

PAPER • OPEN ACCESS

Univariate deep learning framework for short-term SST forecasting at high spatio-temporal scales

To cite this article: Jagdish Prajapati *et al* 2025 *Environ. Res. Commun.* **7** 075002

View the [article online](#) for updates and enhancements.

You may also like

- [The making of rights of nature: nine patterns in a decade of empirical research on social-ecological drivers and actors](#)
Ilkhom Soliev, Frauke Pirscher and Marie Schreiber
- [Charting plastic shores: environmental NGOs bridging historical data gaps for coastal litter management in Cyprus](#)
Ioannis Savva, Maria Kola, Natasa Ioannou et al.
- [The bright side of executive financial background: a perspective based on corporate ESG performance](#)
Xinxin Yu, Yangguang Hu and Shuran Feng



www.hidenanalytical.com
info@hiden.co.uk

HIDEN ANALYTICAL

Instruments for Advanced Science

Mass spectrometers for vacuum, gas, plasma and surface science



Dissolved Species Analysis

Hiden offers MIMS capabilities in the form of a benchtop HPR-40 DSA system for laboratory-based research and the portable case mounted pQA for applications that favour in-situ measurements in the field. Both are supplied with a choice of membrane material and user-changeable sample inlets.



Gas Analysis

The QGA and HPR-20 series gas analysers are versatile tools designed for a broad spectrum of environmental applications, including pollution monitoring, biogas analysis, and sustainable energy research.



PAPER

OPEN ACCESS

RECEIVED
13 March 2025

REVISED
4 June 2025

ACCEPTED FOR PUBLICATION
24 June 2025

PUBLISHED
2 July 2025

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Univariate deep learning framework for short-term SST forecasting at high spatio-temporal scales

Jagdish Prajapati¹ , Balaji Baduru^{1,2,3} , Athul C R^{1,4} , Biswamoy Paul^{1,5} , Vinod Daiya¹ and Arya Paul^{1,*}

¹ Indian National Centre for Ocean Information Services (INCOIS), Ministry of Earth Sciences (MoES), Hyderabad, India

² Indian Institute of Tropical Meteorology (IITM), Ministry of Earth Sciences (MoES), Pune, India

³ Department of Marine Geology, Mangalore University, Mangalagangothri, Karnataka, India

⁴ KUFOS-INCOIS Joint Research Centre, Faculty of Ocean Science and Technology, Kerala University of Fisheries and Ocean Studies, Kochi, India

⁵ Centre for Earth, Ocean and Atmospheric Sciences, University of Hyderabad, Hyderabad, India

* Author to whom any correspondence should be addressed.

E-mail: aryapaul@incois.gov.in

Keywords: data-driven forecasting, deep-learning, long short term memory, sea surface temperature, variational mode decomposition

Abstract

We present a univariate hybrid machine learning framework to predict daily high resolution Sea Surface Temperature (SST) near the Gulf of Kutch region at a resolution of ~ 5.5 km. The hybrid model integrates Intrinsic Mode Functions (IMF) derived from variational mode decomposition with a Long Short-Term Memory (LSTM) network to augment predictive skill. The predicted SST demonstrates impressive performance up to lead times of 7 days. Using statistical metrics like Kullback-Leibler divergence and mutual information, we show that the SST predicted from the hybrid model with a lead time of 3 days decisively outperforms the high-resolution GLORYS SST reanalysis let alone forecast skills of a data-assimilated dynamical model. Using conditional probability, we show that the SST forecasts from the hybrid model are quite reliable over the entire range of SST observations in the study domain. In contrast, the reliability of GLORYS falters in the lower range of SST observations. Also, the hybrid model excels in capturing fine-scale SST features, such as SST fronts, and detecting Marine Heatwaves (MHWs) up to 3 days in advance. These capabilities hold significant applications for Potential Fishing Zone identification and coral bleaching alerts. The hybrid model framework is also adept at forecasting location specific high frequency (3 hourly) SST with a lead time of a day.

1. Introduction

Sea Surface Temperature (SST), which is perhaps the most widely used and ubiquitous ocean state variable, not only plays a vital role in the dynamics of the ocean but also strongly influences atmospheric analysis and weather forecasts. Accurate short-term prediction of SST can serve a suite of applications. For example, short-term accurate prediction of SST over a basin can improve the regional weather forecast. The predicted SST can be treated as synthetic observations and improve the quality of ocean state forecast through data assimilation. Predicted SST fronts can assist offshore fishermen in identifying the inhabited regions of target fish because different water temperatures attract and concentrate different fish species (IOCCG 2009). Thus, SST plays a crucial role in aquaculture, particularly in the survival of many aquatic species like fish and shrimp. The predicted SST can also help to identify appropriate sites for moving marine cages (Roxanne 2016). Additionally, several studies have shown that hurricanes and coral bleaching can result from abnormally warm water (Liu *et al* 2013, McTaggart-Cowan *et al* 2015, Mohanty *et al* 2021). Impending knowledge of abnormally warm water can aid in issuing marine heatwave forecasts. In short, SST prediction can support various management and adaptation initiatives for marine ecosystems and tourism. From all these considerations and many more, there is a growing demand for spatially comprehensive maps of predicted SST, particularly those with higher accuracy and resolution that can demonstrate an increasing number of applications (Aparna *et al* 2018).

There are primarily two approaches through which spatial maps of SST are predicted. The first approach involves SST predictions through ocean general circulation models like MOM, HYCOM, NEMO, MITgcm, ROMS etc, which are often equipped with data assimilation. These models are formulated based on the complex mechanical and thermodynamic interactions among the ocean state variables. The complexity of the dynamics is further compounded when chemical and biological state variables, which significantly modulates the SST, are introduced into the models. These dynamical models have made giant leaps over the past decades in their efficacy but are still fraught with issues like inaccurate initial conditions, approximate fluxes etc. Another issue that plagues these models is the empirical parameterizations of various processes like mixing, bulk flux parameterizations etc, which casts a strong imprint on the SST evolution. Most of these parameterization schemes have been designed based on expeditions carried out in the Pacific (Large *et al* 1994) and Atlantic Ocean (Halliwell 2004). Consequently, these parameterization schemes do not fare equally well across all basins and across all seasons. The models are also often of lower resolution—particularly the basin scale models—than the high-resolution (~2 km) satellite SST products resulting in reduced representation of the processes. The imprints of all these deficiencies on the SST predictions result in a less than accurate SST prediction leaving scope for improvements. With increasing lead times and no futuristic SST observations to constrain the model from diverging through data assimilation, these errors grow in time during the SST forecast.

The second approach in SST prediction leans towards a data-driven approach. This approach has gained traction in the last couple of decades owing to the availability of SST measurements for long periods. The satellite measurement of SST started in 1964 through the Nimbus-1 satellite at a resolution of ~8.5 km. The advent of Advanced Very High Resolution Radiometer (AVHRR) in the late 70s improved the SST coverage and resolution which was further augmented with the Advanced Microwave Scanning Radiometer (AMSRE) in late 90s that could measure SST even on cloudy conditions albeit at lower resolution. The availability of SST data over prolonged periods facilitated a surge during the last decade or so in the use of various data-driven techniques including machine learning (ML) to predict SST. These data-driven methods have an inherent advantage over the dynamical models. They do not suffer from the same maladies that the dynamical models do. Unlike the dynamical models, the data-driven approaches learn the intricate relationships between the variables on their own and are therefore free from explicit assumptions undertaken to ease the solution to the dynamics of the model. In addition, these data-driven models for SST forecasts are computationally faster and can work in a high-resolution setting—the resolution being dictated by the resolution of the observation networks. For example, the Operational Sea Surface Temperature and Ice Analysis (OSTIA) system generates SST (foundation SST to be specific) maps at a resolution of ~5.5 km (Good *et al* 2020) over entire basins. One of the major pitfalls of data-driven approaches is the inherent difficulty in deciphering the influence of various physical processes on SST thereby contributing little to the physical understanding of the SST behavior.

The deep learning methods for SST prediction can be broadly categorized into two main groups. The first group refers to location-specific prediction (Patil *et al* 2016, Sarkar *et al* 2020, Wolff *et al* 2020, Usharani 2023, and many more), while the other produces a regional map of SST values (Sepp and Schmidhuber 1997, Tripathi *et al* 2006, Aparna *et al* 2018, Yang *et al* 2018, Zheng *et al* 2020, Vytla *et al* 2025, and many more). Our focus in this study is on the regional maps of short-term (~few days) SST predictions. Different variants of neural networks have been used over the last decade to predict short-term SST over a region. Of late, long short term memory (LSTM), which is a kind of recurrent neural network (RNN), or some modified variant of LSTM has gained a lot of attention. For example, a fully connected LSTM (Sepp and Schmidhuber 1997) layer with a convolution layer was used to illustrate SST data's spatial and temporal pattern over the Bohai Sea (Yang *et al* 2018). They have shown a spatial map of the difference between ground truth and prediction over the China Sea, mainly in the range from 0.5 °C to 1.0 °C. Zheng *et al* (2020) implemented a deep learning model by cascading and stacking convolutional layers to predict the SST map for the next time step. Shao *et al* (2021) used multivariate empirical orthogonal function on SST and sea surface height anomaly on the machine learning model (Conv1D LSTM) to predict SST over the South China Sea that exhibits an impressive domain-averaged RMSE of ~0.46 °C with a lead time of 10 days. They have also shown that this model predicts SST reasonably well during extreme events exhibiting an RMSE of 0.334 °C for a lead time of 3 days. This model however does not perform well to reproduce the small-scale features. Wei and Guan (2021) used a 3DConv-LSTM to predict daily SST over the China Sea using different lead times from 1 day to 7 days. This study found a good agreement with OSTIA SST till a prediction horizon of 3 days. The RMSE increases from 0.41 °C to 0.53 °C when the lead time is extended from 3 to 7 days. Choi *et al* (2023) used ERA5 and LSTM to predict SST near the Korean Peninsula. Shuang *et al* (2023) proposed a hybrid deep learning model to predict SST at a resolution of ~11 km over the South China Sea that utilizes a data analysis method called Variational Mode Decomposition (VMD), and a deep learning model called Memory in Memory (MIM; Wang *et al* 2019) which is an improved version of LSTM. The VMD decomposes the time-series of a signal into intrinsic mode functions (IMFs) which represent distinct frequency bands. This model predicts SST with improved accuracy (RMSE = 0.247 °C) for a lead time of 7 days. This model, despite its impressive performance, has not been compared against state-of-the-art reanalyses.

Instead, most of the studies on the short-term prediction of SST compare outcomes against other machine learning models. The VMD-MIM model, owing to its relatively low resolution, has also not been tested for the veracity of the fine scale spatial structures like SST fronts.

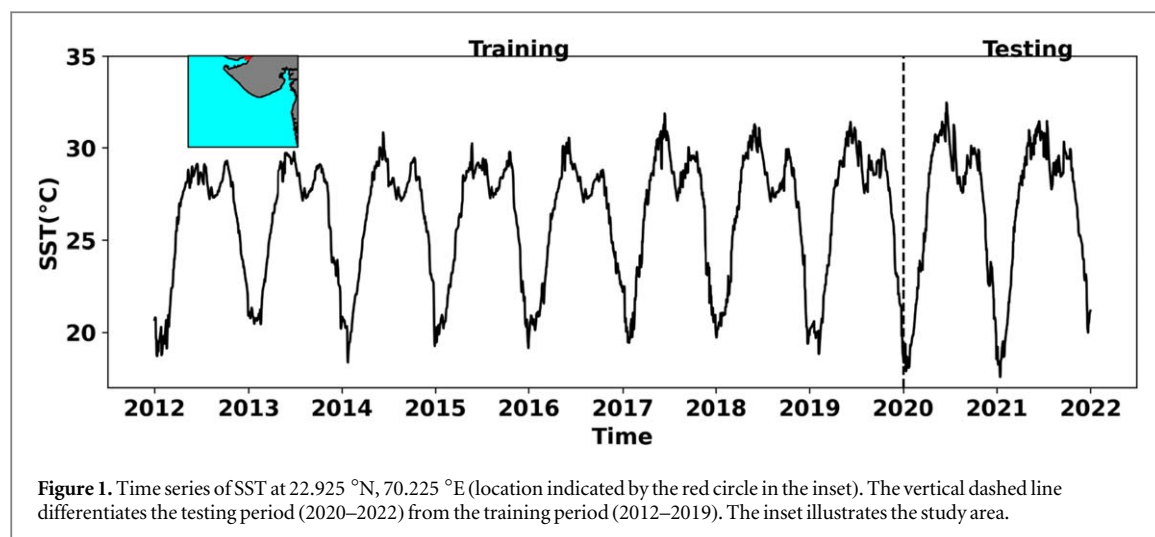
A few ML studies on regional maps of SST predictions have been done in the Indian Ocean as well where our study is based. Using only the past few days of historical daily SST data from the Moderate Resolution Imaging Spectroradiometer (MODIS) at a spatial resolution of 4 km, Aparna *et al* (2018) used Artificial Neural Network (ANN) to predict SST and SST fronts with a lead time of 1 day over a small domain in the northeastern Arabian Sea with reasonable accuracy. Tripathi *et al* (2006) utilized an ANN over a small Indian Ocean region (96 °E–104 °E, 27 °S–35 °S) to predict SST anomalies. A recent study by Vytla *et al* (2025) investigated the performance of various ML techniques for predicting daily SST over three small domains (0.25 million sq km of area each) in the Arabian Sea, Bay of Bengal, and South Indian Ocean regions at a resolution of ~25 km using SST data from Optimum Interpolation Sea Surface Temperature (OISST). The study evaluated various ML models with lead times of 1–7 days. This study identified that the LSTM model fares best when the last 50 days of SST data is provided as input to the ML model. It exhibited a prediction error of ~0.24 °C for a 7-day forecast. In short, there have been several studies across various ocean basins including the Indian Ocean where ML models have been applied for SST predictions with varying degrees of success.

Taking inspiration from Shuang *et al* (2023) and Vytla *et al* (2025), and with an aim to continue the experimentation forward, we ask a set of questions outlined below:

- (1) The reanalysis, derived from the amalgamation of observation and dynamical model, is supposedly the best version of SST. Most existing studies primarily focus on comparing the performance of various machine learning models against each other but do not compare the performance skill of the forecast against reanalyses. Can a hybrid machine learning model for sea surface temperature be developed whose prediction skills outscore the sea surface temperature reanalysis from a state-of-the-art dynamical model equipped with data assimilation? If so, at what lead times? If the hybrid model beats the best available reanalysis over a reasonably long period of analysis, it is implicitly assumed that it shall beat short-term forecasts from dynamical models. A success is a step forward towards operationalization of short-term SST forecasts using machine learning.
- (2) How reliable are the forecasts from the hybrid model? In other words, how often do the forecasts from the hybrid model concur with the observations?
- (3) Aparna *et al* (2018) showed that SST fronts can be predicted with a lead time of 1 day. Can this hybrid model predict the observed sea surface temperature fronts which are fine-scale structures and often misrepresented in state-of-the-art reanalysis with a lead time larger than a day? An extended lead time shall help in issuing Potential Fishing Zone forecasts to the fishermen well in advance.
- (4) Can this hybrid model predict high frequency (3-hourly) sea surface temperature variability observed in buoys? If so, at what lead times? We have not come across any machine learning studies which have attempted to predict high frequency sea surface temperature signals.
- (5) How good does this hybrid model perform during extreme events like cyclones?
- (6) How truthful this hybrid model is for applications like marine heat wave forecasts? This is important given that marine heat waves are known to drive coral bleaching (Hughes *et al* 2017), spread harmful algal blooms (Trainer *et al* 2020) and adversely impact fishing (Mills *et al* 2013).

In this study, we develop a hybrid model similar to VMD-MIM and use the north-eastern Arabian Sea as the testbed to seek answers to the above questions. Instead of using MIM, we use the relatively simpler machine learning model LSTM despite the fact that LSTM suffers from gradient problems in the presence of data abundance and its relative insufficiency to extract spatial features (Rao *et al* 2018, Xu *et al* 2020). We shall show that irrespective of these known concerns, our hybrid model extracts high resolution spatio-temporal features with surprising deftness.

We have organized the rest of the manuscript in the following sequence. Section 2 describes the study area and the various sets of data used in this study. Section 2 also discusses the LSTM model that was implemented to predict the SST. Essential details of the reanalysis against which the predicted SST from the ML model has been compared is also briefly discussed. Section 3 highlights the analysis carried out in this study, followed by the conclusions in section 4.



2. Data and methods

This section presents the details about the study area, the observations used in this study and short descriptions of the numerical models from which reanalysis/forecast data has been used. We shall also outline the hybrid machine learning framework in this section.

2.1. Data and study area

To address the questions we posed in the Introduction, it is important to find a domain that is rich in SST fronts (abundance in fisheries), prone to MHWs, houses a buoy that measures SST at high frequencies and is testament to extreme events like cyclones. Fortunately, the Gulf of Kutch and its surroundings meet all these criteria. It is a known fishing ground for fishermen and had been designated as Marine Sanctuary and Marine National Park in 1980 and 1982 respectively with hundreds of coral species, algae, sponges and mangroves. A buoy is present within its perimeters that measures SST at a frequency of 1 h. This domain has also been experiencing elevated MHW activities with climate change (Chatterjee *et al* 2022) and has witnessed extreme events in recent times like cyclone Vayu in 2019 and cyclone Tauktae in 2021. It is for these reasons we choose the north-eastern Arabian Sea encompassing the Gulf of Kutch as our area of study. The extent of our study area ranges from (67–73) °E and (18–23) °N (see inset of figure 1).

We use near real time level 4 daily SST from OSTIA (Donlon *et al* 2012, Good *et al* 2020) at a horizontal resolution of ~5.5 km over our study area covering the period from 01 Jan 2012 to 31 Dec 2021 to train and test the hybrid machine learning framework for SST map predictions. OSTIA provides gap free gridded SST analysis derived from satellite observations (both infrared and microwave radiometers) and *in situ* data from ships, moored buoys and drifting buoys. The OSTIA SST analysis when compared to all the near-surface (3–5 m) Argo data from the EN4 dataset (~1300 observations) over the entire Indian Ocean (encompassing our domain) shows an impressive mean difference of ~−0.09 °C (table VI.6 in <https://documentation.marine.copernicus.eu/QUID/CMEMS-SST-QUID-010-001.pdf>). Even though OSTIA SST is an analysis generated from multiple observational platforms, we shall treat it as observations in this study. It is also important to mention here that the Argo data has an accuracy of ~0.005 °C (Oka and Ando 2004). In short, the accuracy of OSTIA SST over the Indian Ocean (and hence over our study area) is very high. We plot the daily SST from OSTIA at 70.225 °E, 22.925 °N (location exhibiting maximum SST variability) in figure 1 to apprise the readers of the typical range of variability in SST during the period 2012–2021.

We have also used a buoy located at 67.5 °E, 18.5 °N to train and test the hybrid model at high frequencies (3-hourly). The buoy observes the SST at a depth of 1 m from mean sea level every hour but owing to missing data across multiple time-steps, the buoy data has been averaged over every 3 h. The accuracy of buoy measurement is 0.002 °C (Venkatesan *et al* 2013).

The skills of our SST predictions are assessed against SST from buoys and from eddy-resolving high-resolution (~9 km) Global Ocean Physics Reanalysis (GLORYS12V1; Lellouche *et al* 2021) and eddy-resolving high-resolution (~9 km) forecast from Indian Ocean High Resolution Operational Ocean Forecast System (IO-HOFS) (Francis *et al* 2020) that uses initial condition from the Regional Analysis of Indian Ocean (Baduru *et al* 2019). The GLORYS12V1 (henceforth only GLORYS) reanalysis uses the NEMO ocean model (Madec 2008) with a horizontal resolution of ~9 km at the equator and 50 vertical levels. The European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim atmospheric reanalysis (Dee *et al* 2011) provides the surface

fluxes to GLORYS. It assimilates satellite sea level anomalies, SST, sea ice concentration, and *in situ* temperature and salinity profiles using a reduced-order Kalman filter derived from a singular evolutive extended Kalman (SEEK) filter (Brasseur and Verron 2006) and 3D-Var scheme for bias correction. In contrast, IO-HOOFs uses Regional Ocean Modeling System (ROMS) with a horizontal resolution of ~ 9 km and 40 vertical sigma levels as the ocean engine for the Indian Ocean and assimilates SST, satellite sea level anomaly, *in situ* temperature and salinity using Local Ensemble Transform Kalman Filter (Hunt *et al* 2007, Baduru *et al* 2019, Baduru *et al* 2025). IO-HOOFs provides a 3-hourly ocean state forecast of the Indian Ocean (30°E – 120°E ; 30°S – 30°N).

2.2. Method

LSTM networks, as an enhanced form of RNN, address both the issue of long-distance dependence—an inherent challenge for traditional RNNs—and the problems of exploding or vanishing gradients that are commonly encountered in neural networks (Graves 2012, Tan *et al* 2018, Contractor and Roughan 2021). These networks are particularly effective in processing sequence data, as they are known to reduce the nonlinearity and complexity of sequences through decomposition. Their unique structure includes three key gates—input, forget, and output—that control the flow of information. These gates allow LSTMs to decide what information to keep, discard, or process, helping them retain important context over time (Sarkar *et al* 2020, Hao *et al* 2024). The forget gate removes unnecessary information, the input gate adds new information, and the output gate produces the final result. The cell state serves as a memory that holds critical information across time steps, making it easier for the network to learn patterns in sequential data. Non-linear activation functions like sigmoid and hyperbolic tangent help capture complex relationships, while weights and biases fine-tune the learning process (Hochreiter and Schmidhuber 1997, Wei *et al* 2018, Hao *et al* 2024). They are widely used in areas such as time series forecasting and subsurface temperature prediction, as they effectively model sequential data (Han *et al* 2023, Vytla *et al* 2025). LSTM is implemented in our study following the steps prescribed in Wei *et al* (2018) and Hao *et al* (2024).

We combine the LSTM neural network model with the VMD algorithm to leverage these advantages to predict the SST for different lead times. This combined approach is called the Hybrid model in this study. LSTM architecture has one LSTM layer with 50 neurons and the Relu activation function followed by a fully connected layer. It is trained using 20 epochs, a batch size of 32, a learning rate of 0.001, the Adam optimizer and mean squared error (MSE) as the loss function. VMD is a frequency domain-based, quasi-orthogonal, and completely non-recursive multi-scale signal decomposition technique introduced by Dragomiretskiy and Zosso (2014). It formulates a variational problem, using the Hilbert transform for marginal spectrum computation, Wiener filtering for denoising, and the alternating direction multiplier method to handle unconstrained optimization. VMD effectively breaks down local features with similar frequencies while remaining robust to noise. Just like the Empirical Decomposition Mode (EMD) algorithm (Huang *et al* 1998) that transforms signal decomposition into recursive decomposition modes called intrinsic mode functions (IMFs), the VMD algorithm transforms the signal decomposition into completely non-recursive IMFs. To determine the optimal VMD parameters, we conducted sensitivity experiments varying the number of IMFs from 3 to 7. Based on predictive performance and computational efficiency, 5 modes (4 IMFs + 1 residual) were found optimal for our SST time series. Since decomposition is applied at each grid point (~ 5.5 km resolution), a higher number of IMFs significantly increases computational cost. Other VMD parameters used are $\alpha = 2000$ (moderate bandwidth constraint) and tolerance value of $1.e^{-7}$. In this study, we have decomposed the normalized (minimum-maximum normalization method) SST time series—either from OSTIA or from a buoy—at each grid point into 5 IMFs—4 IMFs and 1 residual (5th IMF). In figure 2, the five IMFs of the SST time series shown in figure 1 are presented for the readers to imbibe the qualitative features of the IMFs. Each of the IMFs represent a specific frequency band of the original SST time series.

Distinct LSTM models, each designed for a pre-designated lead time, are developed for each IMF (five IMFs in our study) at each grid point. At any given grid point and time, IMFs are estimated from the last 8 years (2922 time steps) of SST observations at that location. However, only the most recent 25 IMF values from each IMF are fed into their corresponding LSTM models, which then generate the predicted corresponding IMF value for the specified lead time. In essence, this process produces five predicted IMF values from the five LSTM models at each grid point at a given time. The final SST prediction at that grid point is the sum of these five predicted IMFs. Note that the prediction of SST at any location is independent of SST of all other locations. This procedure is repeated for all time steps of the study period (training and testing) before moving on to the next grid. The hybrid model is therefore a rolling prediction model for SST. We would like to apprise the readers that the IMFs change if the period of decomposition (which is 8 years here) is altered. Figure 3 shows a schematic diagram of our machine learning framework for a given grid location at an instant. We would also like to emphasize that the hybrid model does not have access to the testing data during the training phase, ensuring no leakage or bias.

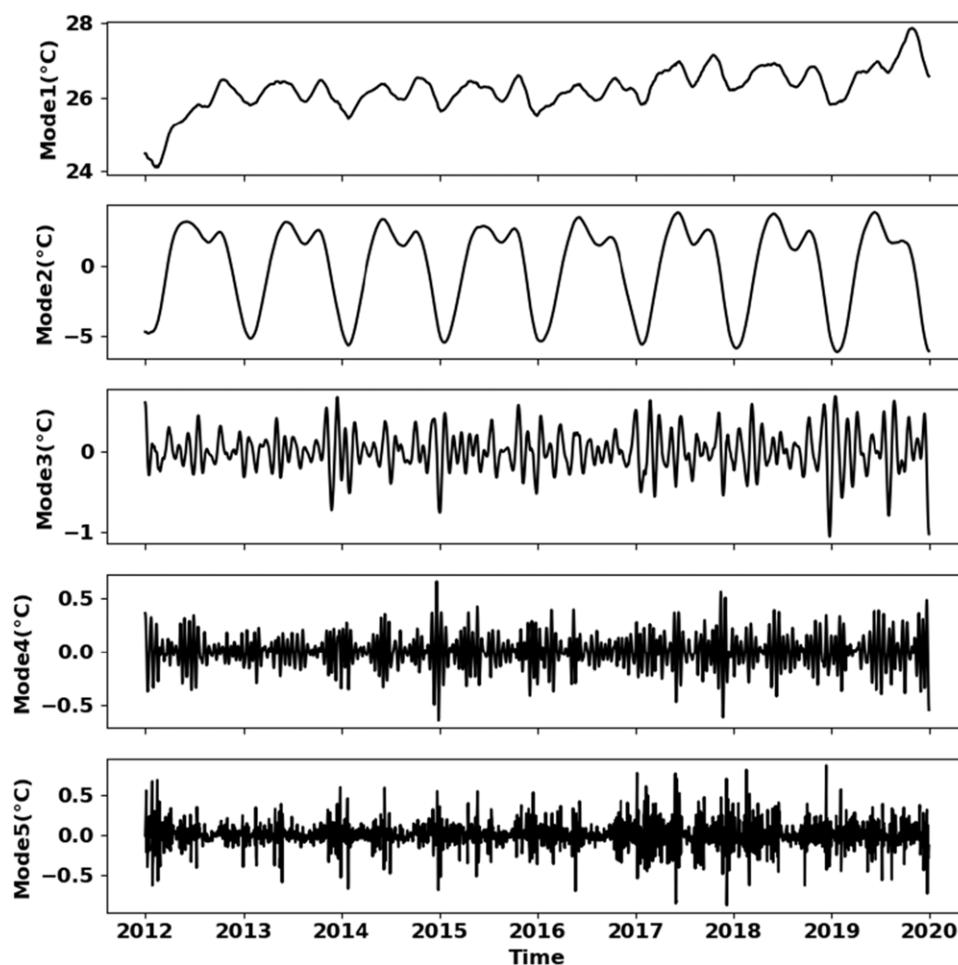


Figure 2. The five IMFs of the SST time series at 22.925 °N, 70.225 °E for the period 2012–2020. This location exhibits the highest SST variability within our domain.

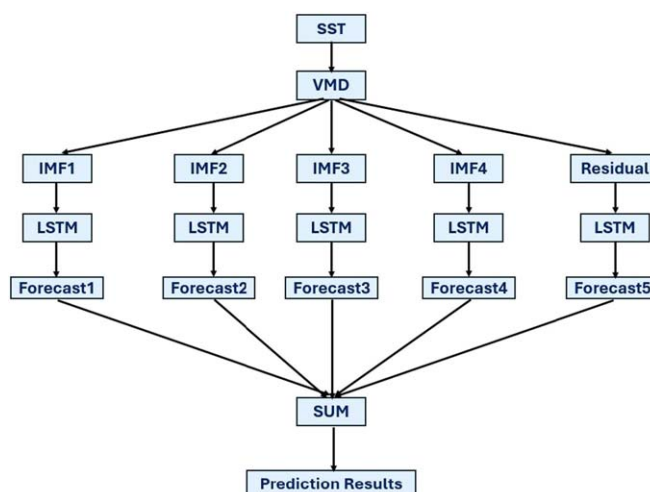
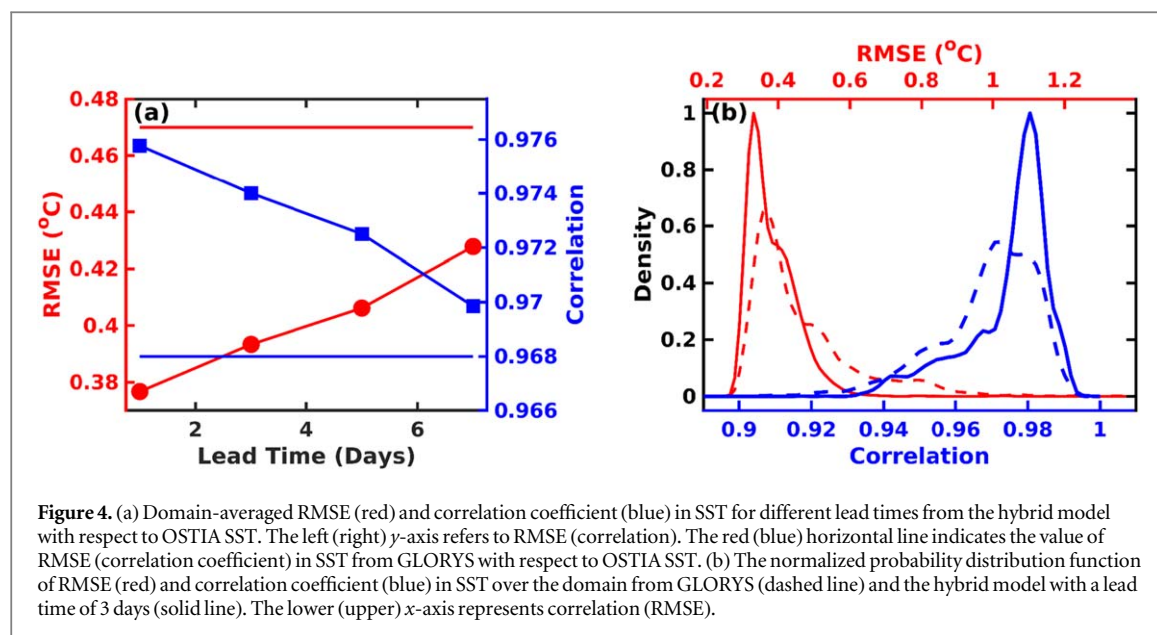


Figure 3. Schematic representation of the workflow of the hybrid model framework at any grid point at an instant.

As an evaluation metric, we use RMSE and Pearson correlation coefficients to evaluate the performance of a model in this study. The lower RMSE (larger correlation coefficient) implies greater accuracy in the model's predictions. The measures like RMSE and correlation coefficient determine the accuracy and skill of the model. However, it is important to draw attention that the non time-dependent performance quality like reliability and sharpness of a forecast is best analyzed through probability densities like marginal and conditional probability



analysis (Murphy 1993). For example, to assess the reliability of the forecast one should look at the conditional distribution of the observation given a forecast. A sharp conditional distribution centered around the given forecast indicates that the forecast is generally aligned with the observations, making it reliable. We shall analyze all these metrics and these distributions in detail in the subsequent section.

3. Results and discussion

We address each of the questions raised in the Introduction in succession. We start with comprehensively addressing the veracity of the forecasted SST maps (3-day lead time) from the hybrid model before moving on.

3.1. Fidelity of the hybrid model SST predictions

It's not surprising that the domain-averaged correlation (RMSE) decreases (increases) with increasing lead times for daily-averaged SST predictions from the hybrid model with a lead time of 3 days when compared to OSTIA SST for the testing period 2020–2021 (see figure 4(a)). Even for a lead time of 3 days, the domain-averaged correlation is as high as 0.974 while the domain-averaged RMSE is 0.393 °C. During 2020–2021, the domain-averaged correlation coefficient for GLORYS over the same study area is 0.968 while the domain-averaged RMSE is 0.47 °C—marginally inferior to the 7th day predictions from the hybrid model (see figure 4(a)). In short, on an average, the hybrid model manages to simulate the SST better than GLORYS 7 days ahead of time. It is also important to assess the skill of the hybrid model at a finer scale and analyze the distribution of large RMSEs across both these models. This shall indicate how abundant large RMSEs are in these two models. This will also indicate how the hybrid model shall fare in location specific forecasts and a measure of the maximum RMSE that can be expected if a location is randomly chosen anywhere within the domain of our study. In figure 4(b), we plot the normalized probability density of the RMSE (red) and correlation coefficient (blue) in SST from GLORYS (dashed line) and hybrid model (3 day lead time; solid line) over the domain to assess the performance of these two systems at a finer scale. The probability density plot demonstrates that the most probable RMSE of the predicted SST is smaller than the most probable RMSE of the GLORYS SST, i.e., the hybrid model manages to forecast a more accurate SST over most of the region compared to the SST reanalysis generated by GLORYS—that too 3 days ahead of time. The tail of the probability density shows that there are more locations in GLORYS that suffer from larger RMSE. It also shows that the largest RMSE that can be found in the hybrid model is about ~0.6 °C in contrast to ~1.1 °C in GLORYS. A similar inference of better skill in the hybrid model emerges from the probability density plot of correlation. There are more locations showing higher correlation in the hybrid model compared to GLORYS. These plots suggest that the 3rd day predicted SST from the hybrid model outperforms the SST reanalysis from GLORYS at both finer and coarser spatial scales. It is however apt to mention here that the number of data used to plot the probability density function differs for these two systems owing to the difference in horizontal resolution. The sample size is larger for the hybrid model because of its high resolution. A detailed investigation is therefore warranted. Unless otherwise mentioned, henceforth all reference to SST from the hybrid model shall refer to the SST from the hybrid model with a lead time of 3 days. In fact, the SST from the hybrid model fares reasonably well up to a lead time of ~7–10 days on a coarser spatial scale similar

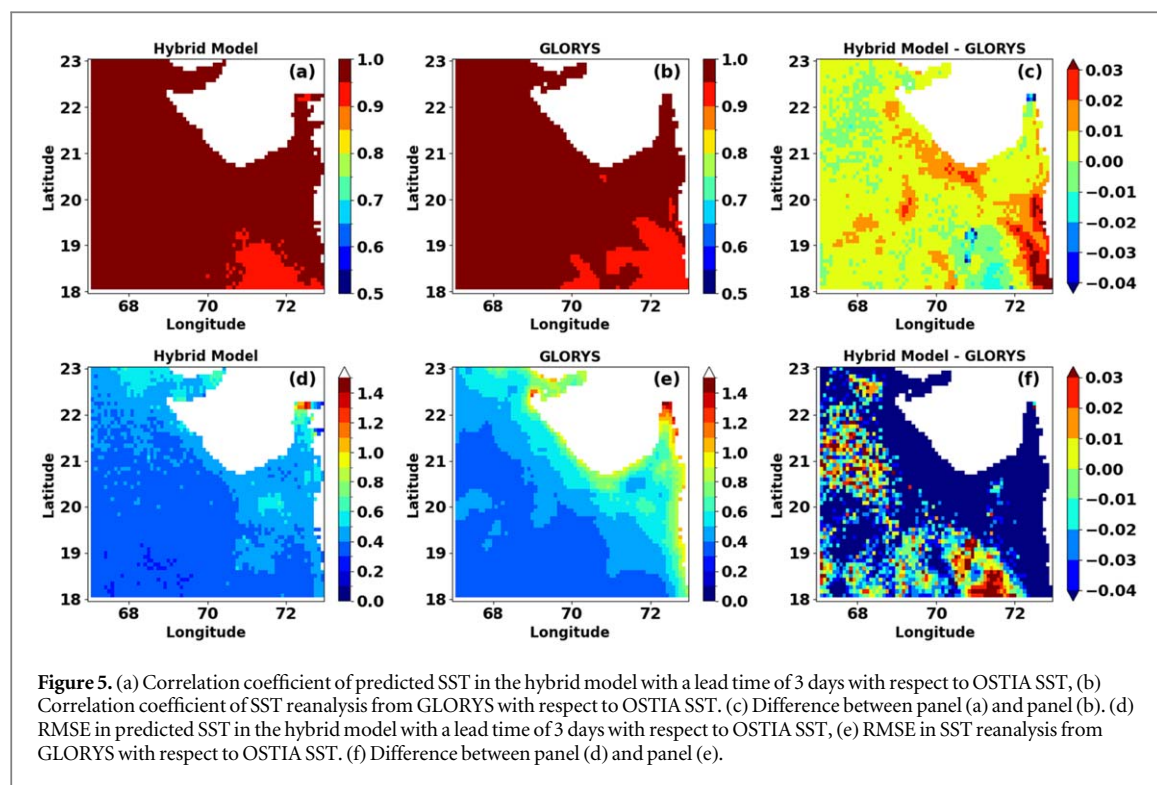
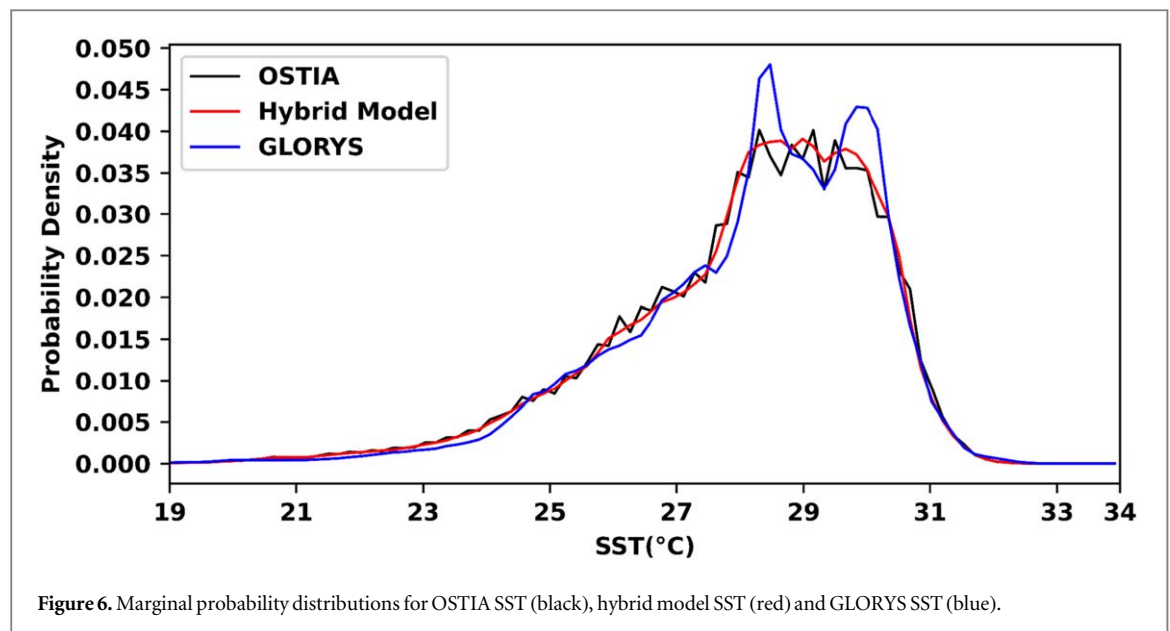


Figure 5. (a) Correlation coefficient of predicted SST in the hybrid model with a lead time of 3 days with respect to OSTIA SST, (b) Correlation coefficient of SST reanalysis from GLORYS with respect to OSTIA SST. (c) Difference between panel (a) and panel (b). (d) RMSE in predicted SST in the hybrid model with a lead time of 3 days with respect to OSTIA SST, (e) RMSE in SST reanalysis from GLORYS with respect to OSTIA SST. (f) Difference between panel (d) and panel (e).

to many other LSTM models (Shao *et al* 2021, Vytla *et al* 2025). Since one of the objectives of this study is to examine if the hybrid model manages to simulate the observed fine scale SST variability, we restrict this investigation to a lead time of 3 days.

In figure 5, we plot the maps of correlation coefficient and RMSE of daily SST from hybrid model and GLORYS with respect to OSTIA SST estimated over the period 2020–2021 along with the difference between the two. To carry out this analysis, the hybrid model SST is projected onto the GLORYS resolution. The correlation coefficients at all locations are larger than 0.9 across both the models. And the RMSE appears smaller for the hybrid model SST across most locations. However, there are locations where GLORYS SST is more aligned with the observations (see figure 5(f)). However, it is evident that the hybrid model SST demonstrates higher (lower) correlation (RMSE) across most of the domain—particularly close to the coast. Except over a region near the southern boundary of the domain, the hybrid model appears to fare better. We would like to alert the readers that the Pearson Correlation Coefficient is not a general measure of dependency and therefore one must dig deeper before asserting the supremacy of the predictive skill of the hybrid model.

We plot the marginal probability distribution of the daily OSTIA SST, hybrid model SST and GLORYS SST during the testing period in figure 6. The most probable SST in OSTIA lies in the range between 28 °C–30 °C, a feature reasonably reproduced by both the hybrid model and GLORYS. Even though both the hybrid model and GLORYS replicate the qualitative features of the observed SST probability density, GLORYS however generates a spurious high density for SST at ~28 °C and ~30 °C. Also, the tails of the probability density of the hybrid model SST is more aligned with the tails of the observed SST, whereas GLORYS shows deviations particularly for lower range of SST. In short, the marginal probability density of the hybrid model SST is better aligned with the marginal probability density of the observed SST. From a statistical perspective, Kullback–Leibler (KL, Kullback and Leibler 1951) divergence is the expectation value of the logarithmic difference between the probability distribution of two variables. Lower values indicate better alignment between the two distributions. In contrast, mutual information (MI, Thomas and Thomas 2005) quantifies the shared information between two variables, reflecting how well one predicts the other. Symmetric Uncertainty (SU, Witten *et al* 2005) is a normalized form of MI (0 to 1) that measures agreement between prediction and observation while accounting for their individual uncertainties. A higher MI signifies enhanced skill. In this analysis, the KL divergence values between the hybrid model SST (GLORYS SST) and the OSTIA SST is 0.0021 (0.0140) suggesting a more accurate probabilistic representation of SST in the hybrid model. Similarly, the SU metric based on MI between the observations and the hybrid model (0.3994) is higher than that between observations and GLORYS (0.3564), implying that the hybrid model retains more information about the observed SST, leading to better predictive performance. These statistical measures collectively confirm that while both models capture the general characteristics of the observed SST distribution, the hybrid model exhibits superior agreement in both distributional shape and predictive accuracy. This is surprising given that the skill of the predicted SST with a lead time of 3 days from a



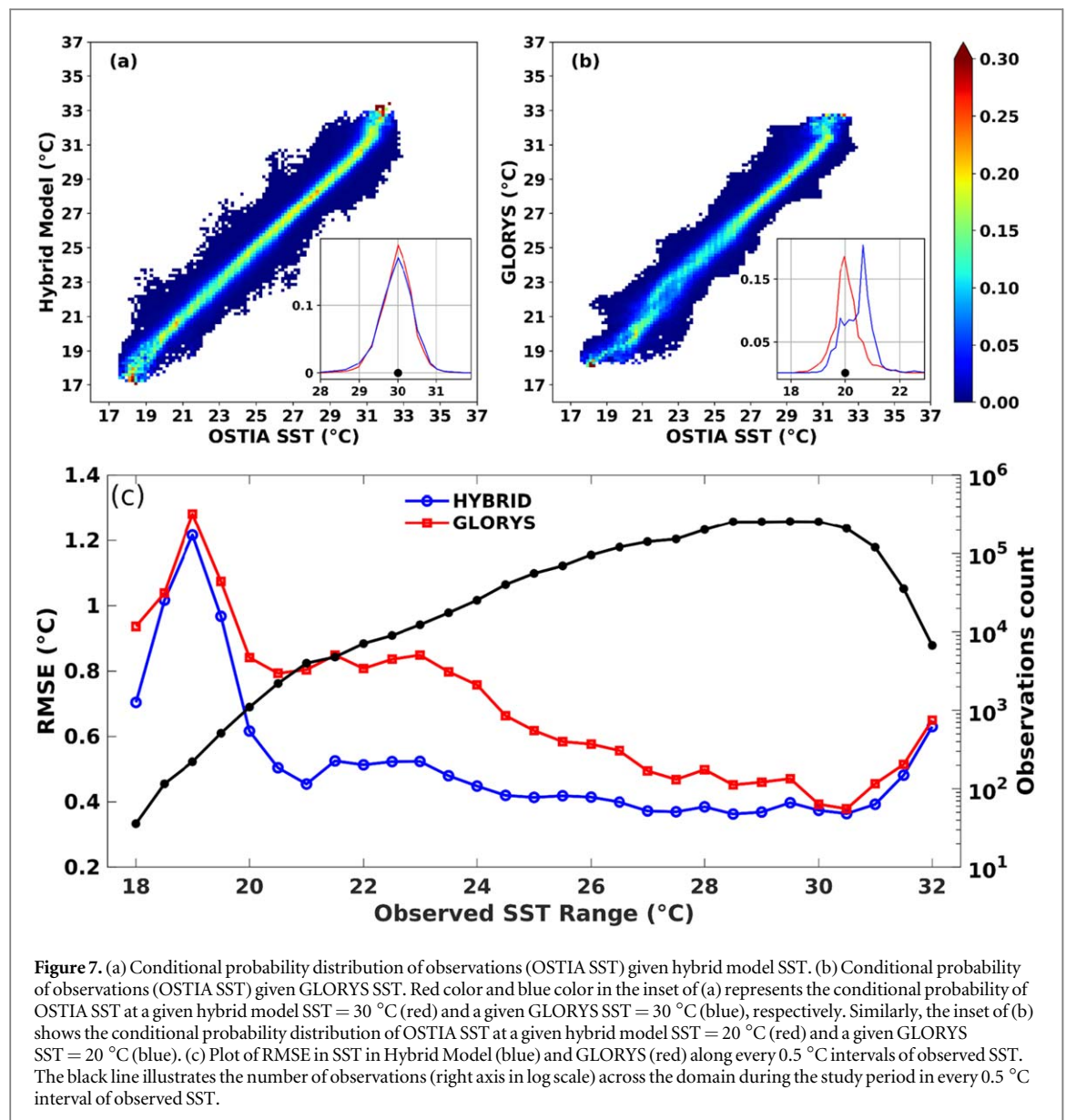
machine learning model is better than the skill of the SST reanalysis from GLORYS—one of the best state-of-the-art reanalysis available today.

3.2. Reliability of SST forecasts

We now ask the question: which of the two systems—hybrid model and Global Ocean Physics Reanalysis (GLORYS)—is more reliable? GLORYS do not produce forecasts but we assume GLORYS to represent the upper bound of predictive skill of forecasts of any dynamical system. To assess the reliability (Murphy 1993) of the SST predictions from the hybrid model versus GLORYS SST reanalysis, we analyze the conditional distribution of the observed OSTIA SST given the hybrid model (figure 7(a)) and the conditional distribution of the observed OSTIA SST given GLORYS SST (figure 7(b)). In both cases, the highest density regions align along the diagonal, indicating a strong correlation between the forecasts and observations. However, notable differences emerge in their probability distributions. The hybrid model (figure 7(a)) exhibits a sharper and more concentrated probability distribution along the diagonal, suggesting that it provides a more reliable and less dispersed prediction of OSTIA SST. In contrast, GLORYS (figure 7(b)) shows a broader spread, particularly at lower temperatures, indicating greater uncertainty in its predictions. The inset plots further illustrate this difference. When the hybrid model (GLORYS) predicts (simulates) 30 °C, the observed SST is generally found to be around 30 °C (see inset of figure 7(a)). However, there are rare occasions across both the models when the model predicts/simulates with an error of ± 1 °C. In contrast, when the hybrid model predicts 20 °C, the observed SST remains largely aligned to the hybrid model prediction. However, when GLORYS ‘predicts’ 20 °C, observed SST exhibits a broader multi-peaked distribution signalling significant deviations from the GLORYS ‘forecast’ (see inset of figure 7(b)). The observed SST is found to be ~ 20.75 °C during most of the time against GLORYS’s estimation of 20 °C. To more effectively illustrate model performance across varying temperature ranges, we pose the following question: What is the RMSE of sea surface temperature (SST) from the Hybrid model and GLORYS, relative to SST from OSTIA, computed within each 0.5 °C interval of observed SST between 18 °C and 32 °C?

Figure 7(c) presents the RMSE in SST from the Hybrid model (blue) and the GLORYS dynamical model (red) across different observed SST intervals, alongside the number of observations (black) within each 0.5 °C temperature bin across the spatio-temporal domain. Both models exhibit qualitatively similar RMSE patterns across the SST spectrum. However, performance notably degrades in the lower temperature range (18 °C–20 °C), where the Hybrid model shows a slight improvement over GLORYS.

This temperature range (18 °C–20 °C) corresponds to approximately 33 grid points within the study domain, all located near the coast over shallow bathymetry (7–16 m depth). These results suggest that both models face challenges in shallow coastal environments, particularly under lower temperature conditions. It is also important to note that observations in this low-temperature range are relatively sparse, as temperatures below 20 °C occur only briefly. Consequently, the Hybrid model—being data-driven—may be more susceptible to errors in this range due to limited training data. These findings suggest that the hybrid model provides a very reliable forecast that outsmarts the hindcasts from one of the best state-of-the-art SST reanalysis. It further



signifies that no existing SST forecasts from any dynamical model can compete with the forecasts from the hybrid model if our assumption of GLORYS representing best possible forecasts stand true.

3.3. Prediction of SST fronts

Given that the hybrid model predicts SST over our study area with a lead time of 3 days with excellent accuracy and reliability, we assess the ability of the model to predict small-scale features like SST fronts which are necessary for Potential Fishing Zone (PFZ) forecasts. The SST fronts are identified using Cayula Cornillon Algorithm (Cayula and Cornillon 1992). Grossly stating, a bimodal distribution is sought for within a spatial scale of 32 times the grid resolution. If such a distribution exists, the SST fronts are detected within this algorithm if there is a difference of a certain threshold (set to 0.3 °C in this study) between the two means of the bimodal distribution. Since GLORYS SST is of a relatively coarser resolution, it is projected on to the same resolution of OSTIA before estimating the SST fronts. In figure 8, SST fronts with a lead time of 3 days (red) are plotted along with the observed SST fronts (black) and SST fronts from GLORYS (blue) for a particular day (07 Feb, 2020) during the study period. The predicted SST fronts from the hybrid model are remarkably aligned with most of the observed SST fronts. In contrast, GLORYS generates some spurious SST fronts. The spurious SST fronts in GLORYS may be an offshoot of the downscaling of GLORYS SST and too much should not be read into it. The qualitative aspect of the plot remains similar for other days during the study period as well (not shown). In short, the hybrid model SST fronts are accurate, i.e., the hybrid model manages to forecast the small-scale behaviour of SST with a lead time of 3 days. This raises the possibility of providing robust PFZ forecasts that are likely to benefit millions of fishermen.

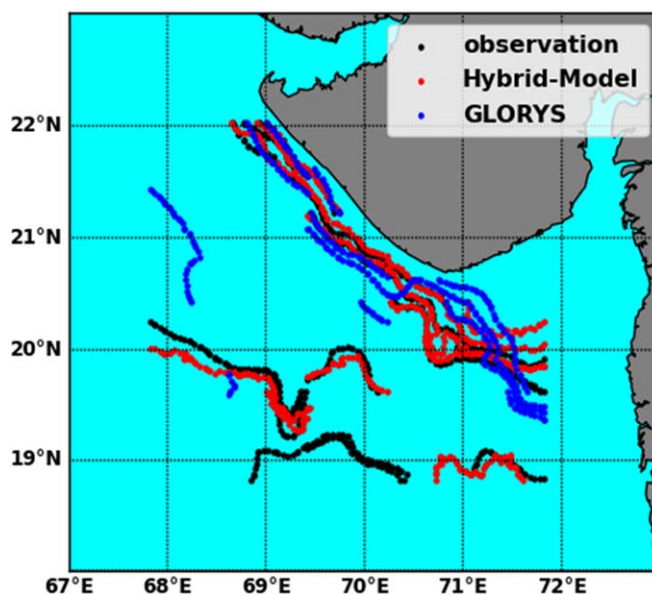


Figure 8. SST fronts from (a) OSTIA (black), (b) GLORYS (blue) and (c) hybrid model (red) with a lead time of 3 days on 07 February 2020.

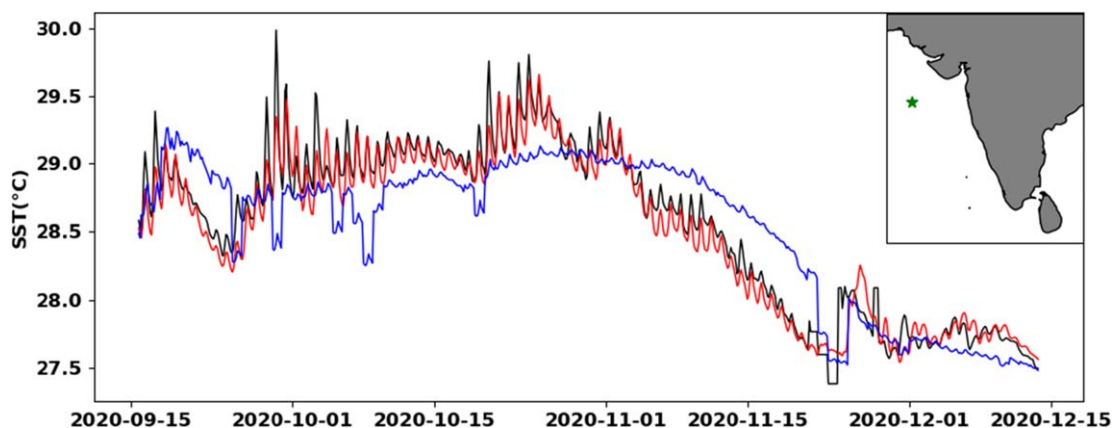
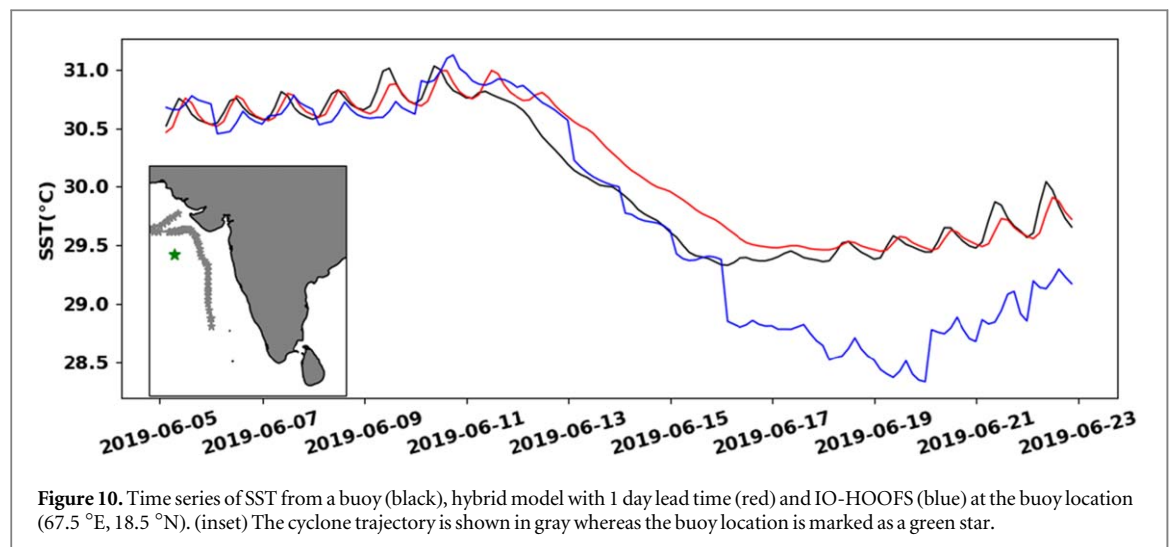


Figure 9. Time series of SST from a buoy (black), IO-HOOFs (blue), and from the hybrid model with a lead time of 24 h (red). The green star in the inset marks the location of the buoy at 18.5 °N, 67.5 °E.

3.4. High frequency SST predictions

Now that it is comprehensively demonstrated that the hybrid model has excellent predictive skill in forecasting daily-averaged SST, it is important to assess how the hybrid model fares in predicting high frequency SST variability. In figure 9, we plot the 3 hourly SST predicted by the hybrid model with a lead time of 24 h (red) along the observed 3-hourly SST (black) from a buoy located at 18.5 °N, 67.5 °E. A hybrid model was configured only for this specific location to predict 3-hourly SST with a lead time of 1 day (24 h). The hybrid model for this analysis was trained for 1 year (mid-September 2019 - mid-September 2020) whereas the testing period extended for the subsequent 3 months. The hybrid model manages to predict the 3-hourly SST with remarkable accuracy with a lead time of a day. The correlation is almost unity (0.98) and the RMSE is 0.137 °C. Expectedly, the skill of the hybrid model declines with increasing lead times (not shown). To provide the readers with an insight about how dynamical models behave at such high frequencies, we have also overlaid 3-hourly SST from our inhouse operational forecast system IO-HOOFs (blue) in figure 9. It is evident that the forecast skill of the dynamical model lags far behind the hybrid model even for high frequency forecasts. The RMSE in IO-HOOFs SST is ~2.5 times larger and the correlation considerably decreases. All the above analyses show that predictions in SST from the hybrid model agree with the observed SST features across low (daily-averaged) and high (3 hourly averaged) frequencies.

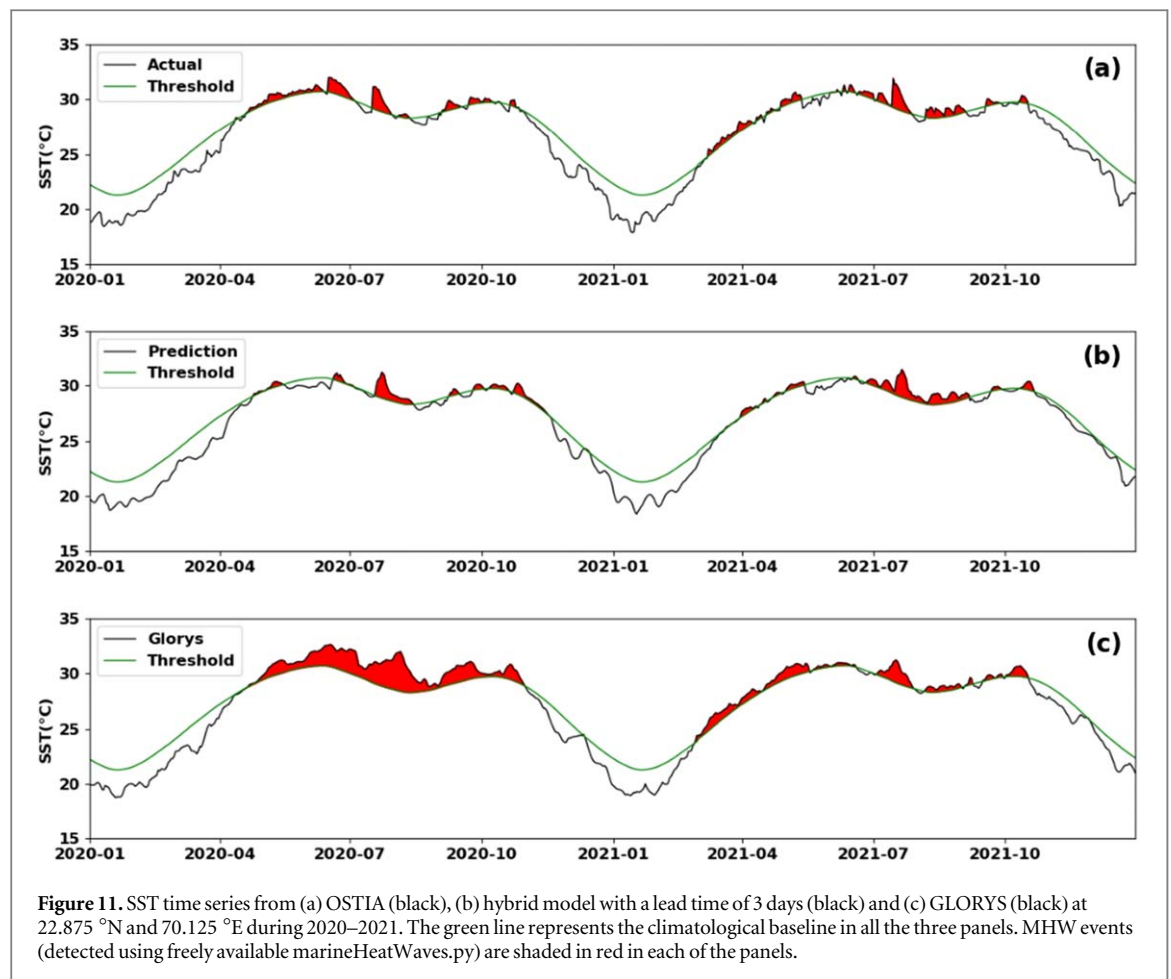


3.5. Prediction skill during extreme events

Since the hybrid model SST manages to mimic the observation at both low and high spatio-temporal scales, it is important to assess if the high frequency predictions from this hybrid model stays true during extreme events. These evaluations are difficult during extreme events because of primarily two reasons—1) there should be an observation platform somewhere in the vicinity of the cyclone trajectory to observe the high frequency SST evolution and 2) the instrument shall remain operational during the extreme events which seldom happens. Satellite SST observations are not useful; it often turns out to be unreliable because of cloud cover and rain during tropical cyclones (Wentz *et al* 2000) and it does not observe high frequency SST. Fortunately, the same buoy (67.5 °E, 18.5 °N), located about 250 kms away from the trajectory of the very severe cyclone Vayu in the Arabian Sea, observed SST at a temporal frequency of 3 h (see inset of figure 10 for the buoy location with respect to the cyclone trajectory). A hybrid model was configured only for this specific location to predict 3-hourly SST during the Vayu cyclone with a lead time of 1 day (24 h). The hybrid model was trained for a year using historical 3-hourly-averaged buoy SST observations (30 Jan 2018 to 30 Jan 2019), validated for 3 months (31 Jan 2019 to 01 May 2019) and tested for another 3 months (02 May 2019 to 31 July 2019). The testing period encompasses the cyclone passage period (June 10, 2019–June 17, 2019). In figure 10, we plot the observed SST (black) and the predicted SST with a lead of 1 day (red) along with the 3 hourly SST forecasts from IO-HOOFs (blue) at this location. The sudden drop of ~ 1.5 °C in the observed SST on June 12, 2019 bears testament to the observed fact that the cyclone Vayu reached the buoy vicinity on the early morning of June 12, 2019. The hybrid model facilitates a similar drop but suffers from a lag of ~ 27 h which is longer than the prediction horizon of 24 h thereby making it unsuitable for SST predictions during extreme events. This is not surprising given that the hybrid model did not experience any extreme events during the course of its training. Nevertheless, the RMSE remains low (0.18 °C). The 3-hourly forecast from IO-HOOFs with a lead time of 1 day appears to be out of phase with the observations by a few hours and the drop in SST is more pronounced. The strong winds (~ 36 m s $^{-1}$) associated with the cyclone facilitates a SST drop by ~ 2.75 °C—about twice the drop witnessed in the observation—possibly an outcome of inaccurate parameterization in the model dynamics. This results in a large RMSE of 0.5 °C in the model SST during the cyclone period. In summary, this analysis illustrates that the hybrid model, equipped with only historical SST inputs, is insufficient for SST predictions during extreme events, likely because of the model's limited exposure to extreme events during training. However, it may be possible to predict the high frequency SST evolution under this hybrid framework by including additional predictors like sea surface height, mixed layer depths, cyclone characteristics etc (Cui *et al* 2023).

3.6. Marine heat wave (MHW) forecast skill

It will be interesting to explore how good the hybrid model is in predicting MHWs which is defined as a discrete, prolonged, and anomalous warm water event (Hobday *et al* 2016). We used 30 years of OSTIA SST from 1991 to 2020 as the climatological baseline. MHW is identified by the threshold value of the 90th percentile of daily climatology, and the MHW is announced if the SST persists above this threshold for a minimum of 5 days. The MHW events are deemed to be separate events if a gap of two days or more exist between two detections. In figures 11(a)–(c), we have plotted OSTIA SST, hybrid model SST with a lead of 3 days, and GLORYS SST reanalysis at 22.875 °N and 70.125 °E along with the climatological baseline (green) during the period 2020–2021. This location is known to witness frequent MHWs—about 2–2.5 MHWs every year on an average (see figure 1 in Saranya *et al* 2022). The observed MHWs occur each year mostly during April to November at



this location. It shows multiple MHW events during this period with some events persisting for a couple of weeks. The hybrid model predicts the MHWs with reasonably good accuracy with a lead time of 3 days. The predicted MHWs are mostly aligned with the observed MHWs. It manages to predict most of the pronounced MHWs but falters at predicting some minor MHWs particularly during the onset of boreal summer. In contrast, GLORYS estimates a single MHW event during 2020 and therefore fails to delineate the multiple MHWs that are observed during this period. This reflects a persistently warm SST in GLORYS at this location. In short, the hybrid model manages to predict the pronounced MHWs with reasonable certainty with a lead time of 3 days.

4. Conclusions

We have developed a univariate hybrid machine learning model for prediction of regional SST maps. The historical SST time-series at each location is decomposed into IMFs which are thereafter fed to the machine learning model LSTM with a specified lead time. The predicted SST is then reconstructed from the output IMFs. Likewise, this process is repeated for each location within the domain to generate the predicted regional map of SST. Unlike many machine learning models wherein a suite of variables act as predictors, only historical SST acts as the predictor in our hybrid model without compromising on the skill of the predictand. There are a multitude of advantages in this approach. This approach aids to generate high resolution SST maps which were otherwise not possible owing to nonavailability of high resolution datasets of other variables. For example, daily sea surface height anomaly, which often acts as a predictor to SST (Shao *et al* 2021, Cui *et al* 2023), is only available at a resolution of $1/8^\circ$ which would have limited the resolution of the final SST regional map. The use of only historical SST data also uses less computational resources and executes the predictions at a faster rate. Even though the present study is designed to forecast SST over a small domain, it can be seamlessly extended to a global domain. The success of this framework demonstrates that a basin-scale SST prediction system can be treated as a collection of localized independent low-dimensional systems insulated from the influence of all other state variables and forcings. This challenges our traditional understanding of high-dimensional dynamical systems, like oceans, where strong cross-correlations between different state variables and finite auto-correlation length scales are crucial for their temporal evolution.

We show that the forecasted daily SST from the hybrid model with a lead time of 3 days comes with a very high correlation and low RMSE—both close to the coast and in open waters—when compared against daily OSTIA SST. The skill of this predicted SST outperforms the skill of the high-resolution SST reanalysis from GLORYS. The skill remains higher till a prediction horizon of 7 days. In short, a high resolution daily SST regional map ($1/20^\circ$ spatial resolution) can be generated a week prior to the availability of a SST reanalysis from a state-of-the-art numerical model equipped with data assimilation.

The predicted regional SST map is adept at reproducing the fine-scale spatial features of the observations. It predicts the SST fronts with remarkable accuracy with a lead time of 3 days which can be beneficial to the fishermen community to locate potential fishing zones. The hybrid model is also skillful at predicting high frequency SST signals during normal conditions with a lead time of 1 day. The correlation (RMSE) of the predicted SST remains high (low). We have also shown how this hybrid model detects the MHWs about 3 days prior to its occurrence. This prior knowledge may be used to mitigate harmful events like coral bleaching. The hybrid model is however inept to predict high frequency SST signals during extreme events.

The daily predicted SST maps raise interesting possibilities. For example, these predicted SSTs may be treated as synthetic observations and assimilated in an ocean model during the forecast which is likely to improve the ocean state forecasts of the numerical model. These synthetic observations shall not only improve the surface temperature during forecasts but also the vertical temperature structures of the ocean. With the proposed advent of microsatellites, this framework may be used to predict SST maps at a higher temporal frequency like 3 h or 6 h. Such high frequency synthetic SST maps may be used to force atmospheric models to improve the atmospheric forecasts.

Acknowledgments

The authors duly acknowledge the funding received from the Ministry of Earth Sciences, Govt. of India. We extend our sincere thanks to the ICT division at INCOIS for managing the computing facility and timely installation of the required modules. We acknowledge Dr R. Venkat Seshu, INCOIS for providing the buoy data. This is INCOIS contribution number 578.

Data availability statement

All the data except one buoy data is publicly available. The buoy data can be shared upon reasonable request. The data that support the findings of this study are available upon reasonable request from the authors. https://data.marine.copernicus.eu/product/SST_GLO_SST_L4_REP_OBSERVATIONS_010_011/description https://data.marine.copernicus.eu/product/GLOBAL_MULTIYEAR_PHY_001_030/services <https://doi.org/10.5281/zenodo.15003614> https://odis.incois.gov.in/essdp/searchMetadata?queryText=&orgName=NIOT&contactName=&start_date=&end_date=&north=90&south=-90&west=-180&east=180

Authors contribution

Conceptualization: AP, BB and JP; **Writing:** AP, JP, BB; **Analysis:** JP, BB, ACR, BP, and VD; **Interpretation:** JP, BB and AP.

ORCID iDs

Jagdish Prajapati  <https://orcid.org/0000-0002-5279-9729>

Balaji Baduru  <https://orcid.org/0000-0001-9641-4109>

Biswamoy Paul  <https://orcid.org/0000-0003-1742-6799>

Vinod Daiya  <https://orcid.org/0009-0002-1870-0374>

Arya Paul  <https://orcid.org/0000-0002-6635-3058>

References

- Aparna S, D'souza S and Arjun N B 2018 Prediction of daily sea surface temperature using artificial neural networks *Int. J. Remote Sens.* **39** 4214–31
- Baduru B, Paul B, Banerjee D S, Sanikommu S and Paul A 2019 Ensemble based regional ocean data assimilation system for the Indian Ocean: implementation and evaluation *Ocean Modell.* **143** 101470

- Baduru B, Paul B, Athul C R, Paul A and Francis P A 2025 Improving Indian Ocean analysis using ROMS with sea level anomaly assimilation *J. Earth Syst. Sci.* **134** (2) 1–19
- Brasseur P and Verron J 2006 The SEEK filter method for data assimilation in oceanography: a synthesis *Ocean Dyn.* **56** 650–61
- Cayula J F and Cornillon P 1992 Edge detection algorithm for SST images *JTECH* **9** 67–80
- Chatterjee A, Anil G and Shenoy L R 2022 Marine heatwaves in the Arabian Sea *Ocean Sci.* **18** 639–57
- Choi H-M, Kim M-K and Hyun Y 2023 Deep-learning model for sea surface temperature prediction near the Korean Peninsula *Deep-Sea Res. II: Top. Stud. Oceanogr.* **208** 105262
- Contractor S and Roughan M 2021 Efficacy of feedforward and LSTM neural networks at predicting and gap filling coastal ocean timeseries: oxygen, nutrients, and temperature *Front. Mar. Sci.* **8** 1–17637759
- Cui H, Tang D, Mei W, Liu H, Sui Y and Gu X 2023 Predicting tropical cyclone-induced sea surface temperature responses using machine learning *Geophys. Res. Lett.* **50** 1–11
- Donlon C J, Martin M, Stark J, Roberts-Jones J, Fiedler E and Wimmer W 2012 **116** 140–158
- Dee D P, Uppala S M, Simmons A J, Berrisford P, Poli P, Kobayashi S and Vitart F 2011 The ERA-interim reanalysis: configuration and performance of the data assimilation system *QJR Meteorol. Soc.* **137** 553–97
- Dragomiretskiy K and Zosso D 2014 Variational mode decomposition *IEEE TSP* **62** 31–544
- Francis P A et al 2020 High-resolution operational ocean forecast and reanalysis system for the Indian ocean *Bull. Am. Meteorol. Soc.* **101** E1340–56
- Good S et al 2020 The current configuration of the OSTIA system for operational production of foundation sea surface temperature and ice concentration analyses *Remote Sens.* **12** 720
- Graves A 2012 *Long Short-Term Memory* (Springer) (https://doi.org/10.1007/978-3-642-24797-2_4)
- Halliwel G R 2004 Evaluation of vertical coordinate and vertical mixing algorithms in the HYbrid-coordinate ocean model (HYCOM) *Ocean Modell.* **7** 285–322
- Han Zhang H, Mengyuan J, Haoyu Z, Longjie L, Yunxia Z, Jie T, Di T and Yanmin Z 2023 Deep learning approach for forecasting sea surface temperature response to tropical cyclones in the Western North Pacific *Deep-Sea Res. I: Oceanogr. Res. Pap.* **197** 104042
- Hao Y, Lu J, Peng G, Wang M, Li J and Wei G 2024 F10.7 daily forecast using LSTM combined with VMD method *Space Weather* **22** e2023SW003552
- Hobday A J et al 2016 A hierarchical approach to defining marine heatwaves *Oceanogr. Prog.* **141** 227–38
- Hochreiter S and Schmidhuber J 1997 Long short-term memory *Neural Comput.* **9** 1735–80
- Huang N E et al 1998 The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis *Proceedings of the Royal Society of London: Mathematical, Physical and Engineering Sciences* **454** 903–95
- Hughes T P, Kerry J T and Wilson S K 2017 Global warming and recurrent mass bleaching of corals *Nature* **543** 373–7
- Hunt B R, Kostelich E J and Szunyogh I 2007 Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter *Phys. D: Nonlinear Phenom* **230** 112–26
- IOCCG 2009 Reports of the international ocean-colour coordinating group, no. 8) (Dartmouth, NS: International Ocean-Colour Coordinating Group *Remote Sensing in Fisheries and Aquaculture* (IOCCG) (<https://doi.org/10.25607/OBP-98r>)
- Kullback S and Leibler R A 1951 On information and sufficiency *The Ann. Math. Stat.* **22** 79–86
- Large W G, McWilliams J C and Doney S C 1994 Oceanic vertical mixing: a review and a model with a nonlocal boundary layer parameterization *Rev. Geophysics* **32** 363–403
- Lellouche J M et al 2021 The copernicus global 1/12° oceanic and sea ice reanalysis *EGU General Assembly Conf. Abstracts EGU21-14961* (<https://doi.org/10.5194/egusphere-egu21-14961>)
- Liu G et al 2013 NOAA coral reef watch 50 km satellite sea surface temperature-based decision support system for coral bleaching management NOAA Technical Report, NESDIS, 143 <https://repository.library.noaa.gov/view/noaa/743>
- Madec G 2008 NEMO Ocean Engine *Note du Pôle de modélisation Institut Pierre-Simon Laplace (IPSL) France* **27** https://frouingroup.ucsd.edu/Most_recent_figs/Refs/NEMO_book_v3_3.pdf<https://www.scrip.org/reference/referencespapers?referenceid=2978123>
- McTaggart-Cowan R, Davies E L, Fairman J G Jr, Galarneau T J Jr and Schultz D M 2015 Revisiting the 26.5 °C sea surface temperature threshold for tropical cyclone development *Bull. Am. Meteorological Soc.* **96** 1929–43
- Mills K E et al 2013 Fisheries management in a changing climate: lessons From the 2012 ocean heat wave in the Northwest Atlantic *Oceanogr.* **26** 191–5
- Mohanty P C, Kushabaha A, Mahendra R S, Nayak R K, Sahu B K, Pattabhi Rama Rao E and Srinivas Kumar T 2021 Persistence of marine heat waves for coral bleaching and their spectral characteristics around Andaman coral reef *Environ. Monit. Assess.* **193** 491
- Murphy A H 1993 What is a good forecast? An essay on the nature of goodness in weather forecasting *WAF* **8** 281–93
- Oka E and Ando K 2004 Stability of temperature and conductivity sensors of argo profiling floats *J. Oceanogr.* **60** 253–8
- Patil K, Deo M C and Ravichandran M 2016 Prediction of sea surface temperature by combining numerical and neural techniques *JTECH* **33** 1715–26
- Rao P, Jiang W, Hou Y, Chen Z and Jia K 2018 Dynamic change analysis of surface water in the yangtze river basin based on MODIS products *Remote Sens.* **10** 1025
- Roxanne L 2016 *Exploring the Future of Marine Farming in New Zealand Under Climate Change Conditions: Using Sea Surface Temperature* (Canterbury:) (Lincoln University)
- Sarkar P P, Janardhan P and Roy P 2020 Prediction of sea surface temperatures using deep learning neural networks *SN Appl. Sci.* **2** 1458
- Saranya J S, Roxy M K, Dasgupta P and Anand A 2022 Genesis and trends in marine heatwaves over the tropical Indian Ocean and their interaction with the Indian summer monsoon *J. Geophys. Res. Oceans* **127** e2021JC017427
- Sepp H and Schmidhuber J 1997 Long short-term memory *Neural Comput.* **9** 1735–80
- Shao Q, Li W, Han G, Hou G, Liu S, Gong Y and Qu P 2021 A deep learning model for forecasting sea surface height anomalies and temperatures in the South China Sea *J. Geophys. Res. Oceans* **126** 1–18
- Shuang X, Dejun D, Xuerong C, Xunqiang Y, Shumin J, Haidong P and Guansuo W 2023 A deep learning approach to predict sea surface temperature based on multiple modes *Ocean Modell.* **181** 102158
- Tan Y, Hu Q, Wang Z and Zhong Q 2018 Geomagnetic index Kp forecasting with LSTM *Space Weather* **16** 406–16
- Thomas M and Thomas J A 2005 *Elements of Information Theory* (Wiley) 13–55
- Trainer V L, Kudela R M, Hunter M V, Adams N G and McCabe R M 2020 Climate extreme seeds a new domoic acid hotspot on the US West Coast *Front. Clim.* **2** 571836
- Tripathi K C, Das M L and Sahai A K 2006 Predictability of sea surface temperature anomalies in the Indian Ocean using artificial neural networks *IJMS* **35** 210–20 <http://nopr.niscpr.res.in/handle/123456789/1518>
- Usharani B 2023 ILF-LSTM:enhanced loss function in LSTM to predict the sea surface temperature *Soft Comput.* **27** 13129–41

- Venkatesan R, Shamji V R, Latha G, Mathew S, Rao R R, Muthiah A and Atmanand M A 2013 *In situ* ocean subsurface time-series measurements from OMNI buoy network in the Bay of Bengal *Curr. Sci.* **104** 1166–1177
- Vytla V *et al* 2025 Forecasting of sea surface temperature using machine learning and its applications *J. Earth Syst. Sci.* **134** 25
- Wang Y, Zhang J, Zhu H, Long M, Wang J and Yu P S 2019 Memory in memory: a predictive neural network for learning higher-order non-stationarity from spatiotemporal dynamics *2019 IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)* 9146–54
- Wei L, Zhong Q, Lin R, Wang J, Liu S and Cao Y 2018 Quantitative prediction of high-energy electron integral flux at geostationary orbit based on deep learning *Space Weather* **16** 903–16
- Wei L and Guan L 2021 Seven-day sea surface temperature prediction using a 3DConv-LSTM model *Front. Mar. Sci.* **9**
- Wentz F J, Gentemann C, Smith D and Chelton D 2000 Satellite measurements of sea surface temperature through clouds *Science* **288** 847–50
- Witten I H, Frank E, Hall M A, Pal C J and Data M 2005 Practical machine learning tools and techniques *Data Mining 2* (Elsevier) 403–13
- Wolff S, O'Donncha F and Chen B 2020 Statistical and machine learning ensemble modelling to forecast sea surface temperature *J. Mar. Syst.* **208** 103347
- Xu L, Li Q, Yu J, Wang L, Xie J and Shi S 2020 Spatio-temporal predictions of SST time series in China's offshore waters using a regional convolution long short-term memory (RC-LSTM) network *Int. J. Remote Sens.* **41** 3368–89
- Yang Y, Dong J, Sun X, Lima E, Mu Q and Wang X 2018 A CFCC-LSTM model for sea surface temperature prediction *IEEE Geosci. Remote Sens. Lett.* **15** 207–11
- Zheng G, Li X, Zhang R H and Liu B 2020 Purely satellite data-driven deep learning forecast of complicated tropical instability waves *Sci. Adv.* **6** eaba1482