangzhou, and Zhuhai, where rapid industrialization and urbanization have adversely impacted water quality in downstream of the Pearl River system[47]. Eutrophication has produced serious red tide events or harmful algal blooms in the estuary and its adjacent coastal waters since the 1980s, causing economic losses[48].

High rates of missing observations prevent completing a picture of the spatiotemporal Chl_a relationships for the Pearl River Estuary, impeding forecasting and analysis. However, STIMP enabled large-scale spatiotemporal imputation of the Pearl River Estuary data. STIMP performed exceptionally compared to other imputation methods for one-year imputation in terms of MAE, reducing MAE significantly by 45.90% to 77.35% compared to DINEOF and by 10.20% to 40.38% compared to the second best model (Fig. 2a). Despite the default length of the time series on imputation, we tested the performance of STIMP in imputation over periods of 6 months and one-and-a-half years (Fig. 3a). The results demonstrated that the length of the time series on gap-filling did not affect the superiority of STIMP. For 6 months imputation, STIMP significantly reduced MAE by 44.34–76.54% compared to DINEOF, and 4.64–38.22% compared to the state-of-the-art AI methods. For one-and-a-half years imputation, STIMP achieved MAE reductions of 50.34–76.43% compared to DINEOF, and of 0.33–38.17% compared to the state-of-the-art AI methods.

We selected observations of the entire Pearl River Estuary from 07 February 2015 to 02 February 2016 and used STIMP and the baseline methods, including data interpolating empirical orthogonal function (DINEOF) method, MaskedAE, and Lin-ITP, to reconstruct the Chl_a distribution to investigate why STIMP imputation of the Pearl River Estuary data outperformed the baseline methods. We examined the

https://doi.org/10.1038/s41467-025-62901-9

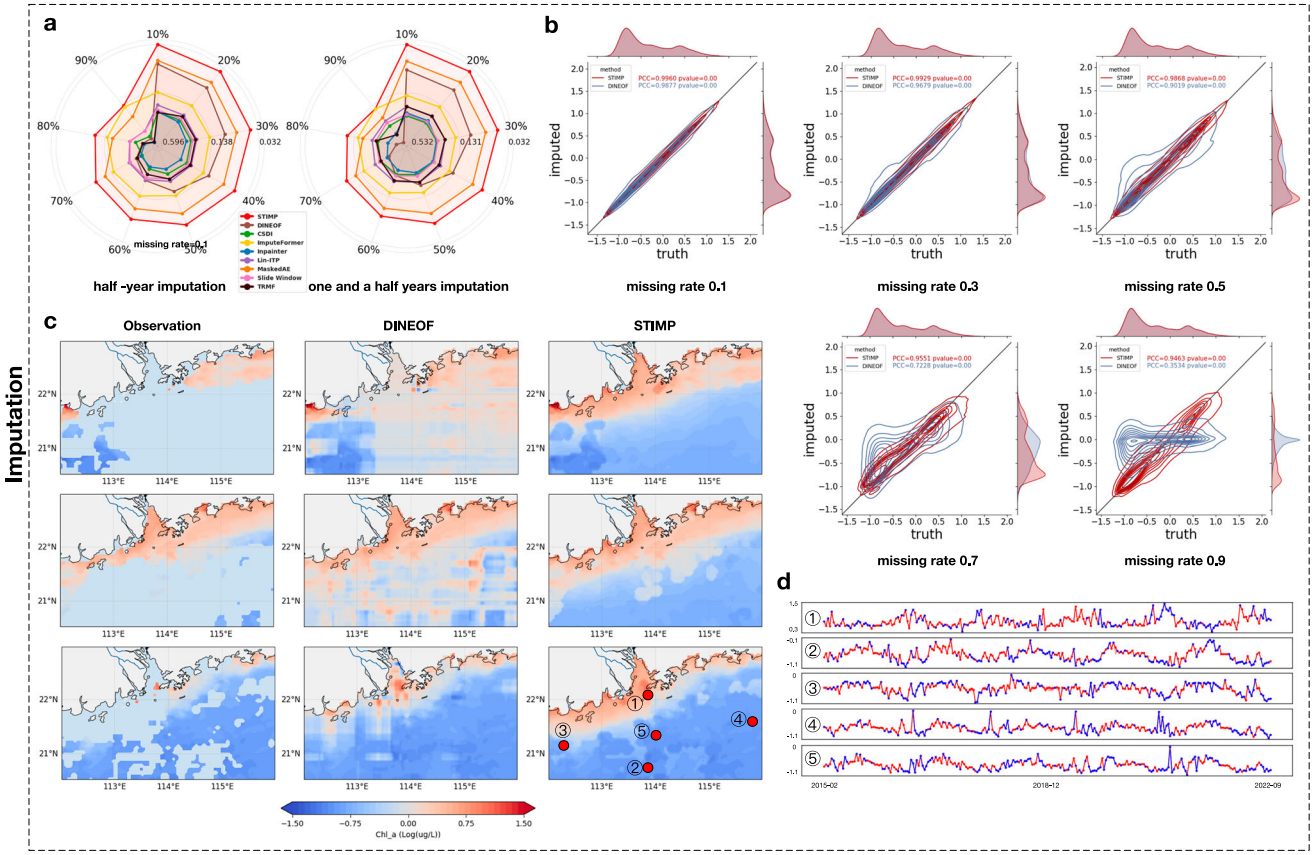

Fig. 3 | SIIMP enables large scale spatiotemporal imputation of chlorophyll-a (Chl_a) in the Pearl River Estuary. a Overall mean absolute error (MAE) between ground truth and imputed Chl_a for different imputation methods over half-year and one-and-a-half-year periods. **b** Pearson correlation coefficient (PCC) between imputed data and ground truth data of STIMP and data interpolating empirical orthogonal function (DINEOF) method at rates of missing data equal to 10%, 30%, 50%, 70%, and 90%. **c** Measured and imputed Chl_a distribution in the Pearl River Estuary. The first column from the left shows the original observations with many missing values. The second and third columns show the corresponding imputed Chl_a from DINEOF and STIMP. **d** Imputed Chl_a from STIMP at five locations. The original observations are represented by blue lines, and the imputed Chl_a are represented by red lines.

correlation between ground truth and imputed data of STIMP and DINEOF using rates of missing data of 10%, 30%, 50%, 70% and 90% (Fig. 3b). We observed that the Pearson correlation coefficients (PCC) of the STIMP imputed data and ground truth data were greater than 0.90 even when the rate of missing data euqalled 90%. DINEOF was prone to fill in gaps with zero values when missing rate was high and could not perform singular value decomposition of the spatiotemporal distribution of Chl_a. When comparing STIMP results to MaskedAE and Lin-ITP results (Supplementary Fig. S3a), we saw that MaskedAE PCC values were comparable to STIMP when rates of missing data were 10% and 50%. However, MaskedAE imputation severely degenerated when rates of missing data were high. In that case, STIMP PCC (correlated to ground truth data) improved by 6.65% over MaskedAE. In addition, the results illustrated that Lin-ITP introduced significant errors which hindered the analysis of the Pearl River Estuary data.

Because STIMP's imputation is based on spatiotemporal relationships, using case studies, we explored whether the imputed values from STIMP aligned with actual spatiotemporal relationships. STIMP effectively preserved the spatial relationships during the imputation compared to DINEOF on 07 February 05 May, and 10 October 2015 (Fig. 3c). STIMP generated larger values close to the coastline, and STIMP generated similar values for most regions. STIMP effectively reconstructed the completed data even when the rate of missing data was high (see 07 February 2015). Moreover, STIMP effectively preserved temporal relationships during imputation. In Fig. 3d, we show the imputation of five single positions from 07 February 2015 to 22 September 2022, where blue represents the observations and red represents the imputation. Compared to simple linear interpolation, STIMP contained more fluctuations, especially for long-term missing data. Overall, we saw that the STIMP results were significantly superior compared to the results from the baseline methods.

STIMP offers spatiotemporal prediction

Predicting Chl_a is critical to diagnosing future sustainable marine environments and formulating effective mitigation schemes, particularly under a changing climate that will warm and stratify the vulnerable ocean. With the continuous and completed spatiotemporal Chl_a distribution imputed from partial datasets of observations, we found that STIMP accurately predicts Chl_a.

STIMP has more advantages in long-term prediction. As shown in Fig. 4a, STIMP gained more improvement for long-term prediction. For 1 year prediction, STIMP achieved 6.54–13.68% mean absolute error (MAE) reduction over baselines. For 2 years prediction, STIMP reduced 13.68–32.25% MAE. For 3 years prediction, STIMP achieved 13.77–32.01% MAE reduction. STIMP outperformed other prediction methods due to its paradigm of first imputing then predicting and three elaborately designed modules. Without imputation, our predictive model results still improved compared to PredRNN (Fig. 4b), indicating that our designed TLT and HSGNN modules are more suitable for capturing the spatiotemporal relationships of Chl_a than the design of PredRNN. We examine the advantages of having these two modules when we describe our ablation study of STIMP later in supplementary. Moreover, imputation significantly improved how well STIMP predicted distributions (Fig. 4b, STIMP w/o imputation versus

https://doi.org/10.1038/s41467-025-62901-9

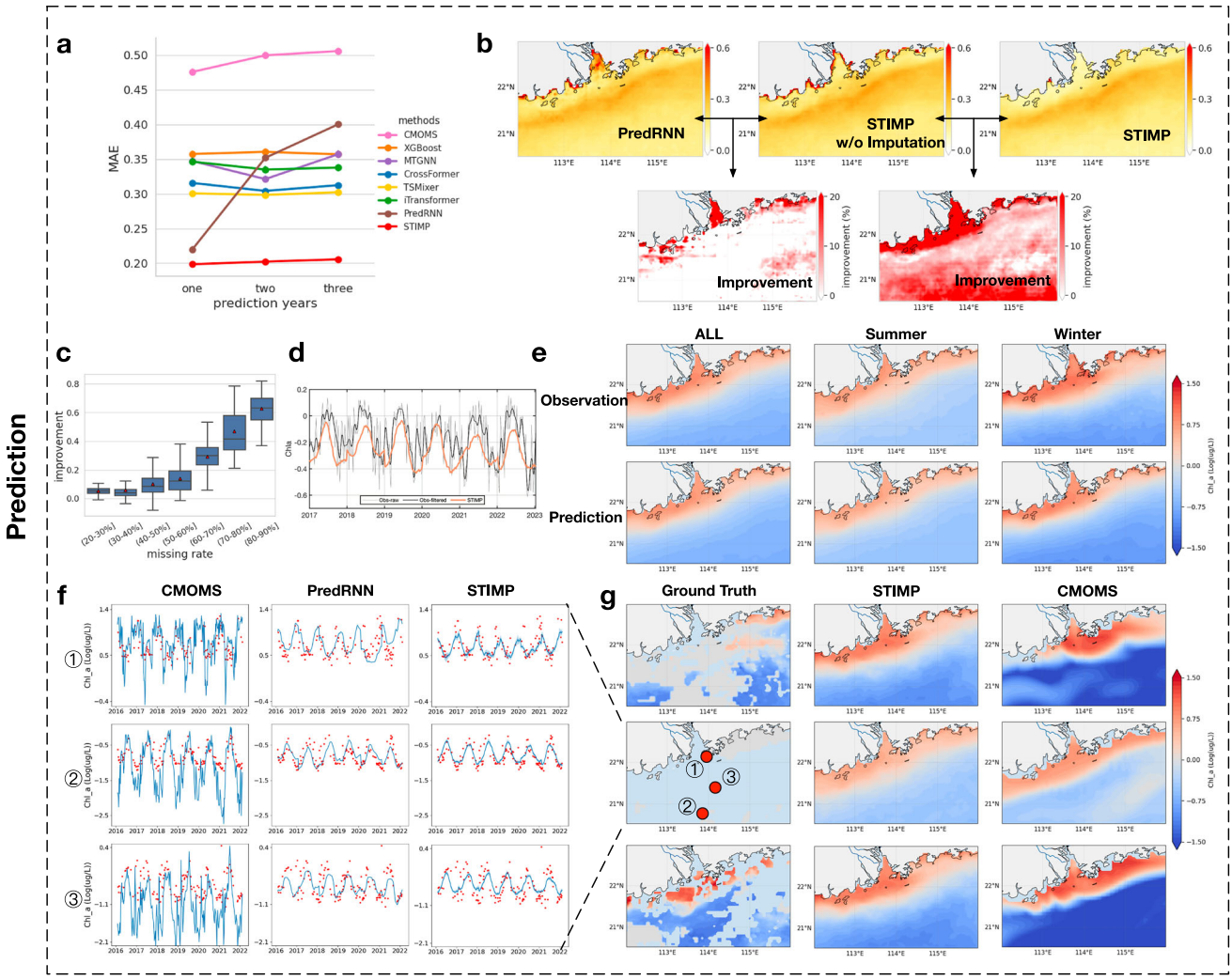

Fig. 4 | SIIMP enables large scale spatiotemporal prediction of chlorophyll-a (Chl_a) in the Pearl River Estuary. a The performance of STIMP and the baseline models in terms of mean absolute error (MAE) for 1-year, 2-year, and 3-year predictions. **b** MAE between the actual values and the predictions made by PredRNN, STIMP without imputation and STIMP. **c** The relationship between the improved performance due to imputation and the rates of missing data. **d** Time series of domain mean Chl_a from January 2017 to December 2022. **e** Horizontal distribution of observed and predicted temporally averaged Chl_a. **f** Predicted Chl_a at three locations. The first to third columns show the Chl_a predicted by CMOMS, PredRNN, and STIMP. The red scatter points represent the observations. The blue shading represents the confidence intervals of STIMP to quantify the prediction uncertainties. Source data are provided as a Source Data file. **g** Predicted Chl_a for the entire Pearl River Estuary. The first column shows the true observations. The second and third columns show the Chl_a predicted by STIMP and CMOMS.

STIMP). In locations with higher rates of missing data, STIMP results often improved more than PredRNN results (Fig. 4c). Our analysis demonstrated that imputing before predicting contributed to STIMP effectively capturing the spatial distribution and the seasonal signal of Chl_a. The maxima of observed annual mean Chl_a are in the river, but large Chl_a concentrations are spread along the coast in the inner shelf and decrease with distance from the coast (Fig. 4d). The time series of the domain-averaged Chl_a has a strong seasonal cycle where Chl_a concentrations in the Pearl River Estuary peak during summer when nutrient inputs from the Pearl River are large (Fig. 4e).

We conducted case studies at three specific locations in the Pearl River Estuary to further demonstrate the superiority of STIMP over other baseline methods: 1) 22°9′N, 113°49′E; 2) 20°41′N, 113°41′E; and 3) 21°16′N, 114°1′E. The three sites correspond to three marine regions delineated by isobaths at 30 m and 50 m below sea level (Fig. 4f). Data are from 2 February 2016 to 2 February 2022. We compared STIMP to CMOMS[15], which is a widely used numerical model for analysing ecosystem dynamics in the Pearl River Estuary. STIMP can not only provide accurate Chl_a prediction (blue lines in Fig. 4i), but also provide confidence intervals, which is visualized as shaded regions around the predicted values, too quantify the prediction uncertainties. STIMP outperformed PredRNN by a large margin, with MAE improving by 30.28%, 22.09%, and 1.83%. CMOMS yielded large MAE because the model severely underestimated the Chl_a. STIMP improved the MAE by 53.78%, 74.63%, and 66.33% compared to CMOMS. In Fig. 4g, we present the prediction results for the entire Pearl River Estuary for 2 February 2016, 16 July 2016, and 8 January 2017. We observed that STIMP predictions consistently agreed with the observations, providing larger predicted values near the coastline. CMOMS, on the other hand, severely underestimated the Chl_a in the region far from the coastline, especially in winter.

Applying STIMP to the global ocean

We examined how effectively STIMP can be applied to global coastal oceans by examining how well the model predicted Chl_a for the northern Gulf of Mexico, Chesapeake Bay and the Yangtze River Estuary.

The Gulf of Mexico, ranking as the ninth largest body of water globally[49], is also among the most economically and ecologically productive water bodies in the world[50]. The average rate of missing

https://doi.org/10.1038/s41467-025-62901-9

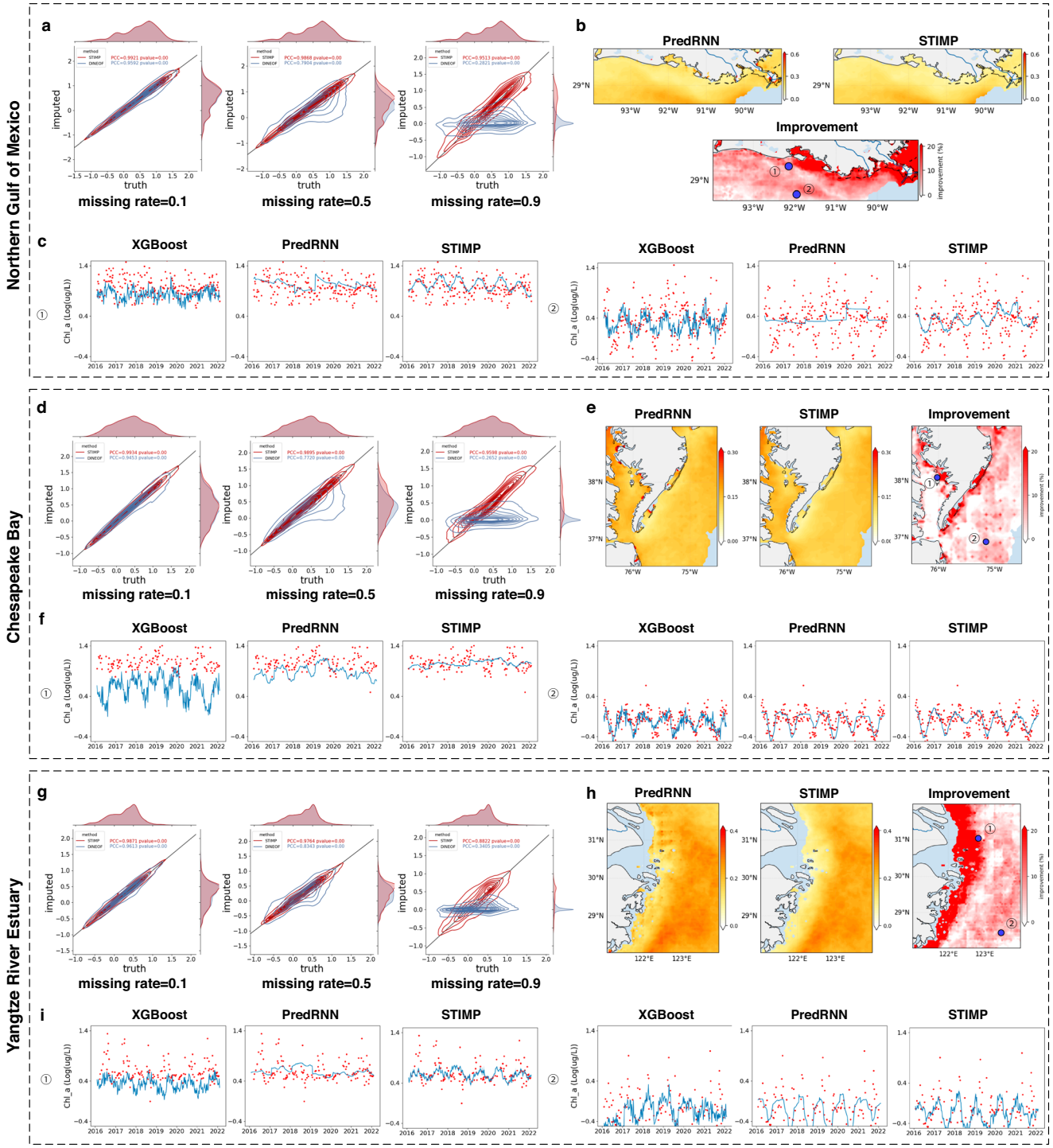

Fig. 5 | STIMP can be applied to other coastal ocean areas. We applied STIMP to other three representive coastal ocean areas: Northern Gulf of Mexico (**a**–**c**), Chesapeake Bay (**d**–**f**) and Yangtze River Estuary (**g**–**i**). **a**, **d**, **g** Pearson correlation coefficient (PCC) between imputed data and ground truth data with the rates of missing data equal to 10%, 50%, and 90%. **b**, **e**, **h** MAE between the actual values and the predictions made by PredRNN (the first column) and STIMP (the second column). The third column indicates how much STIMP reduced the error compared to PredRNN. **c**, **f**, **i** XGBoost, PredRNN, and STIMP predicted Chl_a at two locations. The red scatter points represent the observations. Source data are provided as a Source Data file.

data across all positions for the northern Gulf of Mexico was 18.19%. When imputing data, as shown in Fig. 2a, STIMP performed well compared to other imputation methods, significantly reducing the error by 69.42–74.88% compared to DINEOF and by 8.92–43.04% compared to the second-best model in terms of MAE. The observations for the entire northern Gulf of Mexico were from 07 February 2015 to 02 February 2016. Our findings indicated that the correlation between the STIMP imputed Chl_a and the ground truth data was significantly higher than that of DINEOF imputed Chl_a and ground truth data with rates of missing data equal to 10%, 50%, and 90%, improving PCC by 6.77%, 24.58% and 226.74% Fig. 5a). We also compared STIMP to MaskedAE and Lin-ITP methods (Supplementary Fig. 3b). STIMP improved PCC by 1.07–6.77% compared to MaskedAE, and STIMP improved PCC by 10.90% to 25.75% compared to linear

interpolation (Lin_ITP). For prediction, we showed that STIMP's effective data imputation accurately predicted large-scale spatiotemporal Chl_a. STIMP predictions improved, compared to PredRNN, for most areas of the northern Gulf of Mexico (Fig. 5b). From the Chl_a distribution (Supplementary Fig. 6), we saw that locations close to the river mouth had the largest Chl_a, and the outer shelf had the smallest Chl_a. Hence, we examined results (Fig. 5c) at two positions: 1) 29°29′N, 92°11′W and 28°54′N, 91°59′W, from February 2016 to 2 February 2022. Due to the lack of open-source numerical methods, we compared STIMP with XGBoost and PredRNN for these two positions. At the two positions, STIMP reduced MAE by 26.55% and 16.02% compared to XGBoost and reduced MAE by 6.75% and 10.84% compared to PredRNN. The satellite data from the northern Gulf of Mexico had considerable noise. Predictions generated by PredRNN tended to lose periodicity when confronted by the low signal-to-noise ratio data. In contrast, predictions produced by STIMP successfully captured six complete periods, which corresponded to a period of six years.

Chesapeake Bay is one of the most productive estuaries in the United States. The average rate of missing data across all positions in Chesapeake Bay was 10.49%. For imputation, as shown in Fig. 2a, STIMP performed exceptionally compared to the other imputation methods with regard to MAE. STIMP significantly reduced MAE by 62.08% to 75.63% compared to DINEOF and by 26.91% to 33.46% compared to the second-best model, except when the rate of missing data was 90%. We selected observations from the entire bay from 07 February 2015 to 02 February 2016 for this case study. STIMP demonstrated excellent imputation performance compared to DINEOF, improving PCC by 5.09%, 28.17% and 264.67% (Fig. 5d), MaskedAE and Lin-ITP (Supplementary Fig. 3c). Moreover, STIMP improved its predictions compared to PredRNN for most of Chesapeake Bay (Fig. 5e). The case studies at the two locations at 1) 29°29′N, 92°11′W and 2) 28°54′N, 91°59′W indicated that STIMP more accurately predicted Chl_a than the other two methods. STIMP reduced MAE by 72.10% and 10.39% compared to XGBoost and reduced MAE by 44.34% and 3.99% compared to PredRNN. Overall, STIMP improved MAE by 6.54% compared to PredRNN.

The Yangtze River is the third longest river in the world. The river discharges into the adjacent Yangtze River Estuary near the megacities, Shanghai and Hangzhou. The average rate of missing data across all positions in the Yangtze River Estuary was 54.28%. For imputation, as shown in Fig. 2a, STIMP performed exceptionally compared to the other imputation methods, reducing the MAE by 68.31% to 90.92% compared to DINEOF and by 15.62% to 42.67% compared to the second-best model. The observations for the entire Yangtze River Estuary were from 07 February 2015 to 02 February 2016. STIMP imputed the data well compared to DINEOF (Fig. 5g), MaskedAE and Lin-ITP (Supplementary Fig. 3d). Moreover, STIMP improved predictions compared to PredRNN for most areas of the Yangtze River Estuary (Fig. 5h), especially for the regions closest to the river mouth. As shown in Fig. 4i, at 31°34′N, 122°31′E, STIMP reduced MAE by 45.02% and 9.95% compared to the two baseline methods. At 28°24′N, 123°21′E (outer shelf), STIMP's error improvement was relatively smaller, but MAE still reduced by 14.45% and MAE reduced by 6.59% compared to the two baseline methods.

To summarize, we demonstrated that STIMP can be successfully applied to the global coastal oceans. STIMP is a versatile and effective deep learning framework. STIMP not only reconstructs a complete Chl_a dataset by exploring the spatiotemporal relationships in the partial dataset of observations, STIMP also accurately predicts spatiotemporal distributions at large scales. STIMP even successfully captures periodic data when confronted with low signal-to-noise ratio data (for example, for the northern Gulf of Mexico and the Yangtze River Estuary).

Discussion

In this paper, we presented STIMP, an effective and versatile deep learning framework for analysing Chl_a using satellite remote sensing measurements. Our results demonstrated that STIMP effectively addresses the challenges of satellite remote sensing of Chl_a when diagnosing and prognosing coastal ocean ecosystems by overcoming high rates of missing data, composing complex spatial heterogeneity, and determining diverse dynamic temporal patterns.

Specifically, STIMP imputes historical partial observations before predicting the future of Chl_a for coastal oceans, formulated as $p(\tilde{\mathbf{Y}}|\mathbf{X}^{ob}) = \int_{\mathbf{X}} p_{\Phi}(\tilde{\mathbf{Y}}|\mathbf{X}) p_{\theta}(\mathbf{X}|\mathbf{X}^{ob}) d\mathbf{X}$. Most AI methods, such as PredRNN, are designed to predict based on complete observational datasets. Directly applying these other AI methods to partial datasets results in substantial spatiotemporal relationships being disregarded. STIMP breaks down Chl_a prediction into two steps: 1) $p_{\theta}(\mathbf{X}|\mathbf{X}^{ob})$ to reconstruct a complete spatiotemporal Chl_a distribution from partial observations; and 2) $p_{\Phi}(\tilde{\mathbf{Y}}|\mathbf{X})$ to accurately predict Chl_a based on a set of continuous and complete spatiotemporal relationships. STIMP addresses the difficulty of the task and enables us to evaluate the confidence intervals associated with our predictions. When we were designing the model, we leveraged a diffusion model[51] (an elegant and powerful generative AI method) for the spatiotemporal Chl_a imputation, i.e. $p_{\theta}(\mathbf{X}|\mathbf{X}^{ob})$. Using the diffusion model offered distinct advantages for imputation compared to the traditional interpolating empirical orthogonal function (DINEOF)[52] in Geoscience and other AI-powered data imputation methods. We designed TLT and HSGNN to capture the diverse dynamic temporal patterns found in an entire time series and the spatial heterogeneity of the entire coastal ocean, respectively. STIMP demonstrated capability in large-scale spatiotemporal Chl_a imputation and prediction in four representative coastal oceans. The length of the time series does not affect the superiority of STIMP. Our results indicate that imputation significantly improved the prediction. In addition, STIMP was robust when presented with various rates of missing data, further showing that the model could be applied to global coastal oceans with different rates of missing data.

Despite its advantages, STIMP does have limitations. The advantage of AI-based methods lies in predicting the seasonal and trend components of time series. Many methods[43,53] incorporate seasonal-trend decomposition within each neural block to enhance the predictability of time-series data[54]. Analogous to other AI-based time series methods, STIMP robustly capture the seasonal and trend signals of Chl_a but is incapable of effectively capturing anomaly peak signal. Through extensive analysis in the Pearl River Estuary, compared to weekly and monthly mean predictions, STIMP effectively captures fine-scale variability (Supplementary Figs. S13–S15). We also found that even though the numerical simulation method CMOMS[15] has a large prediction error, it performs better than AI-based data-driven methods in predicting extreme signals. Therefore, we believe that integrating numerical simulation methods with AI-based data-driven methods will be a challenging yet promising direction that we need to address in the future.

Moreover, we focused on satellite remotely sensed data with a spatial resolution of 4 km and a temporal resolution of 8 days. In contrast to CMOMS, STIMP demonstrates low sensitivity to variations in spatial and temporal resolutions. Yet STIMP depends on observational data to capture Chl_a's spatiotemporal patterns. When applied to MODIS's daily Chl_a data with $>90\%$ missing rates, we might need to incorporate additional prior information to address the issue.

In conclusion, based on available satellite observations, STIMP significantly advances large-scale spatiotemporal analysis, diagnostics, and prediction of coastal marine ecosystem health. We expect that STIMP can be broadly applied to future ocean AI-measurement-modelling studies to enhance management and mitigation of worsening coastal ocean ecosystem conditions.

Methods

Rapidly developing satellite remote sensing observations offer opportunities to develop data-driven large-scale spatiotemporal Chl_a prediction method. However, three major challenges must first be solved. First, temporal variations induced by complex physical and biogeochemical stressors are difficult to capture. Second, spatial heterogeneity and relationships are difficult to represent. Third, the high rates of missing satellite observations compromise the completeness of spatiotemporal variations of Chl_a, rendering these observations inadequate to use for accurate prediction.

STIMP is an AI-powered unified framework composed of two steps: 1) spatiotemporal satellite Chl_a data imputation p_θ and 2) Chl_a prediction p_Φ. By doing so, STIMP effectively reconstructs the complete spatiotemporal variations of Chl_a, $\mathbf{X}$, from partial observations, $\mathbf{X}^{ob}$, subsequently providing accurate large-scale spatiotemporal predictions of Chl_a $\tilde{\mathbf{Y}}$ across coastal oceans, such that:

$$p(\tilde{\mathbf{Y}}|\mathbf{X}^{ob}) = \int_{\mathbf{X}} p_\Phi(\tilde{\mathbf{Y}}|\mathbf{X}) p_\theta(\mathbf{X}|\mathbf{X}^{ob}) d\mathbf{X}, \tag{3}$$

where $\mathbf{X}^{ob} \in \mathbb{R}^{T\times N}$, N denotes the number of positions and T denotes the length of time series; $\tilde{\mathbf{Y}} \in \mathbb{R}^{T'\times N}$, T' denotes the predicted length.

We approximate Equation (3) by sampling a finite number of samples and using a summation form, which is formulated as:

$$p(\tilde{\mathbf{Y}}|\mathbf{X}^{ob}) = \sum_{\mathbf{X}} p_\Phi(\tilde{\mathbf{Y}}|\mathbf{X}) p_\theta(\mathbf{X}|\mathbf{X}^{ob}). \tag{4}$$

Here, several datasets $\mathbf{X}$ are generated by well-trained $p_\theta(\mathbf{X}|\mathbf{X}^{ob})$. Based on each dataset, we trained specific $p_\Phi(\tilde{\mathbf{Y}}|\mathbf{X})$. Using Rubin's rules[32], the final Chl_a prediction is obtained by averaging the prediction of specific $p_\Phi(\tilde{\mathbf{Y}}|\mathbf{X})$ based on different $\mathbf{X}$. In this way, our STIMP method not only improves the overall predictive performance through accurate imputation of missing data but also provides confidence intervals to quantify the prediction uncertainties. More details of training $p_\Phi(\tilde{\mathbf{Y}}|\mathbf{X})$ and $p_\theta(\mathbf{X}|\mathbf{X}^{ob})$ can be found in the following two subsections.

Spatiotemporal satellite Chl_a data imputation

We contrived a spatiotemporal denoising diffusion model (STDDM) that we applied to the imputation function, $p_\theta(\mathbf{X}|\mathbf{X}^{ob})$. STDDM is fundamentally a Denoising Diffusion Probabilistic Model (DDPM) specifically designed to approximate the spatiotemporal distribution of Chl_a conditioned on partial observations.

The imputation function conditioned on partial observation. Imputation function intends to leverage deep generative learning to approximate spatiotemporal distribution of Chl_a. The primary challenge in learning the imputation function lies in the complex spatiotemporal distribution of Chl_a, which is challenging to transform from easy-to-sample distribution using neural networks. Inspired by DDPM, we decompose the complex task into L simple tasks. We can manually set a sequence of data distributions perturbed by L levels of signal-to-noise ratio. The objective of each task is to improve the signal-to-noise ratio of the perturbed data distribution from σ_l^2 to σ_{l-1}^2, where $0 \approx \sigma_L < \sigma_{L-1} < \sigma_0$. The final complex spatiotemporal distribution of Chl_a is obtained by the joint distribution $p(\mathbf{X}, \mathbf{X}_{1:L}|\mathbf{X}^{ob})$ defined as a Markov chain with learned Gaussian transitions starting at an easy-to-sample distribution $p(\mathbf{X}_L) = \mathcal{N}(\mathbf{X}_L; \mathbf{0}, \mathbf{I})$:

$$\begin{aligned} p_\theta(\mathbf{X}|\mathbf{X}^{ob}) &= \int p(\mathbf{X}, \mathbf{X}_{1:L}|\mathbf{X}^{ob}) d\mathbf{X}_{1:L} = \int p_\theta(\mathbf{X}|\mathbf{X}_{1:L}, \mathbf{X}^{ob}) p_\theta(\mathbf{X}_{1:L}|\mathbf{X}^{ob}) d\mathbf{X}_{1:L} \\ &= \int p(\mathbf{X}_L) p_\theta(\mathbf{X}|\mathbf{X}_1, \mathbf{X}^{ob}) \prod_{l=L}^{2} p_\theta(\mathbf{X}_{l-1}|\mathbf{X}_l, \mathbf{X}^{ob}) d\mathbf{X}_{1:L}. \end{aligned} \tag{5}$$

In contrast to the noise predictor in DDPM, STDDM is an original data predictor which is utilized in the learned Gaussian transitions $p_\theta(\mathbf{X}_{l-1}|\mathbf{X}_l, \mathbf{X}^{ob})$. Because a time series of Chl_a data usually contain irregular noisy components, estimating the noise is challenging. Note that we can incorporate $p_\theta(\mathbf{X}|\mathbf{X}_1, \mathbf{X}^{ob})$ into $p_\theta(\mathbf{X}_{l-1}|\mathbf{X}_l, \mathbf{X}^{ob})$ and $\mathbf{X}$ can be regarded as clean samples with a signal-to-noise ratio σ_0. Specifically, we define the signal-to-noise ratio $\sigma_l^2 = \bar{\alpha}_l/(1-\bar{\alpha}_l)$, where $\bar{\alpha}_l = \prod_{i=1}^{l} \alpha_i$ and $[\alpha_0, \cdots, \alpha_L]$ is a manually defined noise schedule in which each element regulates the levels of added Gaussian noise to $\mathbf{X}_{l-1}$. Then, the objective of the learned Gaussian transition in the l-th task is formulated as:

$$p_\theta(\mathbf{X}_{l-1}|\mathbf{X}_l, \mathbf{X}^{ob}) = \mathcal{N}(\mathbf{X}_{l-1}; \frac{(1-\bar{\alpha}_{l-1})\sqrt{\alpha_l}}{1-\bar{\alpha}_l}\mathbf{X}_l + \frac{(1-\alpha_l)\sqrt{\bar{\alpha}_{l-1}}}{1-\bar{\alpha}_l}\mathbf{X}, \frac{(1-\alpha_l)\sqrt{\bar{\alpha}_{l-1}}}{1-\bar{\alpha}_l}\mathbf{I}). \tag{6}$$

Unfortunately, the original spatiotemporal distribution of Chl_a, $\mathbf{X}$, is unknown for $p_\theta(\mathbf{X}_{l-1}|\mathbf{X}_l, \mathbf{X}^{ob})$. Hence, we contrive STDDM, an original data predictor, to approximate $\mathbf{X}$ with partial observation $\mathbf{X}^{ob}$ and a perturbed Chl_a distribution, $\mathbf{X}_l$, that is, $p(\mathbf{X}|\mathbf{X}_l, \mathbf{X}^{ob})$. STDMM is learned by the following objective function:

$$\min_\theta \mathbb{E}_{\mathbf{X}_0} \mathbb{E}_{\mathcal{N}(\mathbf{X}_l; \sqrt{\bar{\alpha}_l}\mathbf{X}_0, (1-\bar{\alpha}_l)\mathbf{I})}(||STDDM_\theta(\mathbf{X}^{ob}, \mathbf{X}_l) - \mathbf{X}||^2), \tag{7}$$

In practice, however, three challenges remain in learning the STDMM. First, STDMM cannot differentiate between different subtasks using only $\mathbf{X}^{ob}$ and $\mathbf{X}_l$ as input, leading to difficulties in optimization. To address this issue, we incorporate the task number l as the input to the STDMM. Second, due to we select N locations belong to the ocean as the $\mathbf{X}^{ob}$, STDMM cannot obtain location information, resulting in an inability to obtain spatial heterogeneity and relationships. We construct the spatial graph $\mathbf{G} \in \mathbb{R}^{N\times N}$ to maintain spatial relationships of these locations. By default, if latitude and longitude difference between location i and j are both less than 0.05°, we set $G_{ij} = 1$. By doing so, STIMP can extract information from adjacent locations during the information aggregation of Heterogeneous Spatial Graph Neural Network (HSGNN). Third, we cannot obtain the true values for unobserved Chl_a as a supervised learning signal. As a compromise, we treat the observed data as $\mathbf{X}$, and further mask part of $\mathbf{X}$ with $\mathbf{M}$ to obtain partial observations $\mathbf{X}^{ob} \leftarrow (1 - \mathbf{M}) \odot \mathbf{X}$, where $\odot$ denotes the element-wise product. Because we focus on the performance of reconstructing the unobserved value, the corresponding objective function is given by:

$$\min_\theta \mathbb{E}_{\mathbf{X}} \mathbb{E}_{\mathcal{N}(\mathbf{X}_l; \sqrt{\bar{\alpha}_l}\mathbf{X}, (1-\bar{\alpha}_l)\mathbf{I})}(||\mathbf{M} \odot (STDDM_\theta(\mathbf{X}^{ob}, \mathbf{X}_l, l, \mathbf{G}) - \mathbf{X})||^2), \tag{8}$$

The training process of STDDM is summarized in Algorithm 1.

Algorithm 1. Training Process of STIMP for imputation.

Input : A sample of training data, $\mathbf{X} \in \mathbb{R}^{T\times N}$; Spatial Graph, $\mathbf{G} \in \mathbb{R}^{N\times N}$; the number of denoising steps, L; the number of training iteration, M.

Output : Optimized STDDM

1. **for** $s = 1$ *to* M **do**
2. Get random mask $\mathbf{M}$
3. $\mathbf{X}^{ob} \leftarrow (1 - \mathbf{M}) \odot \mathbf{X}$
4. Sample $l \sim Uniform(1, 2, \cdots, L), \epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
5. Get perturbed value $\mathbf{X}_l \leftarrow \sqrt{\bar{\alpha}_l}\mathbf{X}^{ob} + \sqrt{1-\bar{\alpha}_l}\epsilon$.
6. Estimate missing value with $STDDM_\theta(\mathbf{X}^{ob}, \mathbf{X}_l, l, \mathbf{G})$
7. Update the gradient by $\nabla ||\mathbf{M} \odot (STDDM_\theta(\mathbf{X}^{ob}, \mathbf{X}_l, l, \mathbf{G}) - \mathbf{X}_0)||^2$
8. **end**

Reconstructing the missing values with the imputation function

Now we show how we generate missing Chl_a by leveraging the learned imputation function. During imputation, we reconstruct the missing

values of satellite Chl_a data using the partial observation. Assuming that the STDDM training was good enough to accurately simulate $p(\mathbf{X}|\mathbf{X}_l, \mathbf{X}^{ob})$, the final complex spatiotemporal distribution of Chl_a is obtained by progressively reducing the signal-to-noise ratio. Starting at samples $\mathbf{X}_L$ from the Gaussian distribution, the signal-to-noise ratio can be reduced by getting samples from the following learned Gaussian transition, replacing $\mathbf{X}$ with $STDDM(\mathbf{X}^{ob}, \mathbf{X}_l, l, \mathbf{G})$ in Eq. (6):

$$\mathbf{X}_{l-1} \leftarrow \frac{(1-\bar{\alpha}_{l-1})\sqrt{\alpha_l}}{1-\bar{\alpha}_l}\mathbf{X}_l + \frac{(1-\alpha_l)\sqrt{\bar{\alpha}_{l-1}}}{1-\bar{\alpha}_l} STDDM (\mathbf{X}^{ob}, \mathbf{X}_l, l, \mathbf{G}) + \frac{(1-\alpha_l)\sqrt{\bar{\alpha}_{l-1}}}{1-\bar{\alpha}_l}\epsilon \tag{9}$$

where $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. We summarized the imputation in Algorithm 2.

Spatiotemporal Chl_a prediction

We predict the most probable length-T' sequence of Chl_a $\tilde{\mathbf{Y}}$ with $p_\Phi(\tilde{\mathbf{Y}}|\mathbf{X})$. To achieve this goal, the data distributions in space and time are important and warrant careful consideration. Fortunately, we must obtain a complete spatiotemporal distribution of Chl_a $\mathbf{X}$ is obtained by imputing the Chl_a spatiotemporal satellite data, denoted by $p_\theta(\mathbf{X}|\mathbf{X}^{ob})$ in Eq. (3). We meticulously designed the network architecture to effectively capture spatial correlations and temporal dynamics.

First, we leverage a value embedding network with a 1×1 kernal convolutional layer to encode the information of historical observation of Chl_a, $\mathbf{X}$:

$$\mathbf{H} = Conv(\mathbf{X}). \tag{10}$$

Then, STIMP separately captures the spatial and temporal relationship of the input $\mathbf{H}$ separately. First, the temporal features $\mathbf{H}^{temp}$ are learned by a temporal dependency learning module TLT:

$$\mathbf{H}^{temp} = TLT(\mathbf{H}), \tag{11}$$

Then, the spatial relationships and heterogeneity from the spatial graph, $\mathbf{G}$, are further mined by HSGNN:

$$\mathbf{H}^{spa} = HSGNN(\mathbf{H}^{temp}, \mathbf{G}). \tag{12}$$

Eventually, the predicted Chl_a is obtained using a two-layer fully connected neural network:

$$\tilde{\mathbf{Y}} = F_2^{pre}(SiLU(F_1^{pre}(\mathbf{H}^{spa}))), \tag{13}$$

where $SiLU$ is activation function.

The objective function of training $p_\Phi(\tilde{\mathbf{Y}}|\mathbf{X})$ is given by:

$$\min_\Phi \mathbb{E}_{\mathbf{X}, \tilde{\mathbf{Y}}\sim p_\Phi(\tilde{\mathbf{Y}}|\mathbf{X})}(||\mathbf{Y} - \tilde{\mathbf{Y}}||^2). \tag{14}$$

Algorithm 2. STIMP Imputation

Input : A sample of incomplete observed data, $\mathbf{X}^{ob} \in \mathbb{R}^{T\times N}$, with observation indicator, $\mathbf{M}$; Spatial Graph, $\mathbf{G} \in \mathbb{R}^{N\times N}$; the number of diffusion steps, L; and the optimized STDDM.

Output : Imputed data $\mathbf{X}$

1. Set $\mathbf{X}_L \sim \mathcal{N}(0, I)$
2. **For** $l = L$ *to* 1 **do**
3. $\quad \mathbf{X}_{l-1} \leftarrow \frac{(1-\bar{\alpha}_{l-1})\sqrt{\alpha_l}}{1-\bar{\alpha}_l}\mathbf{X}_l + \frac{(1-\alpha_l)\sqrt{\bar{\alpha}_{l-1}}}{1-\bar{\alpha}_l} STDDM(\mathbf{X}^{ob}, \mathbf{X}_l, l, \mathbf{G})$
4. $\quad \epsilon \sim \mathcal{N}(0, I)$
5. $\quad \mathbf{X}_{l-1} \leftarrow \mathbf{X}_{l-1} + \frac{(1-\alpha_l)\sqrt{\bar{\alpha}_{l-1}}}{1-\bar{\alpha}_l}\epsilon$
6. **end**
7. $\mathbf{X} \leftarrow \frac{(1-\bar{\alpha}_0)\sqrt{\alpha_1}}{1-\bar{\alpha}_1}\mathbf{X}_1 + \frac{(1-\alpha_1)\sqrt{\bar{\alpha}_0}}{1-\bar{\alpha}_1} STDDM(\mathbf{X}^{ob}, \mathbf{X}_1, 1, \mathbf{G})$
8. $\mathbf{X} \leftarrow \mathbf{M} \odot \mathbf{X}^{ob} + (1 - \mathbf{M}) \odot \mathbf{X}$

Network structures of STDDM

STDDM does not limit the network architecture. To make the model for suitable for spatiotemporal imputation, as shown in Supplementary Fig. S16c, we leveraged TLT and HSGNN to handle temporal and spatial relationships, respectively. Specifically, STDDM concatenates the time series of partial observation, $\mathbf{X}^{ob}$; perturbed values, $\mathbf{X}_l$; diffusion step, l; and spatial graph, $\mathbf{G}$ as the input. STDDM utilizes a value embedding network with a 1×1 kernel convolutional layer to encode the information and outputs a d-dimensional latent representations:

$$\mathbf{H} = Conv(\mathbf{X}_l||\mathbf{X}^{ob}), \tag{15}$$

where $\mathbf{H} \in \mathbb{R}^{L\times N\times d}$. The diffusion step, l, is specified by adding the Transformer sinusoidal position embedding (PE)[55] to $\mathbf{H}$:

$$\mathbf{H}^{in} = \mathbf{H} + DE(PE(l)), \tag{16}$$

where DE project consists of two fully connected layers and a Sigmoid Linear Unit (SiLU) function as the activation layer, and we broadcast this embedding vector over timesteps and positions $L\times N$.

To comprehensively capture the global spatiotemporal and geographic relationship of input $\mathbf{H}^{in}$, the temporal features, $\mathbf{H}^{temp}$, is learned using the temporal dependency learning module TLT, and then the temporal features are aggregated through a spatial dependency learning module HSGNN:

$$\mathbf{H}^{temp} = TLT(\mathbf{H}^{in}) \quad \mathbf{H}^{spa} = GN(\mathbf{H}^{temp} + HSGNN(\mathbf{H}^{temp}, \mathbf{G})), \tag{17}$$

where GN means Group Normalization[56] to stabilize the training. More of the original information $\mathbf{H}^{temp}$ is preserved for $\mathbf{H}^{spa}$.

$\mathbf{H}^{out}$ is obtained by leveraging the gated activation unit in CSDI[37], which is formalized as:

$$\mathbf{H}^{out} = sigmoid(Conv(\mathbf{H}^{spa}))*tanh(Conv(\mathbf{H}^{spa})), \tag{18}$$

where $Conv$ is a convolutional layer with a 1×1 kernel. Finally, the predicted imputation target is obtained with

$$\mathbf{X} = \mathbf{F}^{imp}(\mathbf{H}^{out}), \tag{19}$$

where $\mathbf{F}^{imp}$ is a fully connected layer.

Network structures of TLT

TLT is dedicated to effectively capturing the temporal relationships of Chl_a. In TLT, each element in the time series will compute attention with all other elements, thereby preserving and propagating essential information throughout the extended sequence[55]. The long- and short-term temporal dependencies captured by TLT significantly enhance the accuracy of large-scale spatiotemporal Chl_a imputations and predictions.

Specifically, TLT takes the representations of Chl_a from a time series $\mathbf{H}^{in} \in \mathbb{R}^{L\times N\times d}$ as input, yielding the enriched temporal representation $\mathbf{H}^{temp} \in \mathbb{R}^{L\times N\times d}$ by encoding the time dependency in $\mathbf{H}^{in}$. Two key modules in TLT are the Multi-head Self-Attention (MSA) and Position-wise Feed-Forward (PFF). MSA is mainly in charge of capturing temporal relationships and PFF endows the model with nonlinearity. We mathematically represent the overall procedure of TLT as a function as follows:

$$\mathbf{H}^{temp} = LN(PFF(LN(MSA(\mathbf{H}^{in})))), \tag{20}$$

where LN is the standard normalization layer. The key component in MSA is self-attention (SA) along the temporal dimension L, which is

Table 1 | Computational demands of STIMP

Methods	FLOPs (G)	MACs (G)	Parameters (K)	GPU memory-usage (M)
STIMP (imputation)	29.22	13.34	136.77	1450
STIMP (prediction)	1.04	0.48	3.7	762

defined as follows,

$$SA(Q,K,V) = softmax(\frac{QK^T}{\sqrt{d}})V, \tag{21}$$

where $Q = \mathbf{H}^{in}W^Q$; $K = \mathbf{H}^{in}W^K$; $V = \mathbf{H}^{in}W^V$ represent the queries, the keys, and the values, respectively, which are converted from $\mathbf{H}^{in}$, W^Q, W^K, and $W^V \in \mathbb{R}^{d\times d}$ are learnable projection parameters. Essentially, MSA computes the dot product of Q_i and K_j to generate a weight that measures the relevance between Chl_a at times i and j.

MSA utilizes several SA layers: $MSA(\mathbf{H}^{in}) = Concat(h_1, h_2, \cdots, h_n)W^O + \mathbf{H}^{in}$ where $h_i = SA(Q, K, V)$. The residual is introduced to preserve more original Chl_a information. The added PFF networks consist of a fully connected layer and a Rectified Linear Unit(ReLU) function as and activation layer.

Network structures of HSGNN

To leverage the spatial dependency of Chl_a, we contrived HSGNN to obtain the enriched spatiotemporal representation, $\mathbf{H}^{spa} \in \mathbb{R}^{L\times N\times d}$ by encoding the spatial dependency in $\mathbf{H}^{temp}$ with the assistance of $\mathbf{G}$:

$$\mathbf{H}^{spa} = \mathbf{F}^{spa}(\mathbf{W}\cdot\mathbf{H}^{temp}\cdot\mathbf{G}). \tag{22}$$

As shown in Supplementary Fig. S16d (a), the time series exhibits spatial heterogeneity, with significantly different means and variances of Chl_a at different locations. Hence, the same parameter $\mathbf{W}$ for different locations does not meet the requirements.

However, it is impractical to design a separate set of parameters for each different position. For the spatial parameter space, $\theta^{spa} \in \mathbb{R}^{N\times d\times d}$, directly optimizing this parameter spaces brings huge computational burden and data demand, especially when N is large. Instead, we provide a simple yet effective way to maintain a small parameters pool, $\mathbf{P} \in \mathbb{R}^{k\times d\times d}$, which contains k parameter prototypes. k is defined to be much smaller than N. We can generate a spatial embedding $\mathbf{Q} \in \mathbb{R}^{k\times N}$ as a query for all positions, and we can obtain a location-specific parameter using a weighted combination of the parameter prototypes:

$$\mathbf{W}^{spa} = \mathbf{Q}^T\cdot\mathbf{P}. \tag{23}$$

Because spatial heterogeneity is mainly reflected in the mean and variance of the Chl_a time series across different positions, we generate a query, $\mathbf{Q}$, for all positions according to each location, $L \in \mathbb{R}^{1\times N}$, mean, $\mu \in \mathbb{R}^{1\times N}$, and variance, $\varsigma \in \mathbb{R}^{1\times N}$:

$$\mathbf{Q} = \mathbf{H}^T\begin{bmatrix} L \\ \mu \\ \varsigma \end{bmatrix}, \tag{24}$$

where $\mathbf{H} \in \mathbb{R}^{3\times k}$. In doing so, the learned parameter reduces from $\theta^{spa} \in \mathbb{R}^{N\times d\times d}$ to $\mathbf{H} \in \mathbb{R}^{3\times k}$ and $\mathbf{P} \in \mathbb{R}^{k\times d\times d}$.

Then, we leverage $\mathbf{W}^{spa} \in \mathbb{R}^{N\times d\times d}$ to replace the original parameters $\mathbf{W} \in \mathbb{R}^{d\times d}$ in Equation (22), introducing the spatial heterogeneity using location specific parameters:

$$\mathbf{H}^{spa} = \mathbf{F}^{spa}(\mathbf{W}^{spa}\cdot\mathbf{H}^{temp}\cdot\mathbf{G}), \tag{25}$$

where $\mathbf{F}^{spa}$ is a fully connected network. Now, each position obtains information from an adjacent position through $\mathbf{H}^{temp}\cdot\mathbf{G}$. Meanwhile, location-specific parameters, $\mathbf{W}^{spa}$, are utilized to embed the spatial heterogeneity into $\mathbf{H}^{spa}$. As shown in Supplementary Fig. S16d, we leveraged the learned spatial heterogeneous embedding $\mathbf{Q}$ to partition the Pearl River Estuary. The results of the partition were generally consistent with the the isobaths at 30 m and 50 m below sea level, showing that STIMP indeed captured the spatial heterogeneity of different locations in the Pearl River Estuary.

More training details

STIMP employs Adam for stochastic optimization during model training. By default, the number of training epochs in STIMP is 500 and 200 for imputation and prediction respectively, with learning rate $lr = 0.0001$ and a weight decay parameter $\lambda = 0.00001$.

Computational demands

We computed the computational demands of STIMP, including FLOPs (floating-point operations), MACs (multiply-add operations) and Parameters, with Calflops (https://github.com/MrYxJ/calculate-flops.pytorch.git) on Pearl River Estuary. Moreover, we recorded the GPU memory usage of STIMP. Note that we set the batch size of training data is equal to 1 when calculating FLOPs, MACs and GPU memory, which means the minimum computational requirements for running STIMP. The hardware requirements analysis in Table 1 reveals that STIMP maintains low GPU memory demands (≥1.5 GB), compatible with obsolete consumer hardware, including NVIDIA's GTX 1050 (2016 release). This highlights the framework's practical deployability even with limited computational resources.

Data availability

The MODIS Chl_a data used in this study have been publicly available in the Moderate Resolution Imaging Spectroradiometer (MODIS) Aqua projects[33] (https://search.earthdata.nasa.gov/search?q=10.5067/AQUA/MODIS/L3M/CHL/2022). The processed MODIS Chl_a data is available on Zenodo at https://doi.org/10.5281/zenodo.14638405. Source data are provided with this paper. We also uploaded all scripts and materials to reproduce all the analyses at https://stimp-tutorials.readthedocs.io/en/latest/. Source data are provided with this paper.

Code availability

The STIMP software package and source code are available on GitHub at https://github.com/YangLabHKUST/STIMP.git. We also uploaded all scripts and materials to reproduce all the analyses at the same website.

References

1. Costanza, R. et al. The value of the world's ecosystem services and natural capital. *Nature* **387**, 253–260 (1997).
2. Pauly, D. et al. Towards sustainability in world fisheries. *Nature* **418**, 689–695 (2002).
3. Gray, J. S. Marine biodiversity: patterns, threats and conservation needs. *Biodivers. Conserv.* **6**, 153–175 (1997).
4. Dai, Y. et al. Coastal phytoplankton blooms expand and intensify in the 21st century. *Nature* **615**, 280–284 (2023).
5. Gruber, N., Boyd, P. W., Frölicher, T. L. & Vogt, M. Biogeochemical extremes and compound events in the ocean. *Nature* **600**, 395–407 (2021).

6. Breitburg, D. et al. Declining oxygen in the global ocean and coastal waters. *Science* **359**, eaam7240 (2018).
7. Trottet, A., George, C., Drillet, G. & Lauro, F. M. Aquaculture in coastal urbanized areas: a comparative review of the challenges posed by harmful algal blooms. *Crit. Rev. Environ. Sci. Technol.* **52**, 2888–2929 (2022).
8. Zhu, X., Guo, H., Huang, J. J., Tian, S. & Zhang, Z. A hybrid decomposition and machine learning model for forecasting chlorophyll-a and total nitrogen concentration in coastal waters. *J. Hydrol.* **619**, 129207 (2023).
9. Lapointe, B. E., O'Connell, J. D. & Garrett, G. S. Nutrient couplings between on-site sewage disposal systems, groundwaters, and nearshore surface waters of the Florida keys. *Biogeochemistry* **10**, 289–307 (1990).
10. Heffernan, J. B. et al. Hydrologic and biotic influences on nitrate removal in a subtropical spring-fed river. *Limnol. Oceanogr.* **55**, 249–263 (2010).
11. Commission, E. Commission decision (EU) 2017/848 of 17 may 2017 laying down criteria and methodological standards on good environmental status of marine waters and specifications and standardised methods for monitoring and assessment, and repealing decision 2010/477/EU. *J. Eur. Union* **125**, 43–74 (2017).
12. Colella, S., Falcini, F., Rinaldi, E., Sammartino, M. & Santoleri, R. Mediterranean ocean colour chlorophyll trends. *PloS One* **11**, e0155756 (2016).
13. Chassignet, E. P. et al. Us godae: global ocean prediction with the hybrid coordinate ocean model (HYCOM). *Oceanography* **22**, 64–75 (2009).
14. Chassignet, E. P. et al. The HYCOM (Hybrid Coordinate Ocean Model) data assimilative system. *J. Mar. Syst.* **65**, 60–83 (2007).
15. Gan, J., Liu, Z. & Hui, C. R. A three-layer alternating spinning circulation in the South China Sea. *J. Phys. Oceanogr.* **46**, 2309–2315 (2016).
16. MATEAR, R. Parameter optimization and analysis of ecosystem models using simulated annealing: a case study at station P. *J. Mar. Res.* **53**, 571–607 (1995).
17. Allen, M. R., Kettleborough, J. & Stainforth, D. Model error in weather and climate forecasting. In *ECMWF Predictability of Weather and Climate Seminar*, 279–304 (European Centre for Medium Range Weather Forecasts, 2002).
18. Fontana, C., Grenz, C., Pinazo, C., Marsaleix, P. & Diaz, F. Assimilation of seawifs chlorophyll data into a 3D-coupled physical–biogeochemical model applied to a freshwater-influenced coastal zone. *Continental Shelf Res.* **29**, 1397–1409 (2009).
19. Gregg, W. W. Assimilation of SeaWiFS ocean chlorophyll data into a three-dimensional global ocean model. *J. Mar. Syst.* **69**, 205–225 (2008).
20. Sanchez-Gonzalez, A. et al. Learning to simulate complex physics with graph networks. In *International conference on machine learning*, 8459–8468 (2020).
21. Corredor, J. E. *Numerical Models for Operational Ocean Observing*, 113–123 (Springer International Publishing, Cham, 2018). https://doi.org/10.1007/978-3-319-78352-9_6.
22. Wu, G. & Xu, Z. Prediction of algal blooming using EFDC model: Case study in the Daxiang Lake. *Ecol. Model.* **222**, 1245–1252 (2011).
23. Chen, Q., Guan, T., Yun, L., Li, R. & Recknagel, F. Online forecasting chlorophyll a concentrations by an auto-regressive integrated moving average model: feasibilities and potentials. *Harmful Algae* **43**, 58–65 (2015).
24. Fathi, M. M., Awadallah, A. G., Abdelbaki, A. M. & Haggag, M. A new Budyko framework extension using time series SARIMAX model. *J. Hydrol.* **570**, 827–838 (2019).
25. Ai, H., Zhang, K., Sun, J. & Zhang, H. Short-term Lake Erie algal bloom prediction by classification and regression models. *Water Res.* **232**, 119710 (2023).
26. Wang, L. et al. Employing hybrid deep learning for near-real-time forecasts of sensor-based algal parameters in a microcystis bloom-dominated lake. *Sci. total Environ.* **922**, 171009 (2024).
27. Sun, X. et al. Enhanced forecasting of chlorophyll-a concentration in coastal waters through integration of Fourier analysis and transformer networks. *Water Res.* **263**, 122160 (2024).
28. Chen, J., Zhang, M., Cui, T. & Wen, Z. A review of some important technical problems in respect of satellite remote sensing of chlorophyll-a concentration in coastal waters. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **6**, 2275–2289 (2013).
29. Johnson, K. S. & Bif, M. B. Constraint on net primary productivity of the global ocean by Argo oxygen measurements. *Nat. Geosci.* **14**, 769–774 (2021).
30. Park, J.-E., Park, K.-A., Kang, C.-K. & Park, Y.-J. Short-term response of chlorophyll-a concentration to change in sea surface wind field over mesoscale eddy. *Estuaries Coasts* **43**, 646–660 (2020).
31. Mercury, M. et al. Global cloud cover for assessment of optical satellite observation opportunities: a HyspIRI case study. *Remote Sens. Environ.* **126**, 62–71 (2012).
32. Rubin, D. B. *Multiple imputation for nonresponse in surveys*, vol. 81 (John Wiley & Sons, 2004).
33. Esaias, W. E. et al. An overview of MODIS capabilities for ocean science observations. *IEEE Trans. Geosci. Remote Sens.* **36**, 1250–1265 (1998).
34. Alvera-Azcárate, A., Barth, A., Beckers, J.-M. & Weisberg, R. H. Multivariate reconstruction of missing data in sea surface temperature, chlorophyll, and wind satellite fields. *J. Geophys. Res. Oceans* **112** (2007).
35. Ma, C., Zhao, J., Ai, B. & Sun, S. Two-decade variability of sea surface temperature and chlorophyll-a in the northern South China Sea as revealed by reconstructed cloud-free satellite data. *IEEE Trans. Geosci. Remote Sens.* **59**, 9033–9046 (2021).
36. Yu, H.-F., Rao, N. & Dhillon, I. S. Temporal regularized matrix factorization for high-dimensional time series prediction. *Advances in Neural Information Processing Systems 29* (NIPS, 2016).
37. Tashiro, Y., Song, J., Song, Y. & Ermon, S. Csdi: Conditional score-based diffusion models for probabilistic time series imputation. *Adv. Neural Inf. Process. Syst.* **34**, 24804–24816 (2021).
38. Nie, T., Qin, G., Ma, W., Mei, Y. & Sun, J. Imputeformer: Low rankness-induced transformers for generalizable spatiotemporal imputation. In *Proc. of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2260–2271 (2024).
39. Yun, T., Jung, H. & Son, J. Imputation as inpainting: Diffusion models for spatiotemporal data imputation (2023).
40. He, K. et al. Masked autoencoders are scalable vision learners. In *Proc. of the IEEE/CVF conference on computer vision and pattern recognition*, 16000–16009 (2022).
41. Chen, T. & Guestrin, C. Xgboost: A scalable tree boosting system. In *Proc. of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785–794 (2016).
42. Zhang, Y. & Yan, J. Crossformer: Transformer utilizing cross-dimension dependency for multivariate time series forecasting. In *The eleventh international conference on learning representations* (2023).
43. Ekambaram, V., Jati, A., Nguyen, N., Sinthong, P. & Kalagnanam, J. Tsmixer: Lightweight MLP-mixer model for multivariate time series forecasting. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 459–469 (2023).
44. Liu, Y. et al. itransformer: Inverted transformers are effective for time series forecasting. In *Proceedings of the Twelfth International Conference on Learning Representations*, 1–25 (2024).

45. Wu, Z. et al. Connecting the dots: Multivariate time series forecasting with graph neural networks. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, 753–763 (2020).
46. Wang, Y. et al. Predrnn: A recurrent neural network for spatiotemporal predictive learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **45**, 2208–2225 (2022).
47. Jayawardena, A. W. & Lai, F. Time series analysis of water quality data in Pearl River, China. *J. Environ. Eng.* **115**, 590–607 (1989).
48. Chen, C. et al. Remotely sensed assessment of water quality levels in the Pearl River Estuary, China. *Mar. Pollut. Bull.* **54**, 1267–1272 (2007).
49. Kumpf, H., Steidinger, K. & Sherman, K. *The Gulf of Mexico Large Marine Ecosystem: Assessment, Sustainability, and Management*, 704 (Blackwell Science, Inc., Malden, MA, USA, 1999).
50. Spies, R. B., Senner, S. & Robbins, C. S. An overview of the northern Gulf of Mexico ecosystem. *Gulf Mex. Sci.* **33**, 9 (2016).
51. Ho, J., Jain, A. & Abbeel, P. Denoising diffusion probabilistic models. *Adv. neural Inf. Process. Syst.* **33**, 6840–6851 (2020).
52. Liu, X. & Wang, M. Gap filling of missing data for VIIRS global ocean color products using the dineof method. *IEEE Trans. Geosci. Remote Sens.* **56**, 4464–4476 (2018).
53. Wu, H., Xu, J., Wang, J. & Long, M. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting. *Adv. neural Inf. Process. Syst.* **34**, 22419–22430 (2021).
54. Stitsyuk, A. & Choi, J. xpatch: Dual-stream time series forecasting with exponential seasonal-trend decomposition. In *Proc. of the AAAI Conference on Artificial Intelligence*, vol. 39, 20601–20609 (2025).
55. Vaswani, A. Attention is all you need. *Advances in Neural Information Processing Systems* (2017).
56. Wu, Y. & He, K. Group normalization. In *Proc. of the European conference on computer vision (ECCV)*, 3–19 (2018).

Acknowledgements

The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project Reference Number: AoE/P-601/23-N); we thank the support of the Center for Ocean Research in Hong Kong and Macau (CORE), CORE is a joint research center for ocean research between Laoshan Laboratory and HKUST; we thank the Moderate Resolution Imaging Spectroradiometer (MODIS) Aqua projects for providing 8-d Chl_a dataset with spatial resolution of 4 km (https://search.earthdata.nasa.gov/search?q=10.5067/AQUA/MODIS/L3M/CHL/2022).

Author contributions

Fan Z., C.Y. and J.G. conceived the study and designed the STIMP algorithm. Fan Z. implemented the STIMP algorithm. Fan Z. and Fa Z. evaluated the baseline methods. Fan Z. and H.K. processed the data and interpreted the results. Fan Z., H.K., C.Y., and J.G. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-62901-9.

Correspondence and requests for materials should be addressed to Can Yang or Jianping Gan.

Peer review information *Nature Communications* thanks Jesús Morón-López, Xiaofeng Li, and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.