The importance of subsurface thermal effects to IOD prediction revealed by a skillful three-dimensional DL model

Junkai Qian[1,2], Qiang Wang[1,2*], Youming Tang[2], Rong-Hua Zhang[3], Henk A. Dijkstra[4], Yoo-Geun Ham[5], Wansuo Duan[6], Xiangzhou Song[1,2], Gang Huang[6], Suqi Peng[7]

[1]*Key Laboratory of Marine Hazards Forecasting, Ministry of Natural Resources, Hohai University, Nanjing, China*

[2]*College of Oceanography, Hohai University, Nanjing, China*

[3]*School of Marine Sciences, Nanjing University of Information Science and Technology, Nanjing, China*

[4]*Institute Marine and Atmospheric Research Utrecht, Utrecht, Netherlands*

[5]*Department of Environmental Managements, Graduate School of Environmental Studies, Seoul National University, Seoul, South Korea*

[6]*State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China*

[7]*Taishun Meteorological Bureau, Wenzhou, China*

ABSTRACT

The Indian Ocean Dipole (IOD) is a dominant interannual variability mode in the Indian Ocean, significantly influencing global weather, climate and society. Despite extensive research, current numerical models and surface-based deep learning models are limited in their ability to predict IOD events beyond five to seven months. To address this gap, we introduce a novel Multi-

* Corresponding author : Qiang Wang
Email: wangq@hhu.edu.cn

dimensional Air–Sea Coupled Deep Learning Model (MAS-Net), which integrates three-dimensional multivariable data to deliver accurate predictions of the ocean field and the Dipole Mode Index (DMI) up to eight months in advance. Sensitivity and interpretability analyses reveal that subsurface temperature anomalies are critical in improving IOD prediction skill. By integrating subsurface dynamics into a deep learning framework this work hence significantly advances understanding of IOD predictability and provides a robust tool for both operational forecasting and scientific exploration of the IOD. The MAS-Net framework sets a new standard for predictive modelling, offering a versatile approach that can be extended to other climate phenomena.

Key words: Indian Ocean Dipole; Climate prediction; Deep learning; Subsurface temperature; Predictability analysis

https://doi.org/10.1007/s00376-026-5609-4

Article Highlights:

- The MAS-Net is constructed with three-dimensional multivariable data, and the IOD prediction skill is extended to 8 months.
- Sensitivity analyses with MAS-Net reveal that subsurface temperature anomalies play a pivotal role in predicting IOD.
- The three-dimensional sensitive regions for IOD prediction are identified and explained by ocean dynamic processes.

1. Introduction

The Indian Ocean Dipole (IOD), a dominant interannual variability mode in the Indian Ocean, is characterized by a distinct dipole pattern of sea surface temperature anomalies (SSTAs) between the western and eastern regions (Saji et al., 1999; Schott et al., 2009; Webster et al., 1999). Typically, the IOD begins during the boreal summer, peaks in the fall, and declines through the winter, exhibiting a pronounced seasonal phase lock (Saji et al., 2006; Saji et al., 1999). The occurrence of an IOD event can cause extreme weather events, including droughts, floods, and wildfires in adjacent regions (Behera et al., 2005; Meyers et al., 2007), and it also affects areas beyond the Indian Ocean through atmospheric teleconnections (Doi et al., 2020a; Nuncio et al., 2015). Given the backdrop of global warming, the frequency and intensity of these impacts may intensify (Ng et al., 2015). Therefore, there is an urgent need to elucidate IOD physical and dynamical mechanisms and to investigate its predictability.

In recent decades, substantial progress has been made in the understanding of the IOD. The development of the IOD is primarily affected by a positive Bjerknes feedback (Bjerknes, 1969; Hu and Wu, 2025; Xie et al., 1994; Yang et al., 2020), where equatorial zonal wind anomalies modify the thermocline slope, which in turn reinforces the zonal SST gradient and wind anomalies (Li et al., 2003; Rao et al., 2009; Zhang et al., 2022). Furthermore, other internal processes, such as the wind-evaporation-SST feedback (Xie et al., 1994) and the cloud-radiation-SST feedback (Li et al., 2003; Terray et al., 2005), along with external forcings, including the El Niño-Southern Oscillation (ENSO; An, 2004; Cai et al., 2019; Terray et al., 2005; Vecchi et al., 2015), the tropical Atlantic variability (Wang et al., 2009), the Indonesian Throughflow (Tozuka et al., 2007), and stochastic noise (Stuecker et al., 2017), also play a pivotal role in IOD development. These detailed understandings of the underlying mechanisms and have facilitated advancements in numerical simulation (Yao et al., 2016) and extended the IOD prediction skill to approximately five months (Alves et al., 2012; Doi et al., 2020b; Liu et al., 2017). Nevertheless, longer-term prediction of IOD remains a challenge due to the relatively short period of time for which observations are available (Sprintall et al., 2024) and the inaccuracy of the representation of subsurface ocean processes and in numerical models (Chowdary et al., 2016).

Artificial intelligence (AI), particularly deep learning (DL) models, has made significant strides in the field of Earth sciences, demonstrating the predictive performance for high-impact air-sea events that is comparable to, or even surpasses, that of dynamical forecast systems (Bi et al., 2023; Ham et al., 2019; Kochkov et al., 2024; Qin et al., 2022; Shin et al., 2024; Yang et al., 2023; Zhang et al., 2024; Zhou et al., 2023; Liu et al., 2025). The predictive skill of IOD has also been extended by using DL models (Feng et al., 2022; Ling et al., 2022; Liu et al., 2021; Qin et al., 2022; Tao et al., 2024). However, most of these models rely on a "surface-to-point" approach, which maps the two-dimensional data to directly output the Dipole Mode Index (DMI). This approach has two key shortcomings: 1) neglecting the three-dimensional ocean thermal structure may cause models to overlook critical subsurface dynamical processes that are essential for IOD development (Bjerknes, 1969; Xie et al., 1994), which constrains the predictive accuracy of IOD and investigation of different sources of predictability; 2) the "surface-to-point" DL models only output the DMI, whereas similar DMI amplitudes may correspond to different spatial distributions of SSTA (see Fig. 1), making it difficult to assess the weather and climate effects of SSTA associated with the IOD (Endo and Tozuka, 2016; Zhang et al., 2020; Chen et al., 2024). These limitations may hinder the exploration of specific physical mechanisms influencing IOD predictions in DL models, ultimately restricting the model interpretability.

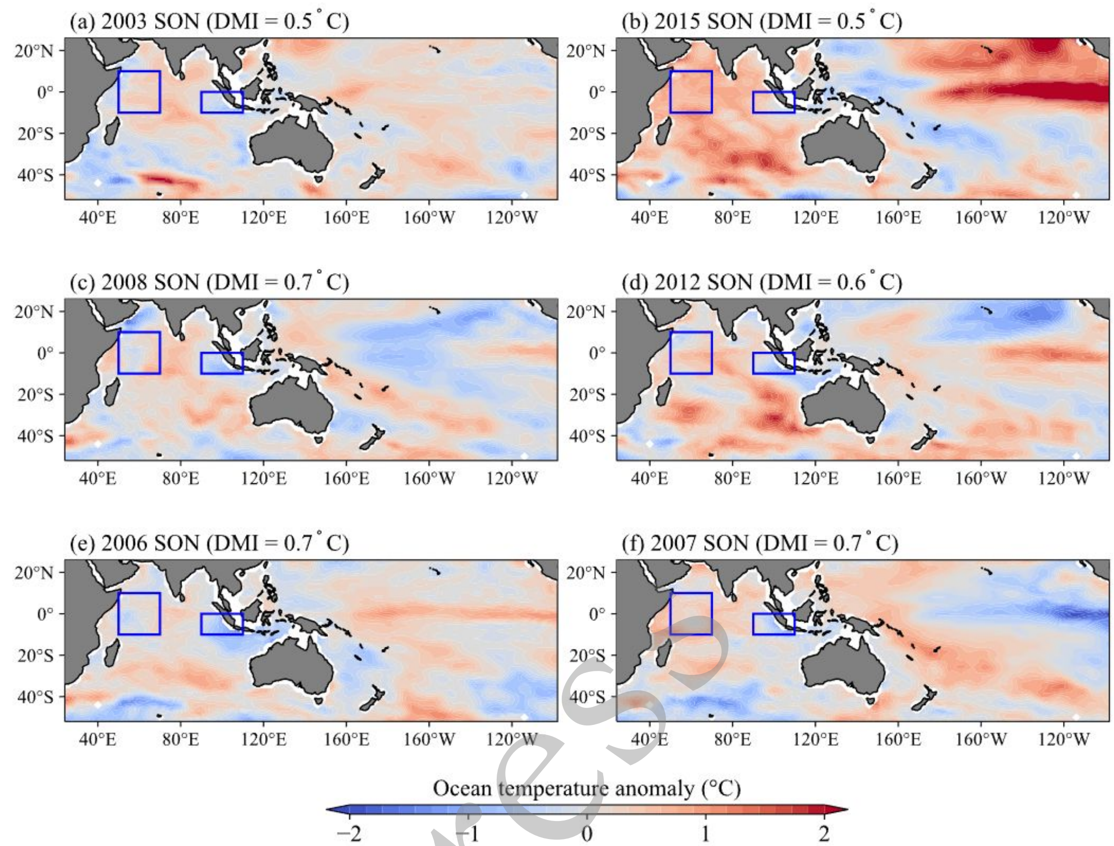

Fig. 1. Sea surface temperature anomalies (SSTAs) and DMI at different years in the Indo-Pacific region, based on GODAS data. Panels show SSTA during 2003 (a), 2015 (b), 2008 (c), 2012 (d), 2006 (e), and 2007 (f) SON. The blue boxes indicate the regions used to calculate the DMI (western pole: 10°S–10°N, 50°–70°E; eastern pole: 10°S–0°, 90°–110°E). While the DMI during these periods are relatively close, the spatial distribution of SSTA shows significant differences, suggesting that despite similar DMI amplitudes, distinct ocean states may lead to varying global impacts.

To address these challenges, this study aims to develop a novel DL-based IOD forecast system that leverages advanced deep learning techniques and integrates three-dimensional oceanic and atmospheric observables from diverse climate simulations and reanalysis datasets. By introducing three-dimensional oceanic data, the effective forecast lead time has been improved to eight months, outperforming current dynamical and DL models. We further employ the sensitivity and interpretability analyses to clarify the specific physical processes that influence IOD predictions, providing deeper insights into the dynamics of IOD. Evidently, this study not only advances predictive skill but also offers new insights into the physical mechanisms underlying IOD, providing a significant contribution to the field of climate prediction.

2. Materials and methods

2.1 Climate simulations and reanalysis dataset

The data used to train Multi-dimensional Air–Sea Coupled Deep Learning Model (MAS-Net) includes three-dimensional ocean temperature anomalies, sea surface height anomalies (SSHA), and wind stress anomalies derived from climate simulations and reanalysis data. We utilize 19 sets of CMIP6 historical simulation datasets (Eyring et al., 2016) from 1850 to 1992 (see Table S1), with a total of 32,604 sample (19 members × 143 years × 12 months) to enable the DL model to learn the underlying physical process. Since CMIP6 dataset provides only 10 m wind speed data, the wind stress $\vec{\tau}$ is calculated using the following formula:

$$\vec{\tau} = \rho C_d \left|\vec{V}\right| \vec{V} (1)$$

where $\rho = 1.225\ kg/m^3$ is the air density and $\vec{V}$ is the 10-m wind speed components, respectively. The drag coefficient C_d equals 0.0012 when $|U| < 11$ m/s and 0.001 × (0.49+0.065$|U|$) when $|U| \geq 11$ m/s.

This study also employs three-dimensional ocean temperature, SSHA and wind stress anomalies data from the SODA reanalysis data (Carton et al., 2008) with the period from 1871 to 1982, as well as from the GODAS reanalysis data (Saha et al., 2010) for 1983 to 2023. All datasets are interpolated to a horizontal spatial resolution of 1° × 1° (covering the region from 24°E to 96°W and 52°S to 28°N) and at corresponding depths (5, 15, 35, 65, 95, 125, and 155 m), resulting in the dimension of 10 × 40 × 120, where the numbers represent the number of variables, latitude, and longitude grid points, respectively. Anomalies are calculated by subtracting the monthly climatological data from 1980 to 2010. Furthermore, each variable at each grid point is standardized using the mean and standard deviation calculated from the training set before inputting into the model.

2.2 NMME hindcasts data

Monthly sea surface temperature (SST) outputs from six dynamic forecast systems derived from the North American Multi-Model Ensemble (NMME) (Kirtman et al., 2014) and their mean (NMME-M) are utilized to evaluate the predictive performance of MAS-Net. The ensemble size for each forecast systems range from 10 to 13 members (Table S3). Anomalies are calculated by subtracting the monthly climatological data from 1980 to 2010.

2.3 MAS-Net model

The MAS-Net model is built upon the SimVP framework (Gao et al., 2022), which ensures reliable convergence and high computational efficiency, thereby reducing training time and memory usage relative to more complex models while maintaining predictive performance (Cao et al., 2024). The model consists of five layers: the input layer, encoder layer, translator layer, decoder layer, and output layer. The encoder layer consists of multiple convolutional layers that extract spatial features from the input data. The translator layer employs U-shape convolutional layers to capture temporal features, while the decoder layer consists of multiple transposed convolutional layers and integrates spatial and temporal features to reconstruct the ocean fields.

Residual connections are added between the encoder and decoder layers to mitigate the gradient vanishing problem (Fig. 2).

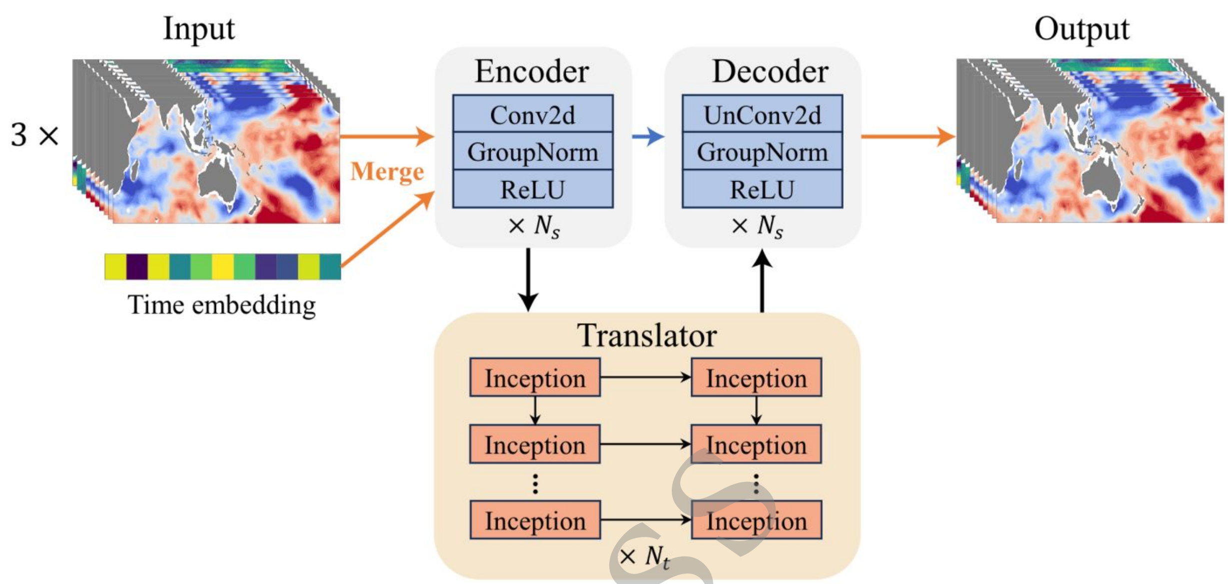

Fig. 2. The structure of MAS-Net. MAS-Net consists of five components: the input layer, encoder layer, translator layer, decoder layer, and output layer. MAS-Net integrates upper ocean temperature anomalies (at depths of 5, 15, 35, 65, 95, 125, and 155 m), SSHA, and wind stress anomalies for the three months, along with time encoding, as inputs to predict air-sea environmental fields for the target month. The encoder layer includes multiple serial convolutional layers that extract key spatial features from the input data. The translator layer uses U-shaped convolutional layers to capture temporal features, while the decoder layer, composed of multiple transposed convolutional layers, integrates spatial and temporal features to reconstruct the air-sea environmental fields. Residual connections are incorporated between the encoder and decoder layers to prevent the gradient vanishing problem.

The upper ocean temperature anomalies, SSHA, and wind stress anomalies for the three months are input to predict the three-dimensional ocean fields at target months, after which the DMI is computed according to its physical definition. To enhance the model's learning of the spatiotemporal dependencies between the input and target months, MAS-Net also incorporates temporal embedding based on the lead time. To effectively predict DMI and air-sea environmental fields, the loss function is as follows:

$$Loss = \frac{1}{M \times N \times C} \sum_{i=1}^{M} \sum_{j=1}^{N} \sum_{k=1}^{C} \left(\hat{y} - y\right)^2 + \left(\hat{y}_{dmi} - y_{dmi}\right)^2 (2)$$

where $\hat{y}$ and y represent the predicted and observed ocean fields, while $\hat{y}_{dmi}$ and y_{dmi} denote the predicted and observed DMI calculated from ocean temperature at depth of 5 m. The M, N, and C denote the number of latitudinal grid points, longitudinal grid points, and variables, respectively, which are 40, 120, and 10. The MAS-Net employs the Adaptive Moment Estimation (Adam) optimization algorithm, with an initial learning rate set at 0.0001, which is updated using a cosine annealing strategy. To improve IOD prediction skill, this study trains separate models for different calendar months and target months. Six ensemble members are generated through different random initial weights, and the ensemble mean is taken as the model output. Subsequently, the bias correction model is fitted to the GODAS-based validation dataset to remove biases in the ensemble mean forecast and is then independently evaluated on the testing dataset (Zhong et al., 2023). The bias correction for each variable at each grid point is formulated as follows:

$$\hat{y}_{bc} = w\hat{y} (3)$$

where $\hat{y}$ and $\hat{y}_{bc}$ denote the uncorrected and bias-corrected time series in the validation set, respectively, and w represents the coefficient estimated using the least-squares method to minimize the discrepancy between $w\hat{y}$ and the observed y. The model excludes an intercept term, as the bias correction aims solely to adjust forecast amplitude while avoiding overcorrection.

When the DMI mapping method is used for model training, the SSTA output from the penultimate layer is further mapped to the DMI through a multilayer perceptron, which serves as the final model output. The optimization process simultaneously minimizes the prediction errors of both the SSTA and DMI, the loss function is as follows:

$$Loss = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(\hat{y}_{ssta} - y_{ssta}\right)^2 + \left(\hat{y}_{dmi} - y_{dmi}\right)^2 \quad (4)$$

where $\hat{y}_{ssta}$ and y_{ssta} represent the predicted and observed SSTA, $\hat{y}_{dmi}$ and y_{dmi} denote the predicted and observed DMI calculated from sea temperature at depth of 5 m.

2.4 Occlusion sensitivity analysis method

The occlusion sensitivity analysis method (Ham et al., 2023) is employed to quantify the sensitivity of each grid point in relation to IOD predictions. Specifically, for each grid point, we assess the impact of removing a 7×7 grid of information centered around it on the prediction. The occlusion sensitivity $O(t,c,w,h)$is a four-dimensional tensor that encompasses time (t), variable (c), longitude (w), and latitude (h). Its calculation formula is as follows:

$$O(t,c,w,h) = |M(X_{in}) - M(X_{in} * Z_7)| \quad (5)$$

where $M(X_{in})$ denotes the predicted DMI from MAS-Net obtained from the input data X_{in}. The $*$ signifies the horizontal convolution operator, while Z_7 indicates that the values of the 7×7 grid points are zero.

3. Results

3.1 Evaluating the predictive skill of MAS-Net

This study develops a Multi-dimensional Air–Sea couple DL model, named MAS-Net, based on multi-layer convolutional neural networks with an encoder-decoder architecture (Gao et al., 2022), to effectively extract complex air-sea dynamic features and predict IOD events (see Fig. 2). Recognizing that the IOD is not an isolated air-sea interaction event within the Indian Ocean and may also be influenced by tropical Pacific Ocean temperature, we set the model domain as the tropical Indo-Pacific region (24°E–96°W, 52°S–28°N). Simultaneously, to consider the role of subsurface information in the IOD prediction, the historical three-month upper ocean temperature anomalies at the depths of 5, 15, 35, 65, 95, 125, and 155 m, as well as sea surface height anomalies (SSHA) and wind stress anomalies, are input into MAS-Net for training. Due to the limited availability of monthly reanalysis data, the training process is divided into two stages. In the initial stage, the model parameters are pre-trained using 19 sets of the Coupled Model Intercomparison Project Phase 6 (CMIP6) historical simulations datasets (see Table S1). Subsequently, the model is fine-tuned with Simple Ocean Data Assimilation (SODA) reanalysis datasets. Finally, the predictive skill of MAS-Net is evaluated using Global Ocean Data Assimilation System (GODAS) datasets from 1993 to 2023 (see Table S2 for details on the data splitting strategy).

The spatial distribution of correlation coefficients between the predicted air-sea fields from MAS-Net and observation data demonstrate good predictive capabilities in the Indo-Pacific region, especially in the equatorial area (Fig. 3), with the majority of grid points exceeding the correlation coefficient of 0.5 up to 7-month lead forecasts. Figure 4 shows a comparison of prediction performance between MAS-Net and six dynamical forecast systems (Table S3) from the North American Multi-Model Ensemble project (NMME) and their ensemble mean (NMME-M) in predicting the DMI. The latter is defined as the SSTA difference between the tropical western Indian Ocean (10°S–10°N, 50°–70°E) and the southeastern Indian Ocean (10°S–0°, 90°–110°E). With a correlation coefficient threshold of 0.5 for effective prediction, MAS-Net achieves skillful DMI forecasts for all calendar months at lead times of up to eight months (black lines in Figs. 4a,b). In contrast, state-of-the-art dynamical forecast systems exhibit useful skill only up to four to five months (red line for NMME-M; other colored lines for different ensemble systems), and the persistence model is limited to lead times of only two months. In addition, for the September-October-November (SON) DMI, MAS-Net maintains its 8-month prediction lead time (Fig. 5).

Notably, as lead times exceed four months, the advantages of MAS-Net over dynamic forecast systems become increasingly pronounced. This indicates that MAS-Net achieves a predictive skill comparable to the state-of-the-art models (Ling et al., 2022; Liu et al., 2021; Ratnam et al., 2025).

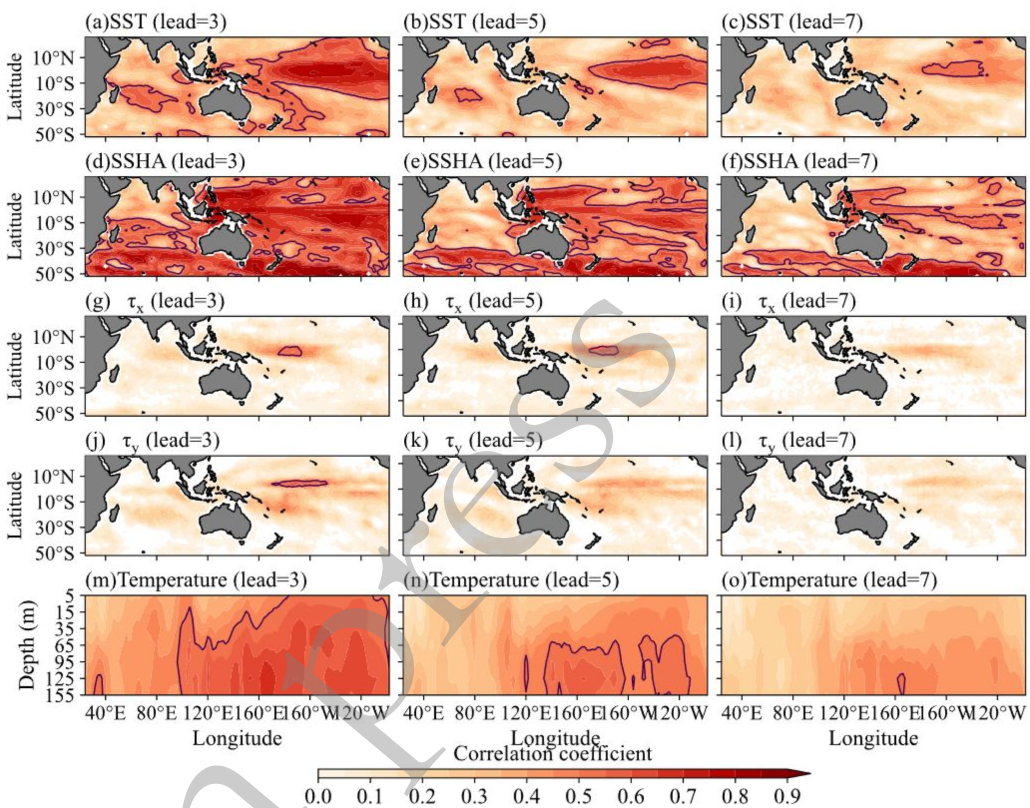

Fig. 3. Spatial distribution of correlation coefficients by MAS-Net at lead times of 3, 5, and 7 months for SSTA (a-c), SSHA (d-f), τx (g-i), τy (j-l), and ocean temperature anomalies in the zonal–depth sections between 10°S and 10°N (m-o). The contour line denotes a correlation coefficient of 0.5.

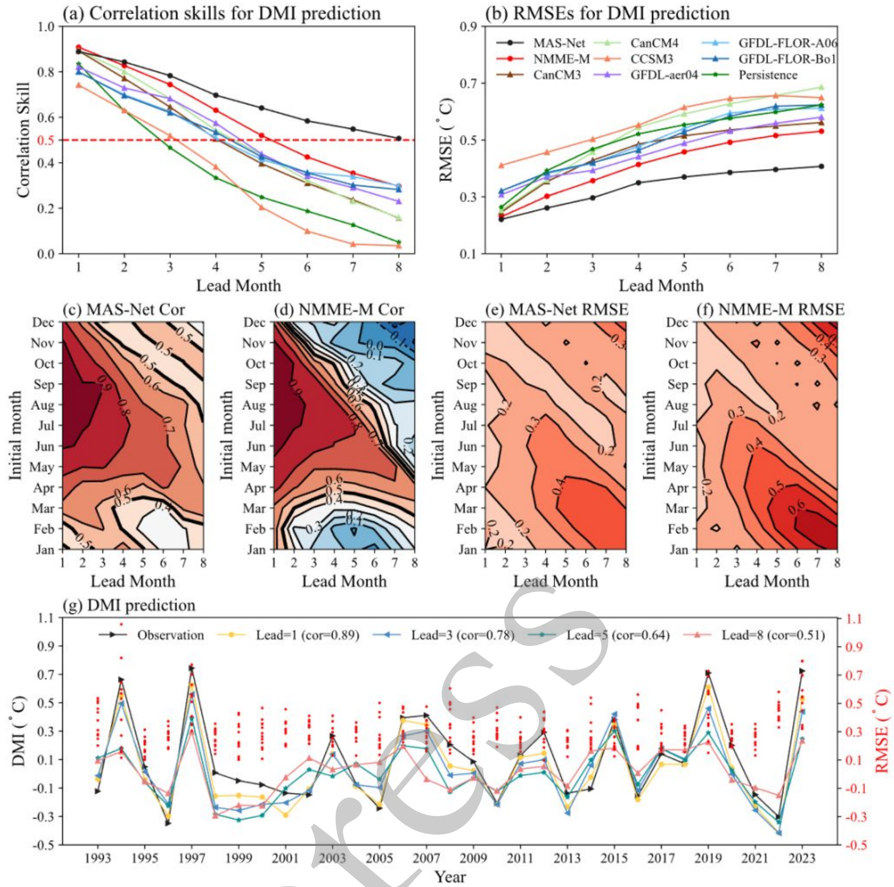

Fig. 4. Predictive skill of MAS-Net in forecasting IOD. (a) Correlation coefficients and (b) Root Mean Square Error (RMSE) of the DMI for all months at different lead times, based on MAS-Net (black lines) and six dynamical forecast systems from the North American Multi-Model Ensemble and their ensemble mean (NMME-M). MAS-Net is validated from 1993 to 2023, while NMME are validated from 1993 to 2018. Testing MAS-Net over the same period as NMME-M (1993–2018) yields very similar predictive performance, so these results are not shown. The red dashed line corresponds to a correlation coefficient of 0.5. The green lines represent the prediction skills of the persistence model. (c-f) Correlation coefficients and RMSE for different initial months and target months, based on MAS-Net and NMME-M. (g) Observation and predicted annual average DMI based on MAS-Net at lead times of 1, 3, 5, and 7 months. Red dots represent the RMSE between the predicted DMI and observation for each month of that year.

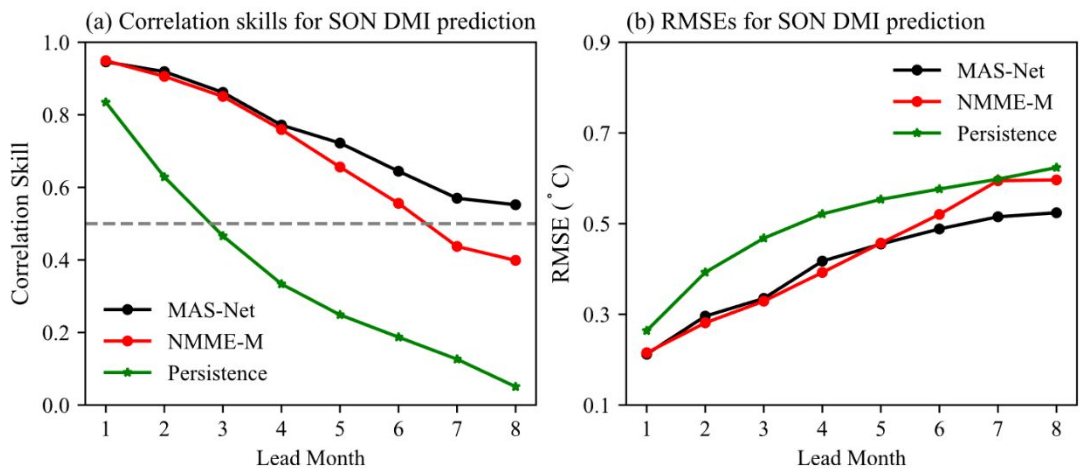

Fig. 5. (a) Correlation coefficients and (b) Root Mean Square Error (RMSE) of the September-October-November (SON) DMI for all months at different lead times, based on MAS-Net and NMME-M. MAS-Net is validated from 1993 to 2023, while NMME-M are validated from 1993 to 2018. The difference in forecast skill between MAS-Net and NMME-M is statistically significant at the 90% confidence level, based on the bootstrap test. The green lines represent the prediction skills of the persistence model. The gray dashed line corresponds to a correlation coefficient of 0.5.

3.2 Comparison of Different DL Prediction Models

Currently, DL model for the IOD prediction are primarily based on the "surface-to-point" (Ling et al., 2022; Liu et al., 2021; Tao et al., 2024) and "surface-to-surface" (Wang et al., 2024) strategies: the former uses sea surface data to directly mapping DMI, and the latter uses the surface data to predict the surface data. In contrast, MAS-Net applies the "volume-to-volume" strategy, which uses three-dimensional data to predict three-dimensional data. To compare the different DL prediction strategies, additional DL models have been developed. Firstly, using the "surface-to-surface" strategy, the MAS-Net$_{surface}$ is constructed, with a systematically lower predictive skill than MAS-Net, showing an effective forecast lead time limited to 7 months (see black dotted lines in Figs. 6a,b). This highlights the importance of incorporating subsurface temperature data to enhance IOD prediction performance (Liu et al., 2022). Furthermore, we construct other two models using the “surface-to-point” and “volume-to-point” mapping DMI strategies (denoted as MAS-Net$_{map+surface}$ and MAS-Net$_{map}$; output the SST map and DMI index). In terms of DMI forecast skill, MAS-Net$_{map}$ (MAS-Net$_{map+surface}$) outperforms MAS-Net (MAS-Net$_{surface}$). However, when examining the SSTA distribution in the equatorial Indian Ocean output by MAS-Net and MAS-Net$_{map}$, we find that the point mapping DMI strategies struggle to predict the SSTA evolution (Figs. 6c-g). Here, SSTA refers to the SSTA fed into the input perceptron layer. This suggests that when the DMI is predicted by point mapping models, it is difficult to capture specific physical processes, limiting the interpretability of the DL model results. We therefore select the MAS-Net to predict the IOD rather than the models based on the point mapping strategy, even though these have a higher DMI prediction skill.

3.3 Evaluating Multi-dimensional Prediction during Strong IOD Events

To further demonstrate the efficacy of MAS-Net, we assess the evolution of the predicted three-dimensional ocean state in the Indian Ocean during the 2019 strong positive IOD (pIOD) event. Figure 7a (light red shading) shows that the MAS-Net well predicts the variability of the DMI, although the predicted amplitude is slightly weaker than observations. More interestingly, the MAS-Net also accurately captures the evolution processes of the negative temperature anomalies in the eastern pole and the positive anomalies in the western pole, as well as the west-deep, east-shallow thermocline structure in the equatorial region (Figs. 7c-h). A similar assessment is also done for the 2022 strong negative IOD (nIOD) event, which illustrates that the developments of the sea temperature anomalies and thermocline are well predicted by the MAS-Net (Figure 8).

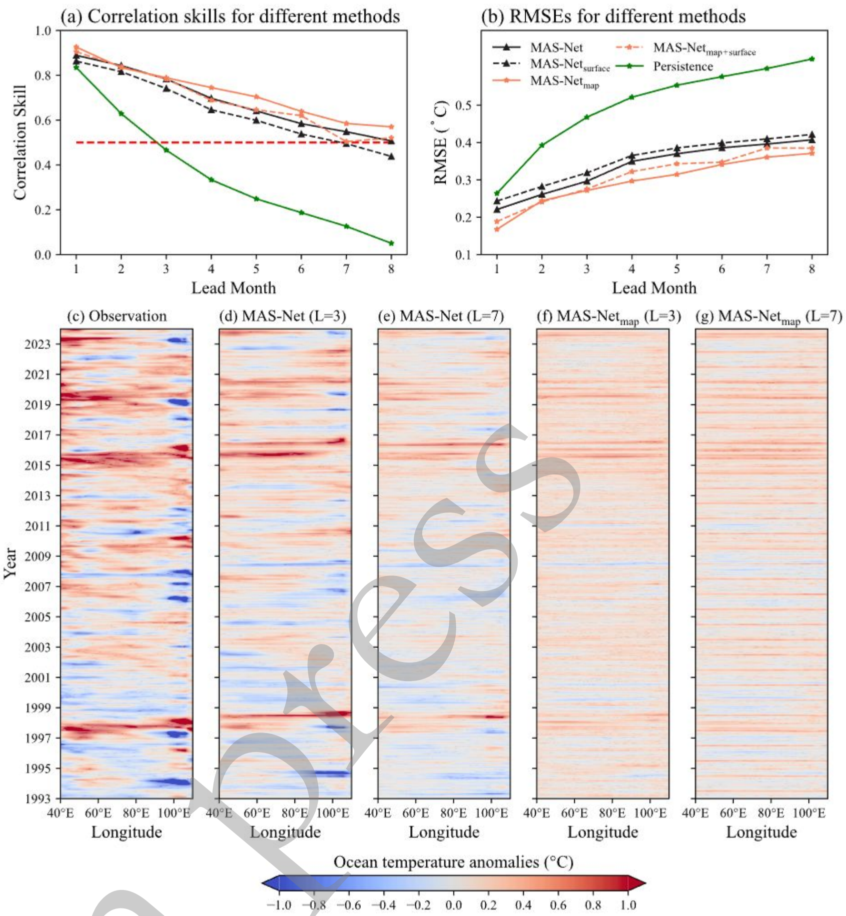

Fig. 6. Correlation coefficients (a) and root mean square error (b) of the DMI for all months at different lead times, based on different prediction approaches. Where MAS-Net and MAS-Net$_{surface}$ are training with all data and only surface data (SSTA, SSHA, τ_x, τ_y), MAS-Net$_{map}$ and MAS-Net$_{map+surface}$ are using the “volume to point” and “surface to point” mapping DMI method. The results in (a–b) are statistically significant at the 90% confidence level, as determined by a bootstrap test. The green lines represent the prediction skills of the persistence model. The red dashed line marks a correlation coefficient of 0.5. (c-g) Evolution of SSTA along the equator (10°S-10°N) based on MAS-Net and MAS-Net$_{map}$ at lead times of 3 and 7 months.

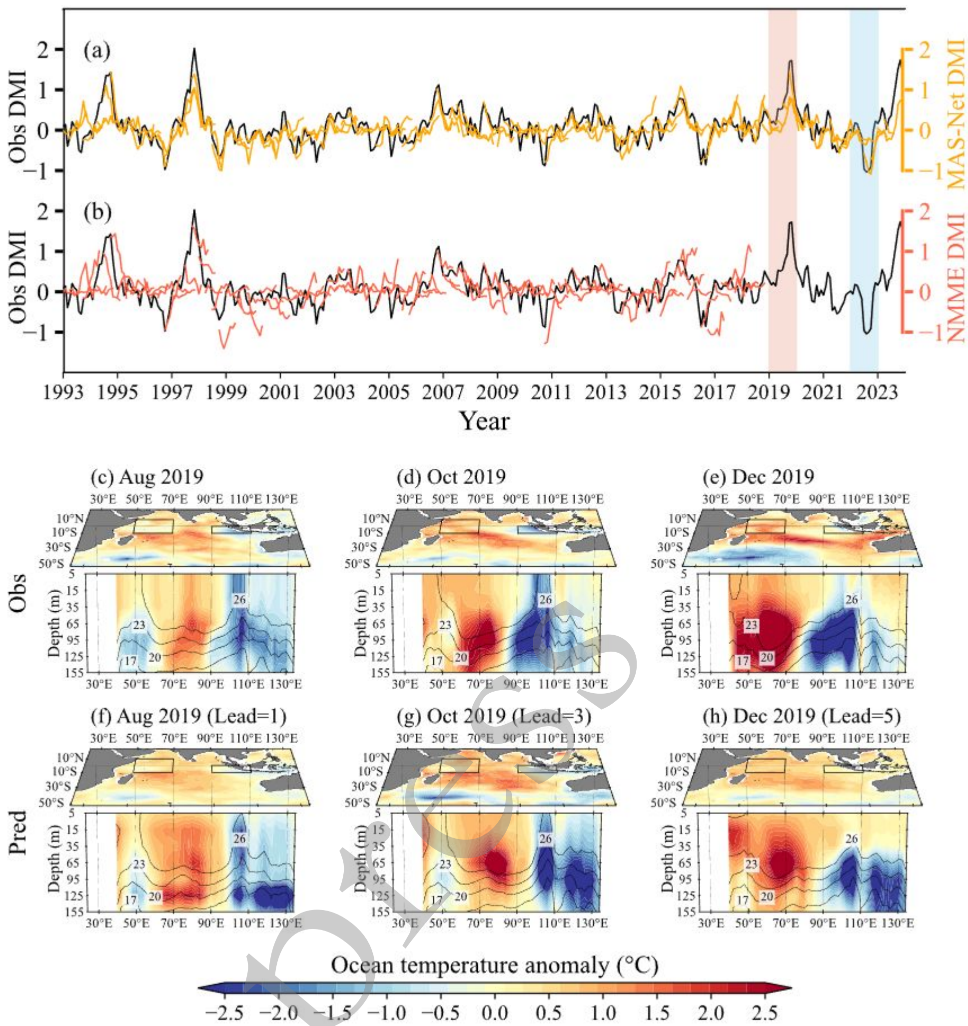

Fig. 7. Evaluation of the IOD prediction. (a-b**)** Observations (black lines) and prediction DMI for the next 8-month when initial in Dec, Mar, Jun, and Sep, based on MAS-Net (orange lines) and NMME-M (red lines). The light red and blue shading indicates the extreme IOD events in 2019 and 2022. (c-h**)** Observed and predicted sea temperature anomalies at lead times of 1, 3, and 5 months, based on MAS-Net. The zonal section represents mean temperature anomalies across the equator region (10°S–10°N; contours denote temperature). The black boxes indicate the regions used to calculate the DMI (western pole region: 10°S–10°N, 50°–70°E; eastern pole region: 10°S–0°, 90°–110°E).

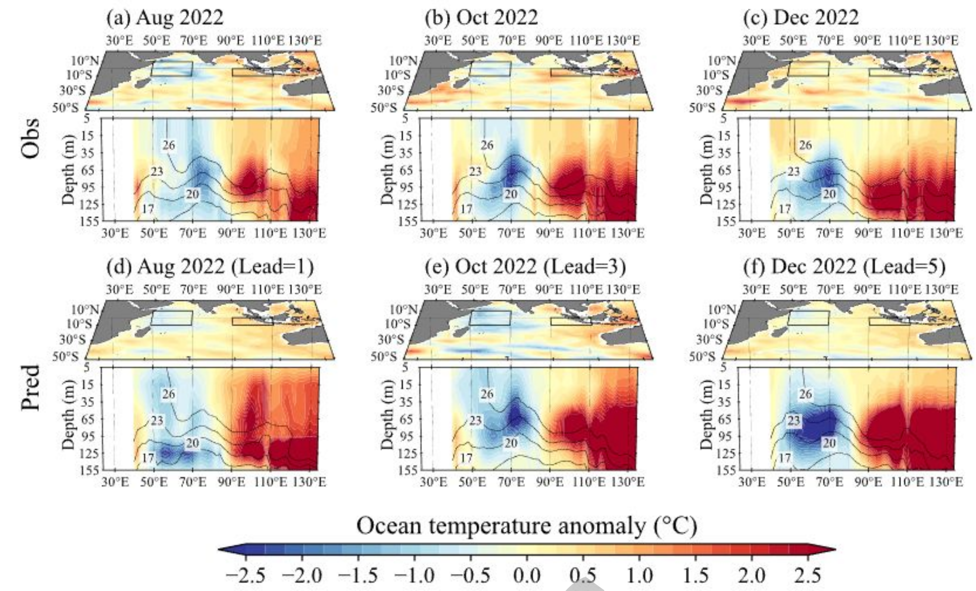

Fig. 8. Observed and predicted sea temperature anomalies for 2022 nIOD events at lead times of 1, 3, and 5 months, based on MAS-Net. The zonal section represents mean temperature anomalies across the equator region (10°S–10°N; contours denote temperature).

3.4 Subsurface temperature anomalies influence the intensity of predicted IOD

To further identify the key factors influencing IOD forecasts, MAS-Net is used to conduct an input variable sensitivity experiment for the target months of SON (hereafter referred to as the target months). To quantify the contribution of a specific variable, we set its value to zero at all grid points while keeping the other variables unchanged. A perturbed forecast is then generated from this modified input, and the contribution of that variable is subsequently quantified by the standardized difference between this perturbed forecast and the unperturbed control forecast. Here, we mainly focus on the strong IOD events, which are defined as when the absolute value of the DMI exceeds one standard deviation for five consecutive months. In total, five pIOD events (1994, 1997, 2006, 2015, 2019) and five nIOD events (1996, 1998, 2010, 2016, 2022) are identified. For pIOD and nIOD events (Fig. 9a), the contributions of SSTA are 20.4% and 11.9%, respectively. However, subsurface temperature anomalies play an even more important role, especially at 35 m, where they account for 23.4% for pIOD events and 21.0% for nIOD events.

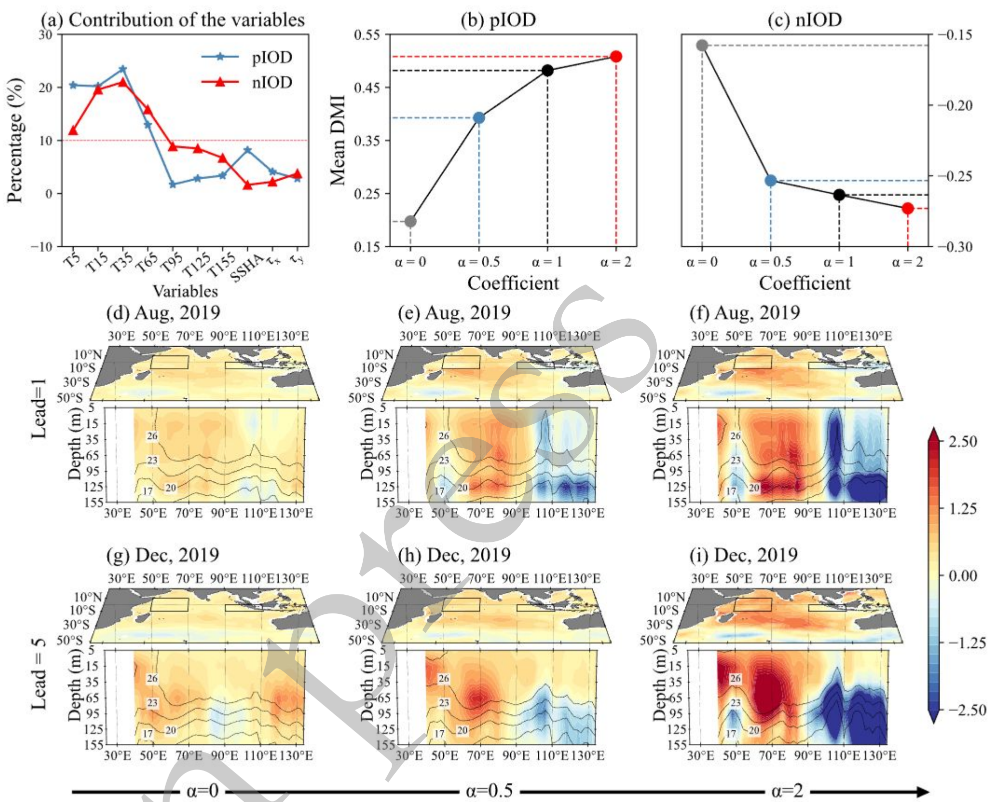

Fig. 9. The key factor of IOD prediction**.** (a) Contribution of different variables to the predictions of positive and negative IOD events across all lead times. (b-c) Predicted annual average DMI by MAS-Net for positive and negative IOD events, with subsurface sea temperature anomalies at depths of 15, 35, 65, 95, 125, and 155 m multiplied by coefficients α of 0 (gray dots), 0.5 (blue dots), 1 (black dots), and 2 (red dots). Here, $\alpha = 1$ indicates no change to the temperature anomalies. (d-i) Predicted sea temperature anomalies for the 2019 pIOD event by MAS-Net, with α values from 0 to 2.

We also conduct sensitivity experiments by changing amplitudes of the subsurface temperature anomalies in the MAS-Net input data to predict IOD events for the target months. In particular, the subsurface sea temperature anomalies (at depths of 15, 35, 65, 95, 125, and 155 m, excluding 5 m) are multiplied by four coefficients (α = 0, 0.5, 1.0, 2.0), and then re-predicted for positive and negative IOD events, where α = 1.0 corresponds to the control prediction. As illustrated in Figs. 9b,c, an increase in the amplitude of subsurface temperature (from α = 0 to 2.0) is accompanied by an enhancement (weakened) in the mean value of the predicted DMI for positive (negative) IOD events. This reflects a robust correlation between subsurface temperature anomalies and the intensity of IOD events, whereby increased subsurface temperature anomalies may contribute to more strong IOD events (Sun et al., 2022).

Taking the 2019 pIOD event as an example, an increase in subsurface temperature anomaly magnitudes leads to stronger SSTA in the Indian Ocean. From the equatorial zonal section, it is found that as the lead time extends from one to five months, absence of subsurface temperature anomalies (α = 0) only causes a minimal SSTA development, failing to form a pIOD event. However, when α = 0.5 and 2, development of sea temperature anomalies can lead to a pIOD event, with the latter showing broader and stronger temperature anomalies and more pronounced inclined isotherms (Figs. 9d-i). Similar sensitivity experiments were conducted for the strong nIOD event in 2022, which show that stronger subsurface temperature anomalies also lead to more intense nIOD events (Fig. 10). This result was consistent with previous findings, which suggested that incorporating subsurface ocean observations into dynamical prediction systems could enhance the forecast skill of the IOD (Doi et al., 2017).

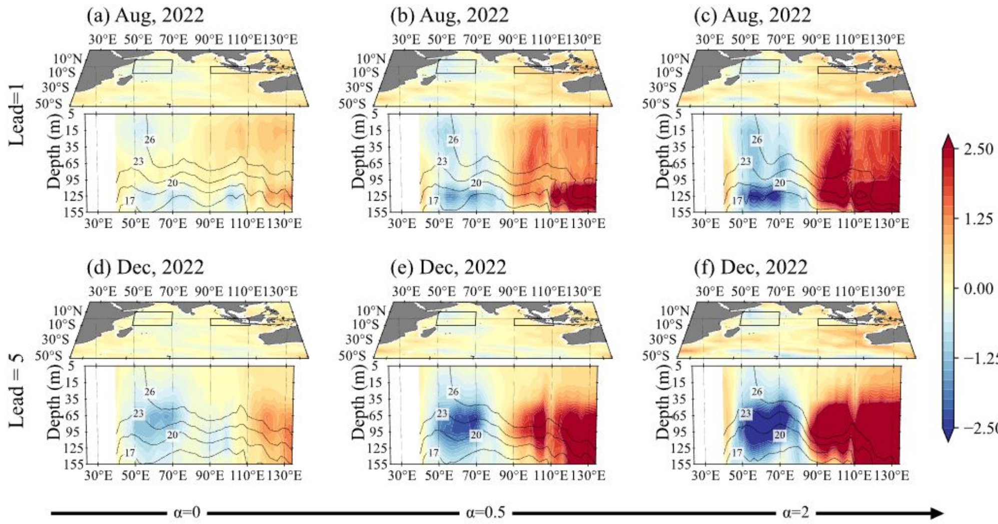

Fig. 10. Predicted sea temperature anomalies for the 2022 nIOD event by MAS-Net, with α values from 0 to 2.

3.5 Key regions for IOD forecast and physical interpretability

The occlusion sensitivity and interpretability analysis method (Ham et al., 2023) is utilized to investigate the key regions of temperature anomalies influencing IOD predictions, in which the surrounding 7×7 grid points of each temperature anomaly variable are set to zero, and the predicted DMI of MAS-Net is compared to the control prediction. Their difference is defined as the sensitivity of this grid point: a larger value means more sensitive. For pIOD events (Figs. 11a,c), the result indicates that at a lead time of 3 months, the sensitivity regions for sea temperature anomalies are primarily concentrated in the upper (5–65 m) tropical Indian Ocean, especially in the tropical eastern Indian Ocean at depth of 15m, suggesting that the sea temperature anomalies in these regions play key roles in the prediction of pIOD events 3 months in advance.

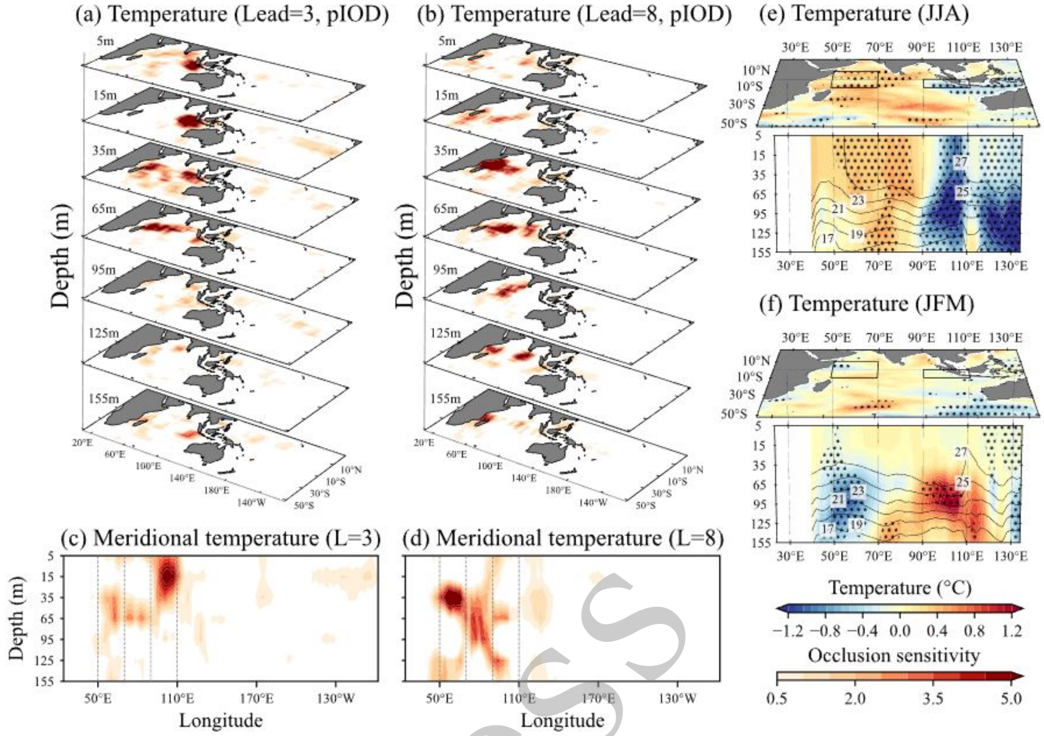

Fig. 11. Sensitivity regions for predicting pIOD events. (a-d) Sensitivity areas of ocean temperature anomalies for predicting five pIOD events (1994, 1997, 2006, 2015, 2019) at lead times of 3 and 8 months. The meridional temperature anomaly indicates that the zonal depth sections on the equator. (e-f) Composite temperature anomalies at lead times of 3 months (initialized in June, July, and August) and 8 months (initialized in January, February, and March) that forecast the September–October–November (SON) DMI. The dotted areas denote regions that are statistically significant at the 90% confidence level. The vertical gray dashed lines, from left to right, correspond to 50°E, 70°E, 90°E, and 110°E.

To evaluate the impact of the different sensitive regions on pIOD prediction, we remove the input subsurface temperature anomalies in the areas with occlusion sensitivity exceeding 1.0 and re-forecast the events at a 3-month lead. The MAS-Net predictions are then compared to the corresponding control predictions and their difference (former minus latter) are shown in Figs.

12a-c. From September to November, the positive SSTA in the eastern pole region gradually intensifies, while the negative one enhances in the western pole region, implying that the absence of sensitive region information impedes the development of pIOD events. Further analysis of the initial input data (the summer temperature anomaly of five pIOD years; Fig. 11e) reveals that the isotherms exhibit west-deep, east-shallow structure and the negative temperature anomaly exists in the most sensitive region. This means that in the sensitivity experiments, setting the sensitive region to zero will lead to an increase in temperature in the subsurface eastern Indian Ocean, which enhances the upwelling of warmer waters to the surface and thereby inhibits the formation of the pIOD event (Deshpande et al., 2014; Gnanaseelan et al., 2024; Nakazato et al., 2021; Zhang et al., 2022).

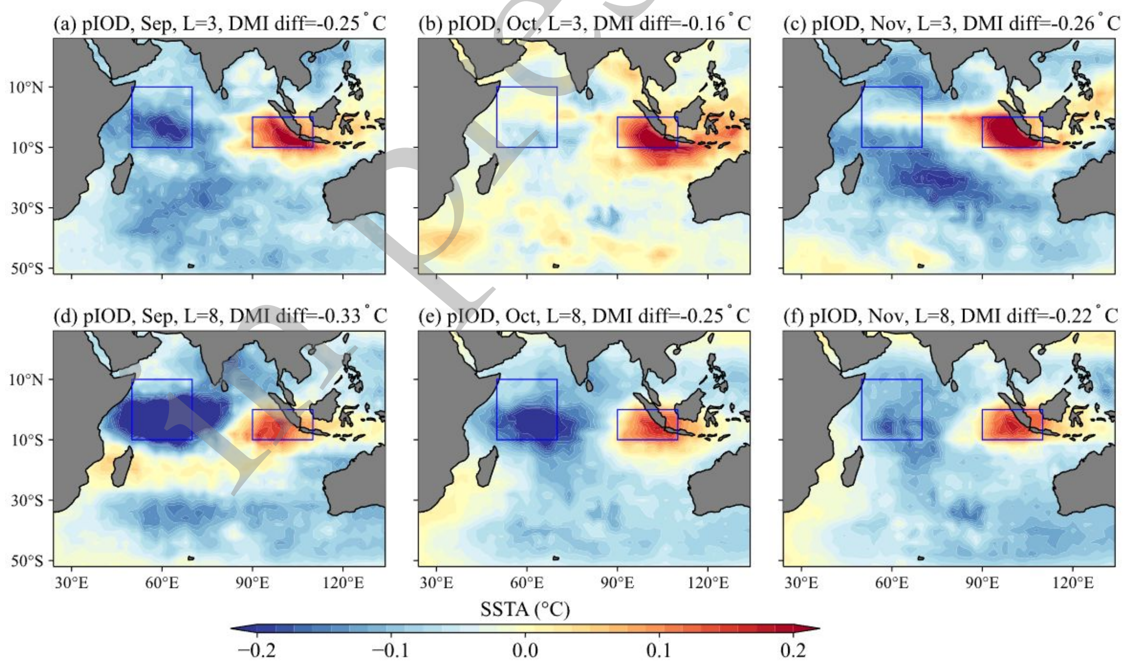

Fig. 12. The influence of occlusion-sensitive regions for pIOD prediction. The difference between the SSTA predicted by MAS-Net with sensitive regions occluded and the control predictions for five pIOD events at lead times of 3 months (a-c) and 8 months (d-f). For lead time of 3 and 8 months, MAS-Net is initialized in June–August and January–March, respectively.

Considering that MAS-Net can improve the prediction skill for IOD events up to 8 months, we therefore apply the occlusion sensitivity method to recognize the sensitive regions for sea temperature anomalies when predicting five pIOD events at lead time of 8 months (Figs. 11b,d). It is found that the sensitive region shifts to a deeper depth, particularly at 35 m in the western equatorial Indian Ocean. Similarly, the input temperature anomalies in these relatively sensitive regions are removed and the five pIOD events are re-predicted 8 months in advance. In Figs. 12d-f, similar to the situation at the 3-month lead time, the pronounced positive anomalies exist in the eastern equatorial pole region, while the western pole exhibits the opposite pattern. As shown in Figure 11f, the initial input data exhibit a west-shallow, east-deep isotherm structure. In this situation, setting the sensitive region to zero will cause the inclined structure of isotherms to become weak and in turn weaken the eastward-propagating ocean waves, which hinders the transition to negative SSTA in the eastern Indian Ocean, and ultimately limits the formation of pIOD events (Feng et al., 2003; Rao et al., 2002; Zhang et al., 2022).

The same experiments are performed for nIOD events, and the results show that the sensitive regions identified through the occlusion sensitivity method are primarily concentrated in the subsurface (Fig. 13). As shown in Fig. 14, removing the initial subsurface temperature anomalies in the relatively sensitive regions leads to substantially weaker predicted nIOD events compared with the control simulation. We note that the reduced negative SSTA over the eastern Indian Ocean in Fig. 14b differs slightly from other experiments, likely because the sensitivity analysis and the occlusion experiments are based on different model configurations. Nevertheless, this discrepancy does not affect our main conclusion that subsurface temperature anomalies in key regions can promote the development of nIOD events (Fang et al., 2024; Huang et al., 2024; Lu et al., 2018).

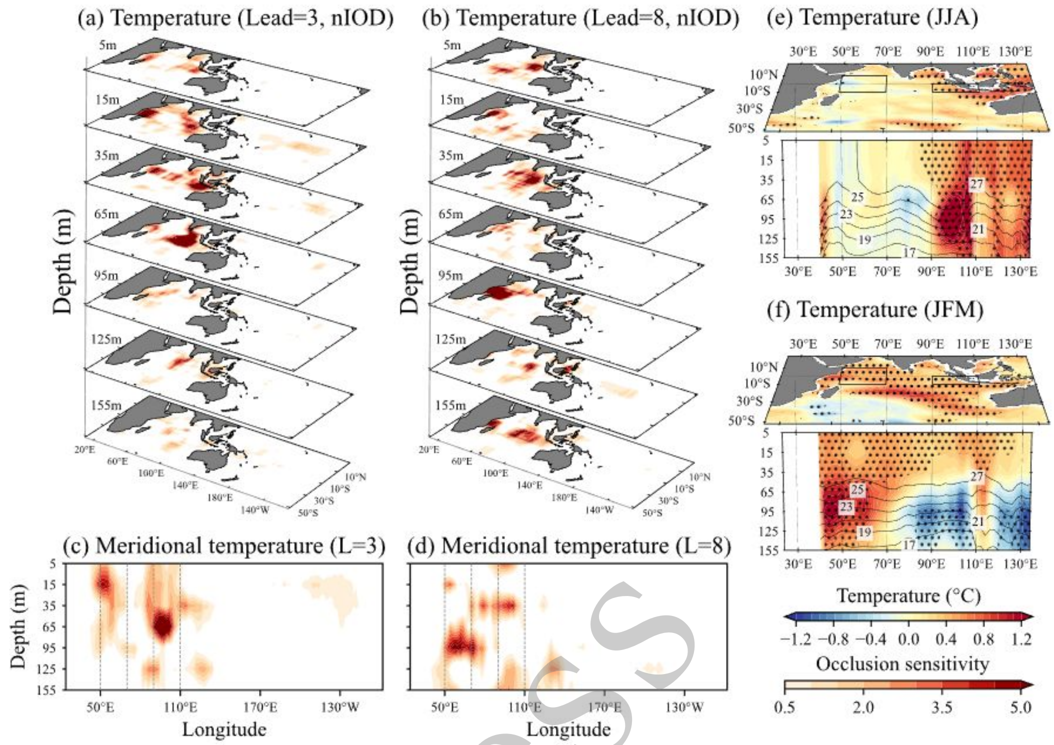

Fig. 13. (a-d) Sensitivity areas of sea temperature anomalies for predicting five nIOD events (1996, 1998, 2010, 2016, 2022) at lead times of 3 and 8 months. The meridional temperature anomaly indicates that the zonal depth sections on the equator. (e-f) Composite of temperature anomaly at lead times of 3 months (initialized in June, July, and August) and 8 months (initialized in January, February, and March) that forecast the September–October–November (SON) DMI. The dotted areas denote regions that are statistically significant at the 90% confidence level. The vertical gray dashed lines, from left to right, correspond to 50°E, 70°E, 90°E, and 110°E.

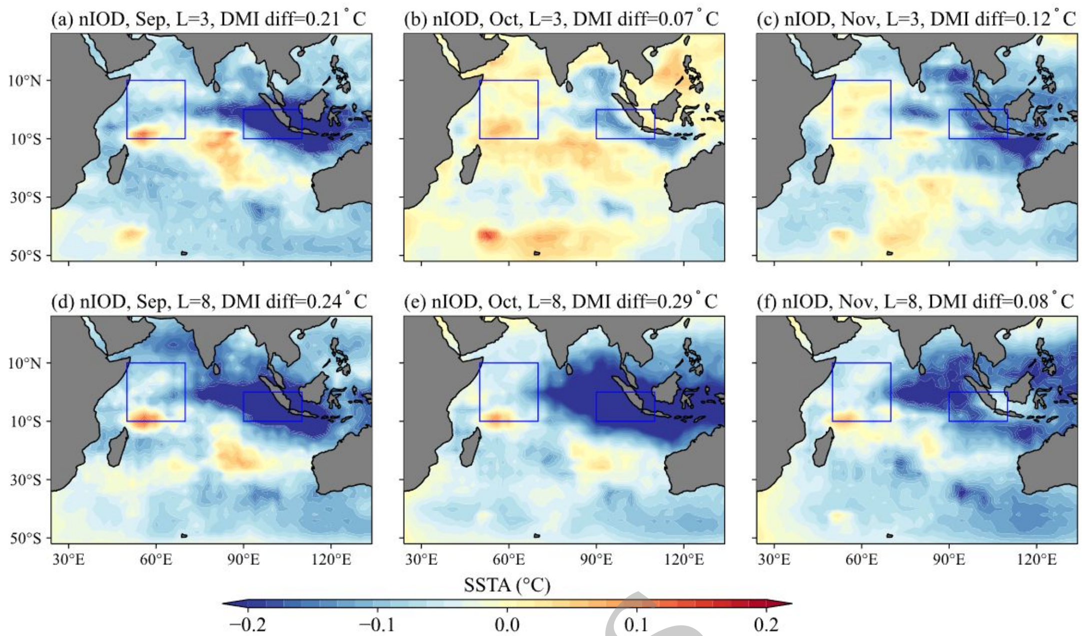

Fig. 14. The different between the SSTA predicted by MAS-Net with sensitive regions occluded and the control predictions for five nIOD events at lead times of 3 months (a-c) and 8 months (d-f). For lead time of 3 and 8 months, MAS-Net is initialized in June–August and January–March, respectively.

4. Discussion and conclusion

By employing advanced deep learning algorithms, this study constructs a Multi-dimensional air–sea Coupled model (MAS-Net), which is able to extend the skillful prediction lead time for IOD events to eight months, outperforming the dynamic forecast systems and other available DL models. To the best of our knowledge, this is the first study to construct a DL model for IOD prediction that uses volume input data. The improvement in MAS-Net's predictions primarily stems from incorporating subsurface sea temperature anomaly information, which enables a more comprehensive representation of the three-dimensional oceanic dynamical evolution. Furthermore, through the sensitivity experiments, the importance of subsurface temperature anomalies in predicting extreme IOD events is highlighted: such temperature anomalies influence the evolution and intensity of IOD events. Absence of subsurface input data will hinder MAS-Net's ability to simulate essential physical processes, making it difficult to predict IOD events. However, this study also has limitations, as it does not explicitly examine the influence of ENSO on IOD prediction; this issue will be addressed in future work. In the further, this study not only introduces a novel framework for predictability research but also shows that DL models can be used in sensitivity experiments to identify key factors influencing prediction skill and improve our understanding of the underlying physical processes.

Acknowledgments. This study was supported by the National Key R&D Program of China (2024YFB4207201), National Natural Science Foundation of China (42476192, 42288101 and 42030410), Jiangsu Innovation Research Group (JSSCTD 202346) and the Postgraduate Research and Practice Innovation Program of Jiangsu Province (422003282).

Data availability. The data related to this study can be downloaded from: CMIP6 dataset, https://esgf-node.llnl.gov/search/cmip6/; SODA dataset, http://iridl.ldeo.columbia.edu/SOURCES/.CARTON-GIESE/.SODA/; GODAS dataset, https://psl.noaa.gov/data/gridded/data.godas.html; NMME dataset, http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/. The deep learning prediction is done using the open-source framework of PyTorch provided by Facebook (https://pytorch.org/). The code of the MAS-Net model can be found at https://doi.org/10.5281/zenodo.14930039.

REFERENCES

Alves, O., H. H. Hendon, L. Shi, J.-J. Luo, M. Balmaseda, and D. Anderson, 2012: How Predictable is the Indian Ocean Dipole? Monthly Weather Review, 140, 3867-3884.

An, S.-I., 2004: A dynamic link between the basin-scale and zonal modes in the tropical Indian Ocean. J Theoretical Applied Climatology, 78, 203-215.

Behera, S. K., J.-J. Luo, S. Masson, P. Delecluse, S. Gualdi, A. Navarra, and T. Yamagata, 2005: Paramount Impact of the Indian Ocean Dipole on the East African Short Rains: A CGCM Study. J Climate, 18, 4514-4530.

Bi, K., L. Xie, H. Zhang, X. Chen, X. Gu, and Q. Tian, 2023: Accurate medium-range global weather forecasting with 3D neural networks. Nature, 619, 533-538.

Bjerknes, J., 1969: Atmospheric Teleconnections from the Equatorial Pacific1. Monthly Weather Review, 97, 163-172.

Cai, W., and Coauthors, 2019: Pantropical climate interactions. Science, 363, eaav4236.

Cao, X., C. Liu, S. Zhang, and F. Gao, 2024: A method for predicting high-resolution 3D variations in temperature and salinity fields using multi-source ocean data. Journal of Marine Science and Engineering, 12, 1396.

Carton, J. A., and B. S. Giese, 2008: A Reanalysis of Ocean Climate Using Simple Ocean Data Assimilation (SODA). Monthly Weather Review, 136, 2999-3017.

Chen, M., M. Collins, J.-Y. Yu, X. Wang, L. Zhang, and C.-Y. Tam, 2024: Emerging influence of the Australian Monsoon on Indian Ocean interannual variability in a warming climate. Npj Clim Atmos Sci, 7, 319.

Chowdary, J. S., A. Parekh, G. Srinivas, C. Gnanaseelan, T. S. Fousiya, R. Khandekar, and M. K. Roxy, 2016: Processes Associated with the Tropical Indian Ocean Subsurface Temperature Bias in a Coupled Model. Journal of Physical Oceanography, 46, 2863-2875.

Deshpande, A., J. Chowdary, and C. Gnanaseelan, 2014: Role of thermocline–SST coupling in the evolution of IOD events and their regional impacts. Clim Dynam, 43, 163-174.

Doi, T., A. Storto, S. K. Behera, A. Navarra, and T. Yamagata, 2017: Improved prediction of the Indian Ocean dipole mode by use of subsurface ocean observations. J Climate, 30, 7953-7970.

Doi, T., S. K. Behera, and T. Yamagata, 2020a: Wintertime Impacts of the 2019 Super IOD on East Asia. Geophys Res Lett, 47, e2020GL089456.

——, 2020b: Predictability of the Super IOD Event in 2019 and Its Link With El Nino Modoki. Geophys Res Lett, 47, e2019GL086713.

Endo, S., and T. Tozuka, 2016: Two flavors of the Indian Ocean dipole. Clim Dynam, 46, 3371-3385.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, 2016: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geosci Model Dev, 9, 1937-1958.

Fang, Y., S. W. Sun, Y. C. Zu, J. H. Wang, and L. Feng, 2024: Characteristics and triggering mechanisms of early negative Indian Ocean Dipole. Acta Oceanol Sin, 43, 59-65.

Feng, M., and G. Meyers, 2003: Interannual variability in the tropical Indian Ocean: a two-year time-scale of Indian Ocean Dipole. Deep Sea Research Part II: Topical Studies in Oceanography, 50, 2263-2284.

Feng, M., and Coauthors, 2022: Predictability of sea surface temperature anomalies at the eastern pole of the Indian Ocean Dipole—using a convolutional neural network model. Frontiers in Climate, 4, 925068.

Gao, Z., C. Tan, L. Wu, and S. Z. Li, 2022: Simvp: Simpler yet better video prediction. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 3170-3180.

Gnanaseelan, C., R. Kakatkar, S. Anila, S. Mohapatra, A. Parekh, and J. Chowdary, 2024: Role of strong subsurface mode on the anomalous basin-wide surface warming of the Tropical Indian Ocean in 2019–2020. Journal of Earth System Science, 133, 143.

Ham, Y.-G., J.-H. Kim, S.-K. Min, D. Kim, T. Li, A. Timmermann, and M. F. Stuecker, 2023: Anthropogenic fingerprints in daily precipitation revealed by deep learning. Nature, 622, 301-307.

Ham, Y. G., J. H. Kim, and J. J. Luo, 2019: Deep learning for multi-year ENSO forecasts. Nature, 573, 568-572.

Hu, X., and B. Wu, 2025: Coastal thermocline feedback incorporating offshore advection: A key driver of Indian Ocean Dipole growth. J Climate.

Huang, K., and Coauthors, 2024: Diversity of strong negative Indian Ocean dipole events since 1980: Characteristics and causes. Clim Dynam, 62, 2017-2040.

Kirtman, B. P., and Coauthors, 2014: THE NORTH AMERICAN MULTIMODEL ENSEMBLE Phase-1 Seasonal-to-Interannual Prediction; Phase-2 toward Developing Intraseasonal Prediction. B Am Meteorol Soc, 95, 585-601.

Kochkov, D., and Coauthors, 2024: Neural general circulation models for weather and climate. Nature, 632, 1060-1066.

Li, T., B. Wang, C. P. Chang, and Y. Zhang, 2003: A Theory for the Indian Ocean Dipole–Zonal Mode*. Journal of the Atmospheric Sciences, 60, 2119-2135.

Ling, F., J. J. Luo, Y. Li, T. Tang, L. Bai, W. Ouyang, and T. Yamagata, 2022: Multi-task machine learning improves multi-seasonal prediction of the Indian Ocean Dipole. Nat Commun, 13, 7681.

Liu, H. F., Y. M. Tang, D. K. Chen, and T. Lian, 2017: Predictability of the Indian Ocean Dipole in the coupled models. Clim Dynam, 48, 2005-2024.

Liu, J., Y. M. Tang, Y. L. Wu, T. Li, Q. Wang, and D. K. Chen, 2021: Forecasting the Indian Ocean Dipole With Deep Learning Techniques. Geophys Res Lett, 48, e2021GL094407.

Liu, M. H., M. J. McPhaden, H. L. Ren, M. A. Balmaseda, and R. Wang, 2022: Oceanic Heat Content as a Predictor of the Indian Ocean Dipole. J Geophys Res-Oceans, 127, e2022JC018896.

Liu, Q., S. Bao, H. Yan, H. Wang, Y. Wang, and R. Zhang, 2025: Enhancing sea surface salinity short-term prediction using physically informed deep learning. Applied Ocean Research, 165, 104832.

Lu, B., and Coauthors, 2018: An extreme negative Indian Ocean Dipole event in 2016: dynamics and predictability. Clim Dynam, 51, 89-100.

Meyers, G., P. McIntosh, L. Pigot, and M. Pook, 2007: The Years of El Niño, La Niña, and Interactions with the Tropical Indian Ocean. J Climate, 20, 2872-2880.

Nakazato, M., S. Kido, and T. Tozuka, 2021: Mechanisms of asymmetry in sea surface temperature anomalies associated with the Indian Ocean Dipole revealed by closed heat budget. Sci Rep, 11, 22546.

Ng, B., W. Cai, K. Walsh, and A. Santoso, 2015: Nonlinear processes reinforce extreme Indian Ocean Dipole events. Sci Rep, 5, 11697.

Nuncio, M., and X. Yuan, 2015: The Influence of the Indian Ocean Dipole on Antarctic Sea Ice*. J Climate, 28, 2682-2690.

Qin, M. J., and Coauthors, 2022: Deep Learning for Multi-Timescales Pacific Decadal Oscillation Forecasting. Geophys Res Lett, 49, e2021GL096479.

Rao, S. A., S. K. Behera, Y. Masumoto, and T. Yamagata, 2002: Interannual subsurface variability in the tropical Indian Ocean with a special emphasis on the Indian Ocean dipole. Deep Sea Research Part II: Topical Studies in Oceanography, 49, 1549-1572.

Rao, S. A., J. J. Luo, S. K. Behera, and T. Yamagata, 2009: Generation and termination of Indian Ocean dipole events in 2003, 2006 and 2007. Clim Dynam, 33, 751-767.

Ratnam, J., S. K. Behera, M. Nonaka, and K. R. Patil, 2025: Skillful prediction of Indian Ocean Dipole index using machine learning models. Applied Computing and Geosciences, 100228.

Saha, S., and Coauthors, 2010: The Ncep Climate Forecast System Reanalysis. B Am Meteorol Soc, 91, 1015-1057.

Saji, N. H., S. P. Xie, and T. Yamagata, 2006: Tropical Indian Ocean variability in the IPCC twentieth-century climate simulations. J Climate, 19, 4397-4417.

Saji, N. H., B. N. Goswami, P. N. Vinayachandran, and T. Yamagata, 1999: A dipole mode in the tropical Indian Ocean. Nature, 401, 360-363.

Schott, F. A., S. P. Xie, and J. P. McCreary, 2009: Indian Ocean Circulation and Climate Variability. Rev Geophys, 47.

Shin, N. Y., D. Kim, D. Kang, H. Kim, and J. S. Kug, 2024: Deep learning reveals moisture as the primary predictability source of MJO. Npj Clim Atmos Sci, 7, 11.

Sprintall, J., and Coauthors, 2024: COVID impacts cause critical gaps in the Indian Ocean observing system. B Am Meteorol Soc, 105, E725-E741.

Stuecker, M. F., and Coauthors, 2017: Revisiting ENSO/Indian Ocean Dipole phase relationships. Geophys Res Lett, 44, 2481-2492.

Sun, S., Y. Fang, Y. Zu, L. Liu, and K. Li, 2022: Increased occurrences of early Indian Ocean Dipole under global warming. Sci Adv, 8, eadd6025.

Tao, Y., C. Qiu, D. Wang, M. Li, and G. Zhang, 2024: Indian Ocean Dipole (IOD) forecasts based on convolutional neural network with sea level pressure precursor. Environmental Research Letters, 19, 104045.

Terray, P., A. S. Fischer, E. Guilyardi, S. Gualdi, and P. Delecluse, 2005: Two Independent Triggers for the Indian Ocean Dipole/Zonal Mode in a Coupled GCM. J Climate, 18, 3428-3449.

Tozuka, T., J.-J. Luo, S. Masson, and T. Yamagata, 2007: Decadal Modulations of the Indian Ocean Dipole in the SINTEX-F1 Coupled GCM. J Climate, 20, 2881-2894.

Vecchi, G. A., N.-C. Lau, Y. Kosaka, L. Wu, S.-P. Xie, and Y. Yang, 2015: Seasonality and Predictability of the Indian Ocean Dipole Mode: ENSO Forcing and Internal Variability. J Climate, 28, 8021-8036.

Wang, C., F. Kucharski, R. Barimalala, and A. Bracco, 2009: Teleconnections of the tropical Atlantic to the tropical Indian and Pacific Oceans: A review of recent findings. Meteorologische Zeitschrift, 18, 445-454.

Wang, L., X. Zhang, L. R. Leung, F. H. Chiew, A. AghaKouchak, K. Ying, and Y. Zhang, 2024: CAS-Canglong: A skillful 3D Transformer model for sub-seasonal to seasonal global sea surface temperature prediction. arXiv preprint arXiv:2409.05369.

Webster, P. J., A. M. Moore, J. P. Loschnigg, and R. R. Leben, 1999: Coupled ocean–atmosphere dynamics in the Indian Ocean during 1997–98. Nature, 401, 356-360.

Xie, S. P., and S. G. H. Philander, 1994: A coupled ocean ‐ atmosphere model of relevance to the ITCZ in the eastern Pacific. Tellus A, 46, 340-350.

Yang, K., W. Cai, G. Huang, G. Wang, B. Ng, and S. Li, 2020: Oceanic processes in ocean temperature products key to a realistic presentation of positive Indian Ocean dipole nonlinearity. Geophys Res Lett, 47, e2020GL089396.

Yang, Y. M., J. H. Kim, J. H. Park, Y. G. Ham, S. I. An, J. Y. Lee, and B. Wang, 2023: Exploring dominant processes for multi-month predictability of western Pacific precipitation using deep learning. Npj Clim Atmos Sci, 6, 157.

Yao, Z., Y. Tang, D. Chen, L. Zhou, X. Li, T. Lian, and S. Ul Islam, 2016: Assessment of the simulation of Indian Ocean Dipole in the CESM—Impacts of atmospheric physics and model resolution. J Adv Model Earth Sy, 8, 1932-1952.

Zhang, R.-H., L. Zhou, C. Gao, and L. Tao, 2024: A transformer‐based coupled ocean‐atmosphere model for ENSO studies. Science Bulletin, S2095-9273 (2024) 00297.

Zhang, Y., J. Li, S. Zhao, F. Zheng, J. Feng, Y. Li, and Y. Xu, 2020: Indian Ocean tripole mode and its associated atmospheric and oceanic processes. Clim Dynam, 55, 1367-1383.

Zhang, Y., and Y. Du, 2022: Oceanic Rossby Waves Induced Two Types of Ocean–Atmosphere Response and Opposite Indian Ocean Dipole Phases. J Climate, 35, 3927-3945.

Zhou, L., and R.-H. Zhang, 2023: A self-attention–based neural network for three-dimensional multivariate modeling and its skillful ENSO predictions. Science Advances, 9, 2827.

Zhong, X., Chen, L., Liu, J., Lin, C., Qi, Y., Li, H., 2023. FuXi-Extreme: Improving extreme rainfall and wind forecasts with diffusion model. arXiv preprint arXiv:2310.19822.

Electronic Supplementary Material

Supplementary Materials for this paper include the following:

Tables S1 to S3

Table S1. List of CMIP6 historical simulations used for training MAS-Net.

No.	Model ID	Model Group	Member	Period
1	ACCESS-CM2	Commonwealth Scientific and Industrial Research Organization	r1i1p1f1	1850-1992
2	BCC-CSM2-MR	Beijing Climate Center, China Meteorological Administration		
3	BCC-ESM1			
4	CAMS-CSM1-0	Chinese Academy of Meteorological Sciences		
5	CanESM5	Canadian Centre for Climate Modelling and Analysis		
6	CAS-ESM2-0	Chinese Academy of Sciences		
7	CMCC-CM2-SR5	Centro euro-Mediterraneo sui Cambiamenti Climatici		
8	CMCC-ESM2			
9	CNRM-CM6-1	Centre National de Recherches Météorologique/Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique		
10	CNRM-ESM2-1			
11	EC-Earth3	European Centre for Medium-Range Weather Forecasts		
12	EC-Earth3-CC			
13	FGOALS-f3-L	Chinese Academy of Sciences		
14	GFDL-ESM4	NOAA Geophysical Fluid Dynamics Laboratory		
15	GISS-E2-1-H	NASA Goddard Institute for Space Studies		
16	MIROC6	Japan Agency for Marine-Earth Science and Technology/Atmosphere and Ocean Research Institute, Japan/ National Institute for Environmental Studies, Japan/RIKEN Center for Computational Science, Japan		
17	MRI-ESM2-0	Meteorological Research Institute, Japan		
18	NESM3	Nanjing University of Information Science and Technology		
19	NorCPM1	NorESM Climate modeling Consortium, Norway		

Table S2. Partition of datasets at different stages

No.	Dataset	Data	Period
1	Pretraining dataset	CMIP6	1860-1982
2	Pretesting dataset	CMIP6	1983-1992
3	Training dataset	SODA	1871-1982
4	Validation dataset	GODAS	1983-1992
5	Testing dataset	GODAS	1993-2023

Table S3. List of the North American Multi-Model Ensemble forecast system used in this study.

<table>
<tr><th>No.</th><th>Model</th><th>Ensemble size</th><th>Lead time</th><th>Period</th></tr>
<tr><td>1</td><td>CanCM3</td><td>11</td><td rowspan="6">8</td><td rowspan="6">1993-2018</td></tr>
<tr><td>2</td><td>CanCM4</td><td>11</td></tr>
<tr><td>3</td><td>CCSM3</td><td>7</td></tr>
<tr><td>4</td><td>GFDL-aer04</td><td>11</td></tr>
<tr><td>5</td><td>GFDL-FLOR-A06</td><td>13</td></tr>
<tr><td>6</td><td>GFDL-FLOR-Bo1</td><td>13</td></tr>
</table>