Check for updates

AGU ADVANCING EARTH AND SPACE SCIENCES

Geophysical Research Letters

RESEARCH LETTER
10.1029/2024GL112308

Skillful Prediction of Indian Monsoon Intraseasonal Precipitation Using Central Indian Ocean Mode and Machine Learning

Lei Zhou[1,2], Yanwei Yu[3], Bingqi Yan[3], Xingyu Zhao[3], Jianhuang Qin[4], Wei Tan[5], Youmin Tang[4,6], Xiaofeng Li[7,8], Xiaojing Li[9], Junyu Dong[3], Dake Chen[2,9], and Raghu Murtugudde[10,11]

[1]School of Oceanography, Shanghai Jiao Tong University, Shanghai, China, [2]Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, China, [3]College of Computer Science and Technology, Ocean University of China, Qingdao, China, [4]College of Oceanography, Hohai University, Nanjing, China, [5]College of Ocean Science and Engineering, Shandong University of Science and Technology, Qingdao, China, [6]Environmental Science and Engineering, University of Northern British Columbia, Prince George, BC, Canada, [7]CAS Key Laboratory of Ocean Circulation and Waves, Big Data Center, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China, [8]Center for Ocean Mega-Science, Chinese Academy of Sciences, Qingdao, China, [9]State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou, China, [10]Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD, USA, [11]Indian Institute of Technology—Bombay, Mumbai, India

Abstract Monsoonal precipitation is dominated by intraseasonal variabilities, whose skillful prediction lead time is currently less than 5 days and remains a grand challenge. Here we show that an intrinsic variability in the Indian Ocean, the Central Indian Ocean (CIO) mode, when combined with a machine learning (ML) algorithm, can produce skillful predictions of intraseasonal precipitation over the monsoon region with a lead time of over 15 days, which is close to the theoretical predictability limit. This remarkable skill improvement stems from the fact that the CIO mode is dynamically related to the intraseasonal monsoon rainfall, while the data-driven ML algorithm suppresses unwanted high-frequency noise. Using the CIO mode and the ML algorithm, the forecast system hybridizes physical fundamentals and versatility of data-driven algorithms. The identification of CIO mode and the verification of its significant contribution to intraseasonal predictions advance our understanding of the coupled monsoon system and also underscores the great potential of ML techniques in weather forecasts and climate predictions.

Plain Language Summary Rainfall during the Indian summer monsoon is dominated by variations with a period of tens of days, which are referred to as intraseasonal variabilities. Current prediction skill of intraseasonal monsoonal rainfall is less than 5 days and it remains a grand challenge in terms of increasing the current prediction skill. Here we show that an intrinsic mode of variability in the Indian Ocean, called the Central Indian Ocean (CIO) mode, when combined with a machine learning (ML) algorithm, can produce skillful predictions of intraseasonal precipitation over the monsoon region with a lead time of over 15 days. This remarkable skill improvement stems from the fact that the CIO mode is dynamically related to intraseasonal monsoon rainfall, while data-driven ML algorithm suppresses disruptive noise with a period shorter than 10 days. Using the CIO mode and an ML algorithm, the forecast system synergizes physical fundamentals and versatility of data-driven algorithm. The identification of CIO mode and the verification of its significant contribution to intraseasonal prediction advance our understanding of the coupled monsoon system and also demonstrate the great potential of ML techniques in weather forecasts and climate predictions.

Key Points:

- The Central Indian Ocean (CIO) mode provides a dynamical basis for the prediction of monsoon intraseasonal rainfall
- The machine learning (ML) algorithm suppresses high-frequency noise while capturing the real-time CIO mode index
- The dynamics and ML hybrid forecast system can skillfully predict monsoon intraseasonal rainfall with a lead time of ~15 days

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

L. Zhou,
zhoulei1588@sjtu.edu.cn

Citation:

Zhou, L., Yu, Y., Yan, B., Zhao, X., Qin, J., Tan, W., et al. (2024). Skillful prediction of Indian monsoon intraseasonal precipitation using Central Indian Ocean mode and machine learning. *Geophysical Research Letters*, *51*, e2024GL112308. https://doi.org/10.1029/2024GL112308

Received 10 SEP 2024
Accepted 8 NOV 2024

© 2024. The Author(s).
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

1. Introduction

The Indian summer monsoon is the wet season of the Indian subcontinent. Bountiful monsoon rainfall nourishes the agricultural system and sustains the food production and socioeconomic development in the monsoon region (Gadgil et al., 1984; Webster et al., 1998). Meanwhile, heavy precipitation also causes severe natural disasters every year with significant loss of life, crops, and property (Ghosh et al., 2012; Goswami et al., 2006; Vittal et al., 2016). Thus, the prediction of monsoon precipitation has been the kernel of weather and climate research over the Indian Ocean region. Although great efforts have been made and enormous progress has been achieved,

19448007, 2024, 24, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024GL112308 by Lei Zhou - Shanghai Jiaotong University , Wiley Online Library on [11/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the skill of monsoon rainfall prediction is still limited [Wang et al., 2015]. The monsoon precipitation is highly variable within each season and is dominated by intraseasonal variabilities (ISVs), commonly known as the monsoon intraseasonal oscillation (MISO) (B. N. Goswami, 2005; Shukla, 2014), which accounts for about 60% of the total variance of monsoon precipitation. The rapid evolution of MISO makes its prediction challenging and reduces the predictability of monsoon rainfall. Thus, the prediction of MISO and the associated rainfall becomes a bottleneck in the already anemic monsoon predictability.

The period of MISOs is generally between 20 and 90 days. The potential theoretical predictability of MISO is estimated to be 20–30 days for both empirical and dynamic forecasts (B. N. Goswami & Xavier, 2003; Qin et al., 2022; Waliser et al., 2003). It is usually longer for dynamical fields such as winds in the lower and upper troposphere (Zhu et al., 2015), but shorter for thermodynamic fields such as rainfall (Li et al., 2023; Zhang et al., 2021; Zhu et al., 2023). In reality, the actual lead time for MISO predictions is typically 7–10 days, which is far below its potential predictability (B. N. Goswami & Xavier, 2003). The prediction skill of intraseasonal rainfall is typically shorter than 5 days (Figure S1a in Supporting Information S1), since the correlations between the predictions from all subseasonal-to-seasonal (S2S) models (Vitart et al., 2017) and TRMM data are larger than 0.5 only in sporadic regions when the lead time is 5 days. Many factors contribute to currently limited prediction skill of MISOs and monsoon intraseasonal rainfall, including the poor simulation of MISOs in current climate models (Goswami et al., 2014; Lin et al., 2008; Meehl et al., 2012; Sperber & Annamalai, 2008), and the uncertainties in the statistical properties of MISOs (Suhas et al., 2013; Webster et al., 1998).

Nevertheless, an essential deficiency that severely curtails the monsoon prediction at intraseasonal timescales has been the lack of an identified natural mode that captures the evolution of MISOs. In contrast, the seasonal-to-interannual predictions of the monsoon benefit from their relations with ENSO (Gadgil et al., 2004; Gill et al., 2015; Kumar et al., 2006) and IODZM (Abram et al., 2008; Ashok et al., 2001; Sabeerali et al., 2014). However, due to the rather weak relation between MISOs and the climate modes over the Indo-Pacific region (Wang et al., 2008, 2015; Webster et al., 1998), MISO predictions are less skillful than seasonal-to-interannual monsoon predictions.

A central Indian Ocean (CIO) mode was established as the first Empirical Orthogonal Function (EOF) mode of the combined intraseasonal sea surface temperature (SST) anomalies and intraseasonal zonal wind anomalies at 850 hPa. Dynamically, barotropic energy conversion is the key to maintain the CIO mode. It was found to have a close dynamic relation with MISOs and intraseasonal rainfall during summer monsoon (Qin et al., 2020; Zhou et al., 2017a, 2017b, 2018). The CIO mode acts like a T-junction which controls the propagation direction of ISVs originating in the western Indian Ocean, that is, the positive CIO mode enhances the easterly wind shear in the tropics and creates an environment favorable for northward-propagating MISOs. More details about the dynamics of the CIO mode are presented in Section 3.1 and in Zhou et al. (2017a). In this study, a forecast system taking advantages of both the CIO mode and machine learning (ML) algorithm is established, which yields skillful predictions of intraseasonal rainfall with a lead time of ∼15 days. The remainder of this article is organized as follows. The data and methods are described in Section 2. Section 3 presents the results of the hybrid forecast system, comparing it with the abilities of S2S models. Finally, Section 4 provides the conclusions and a discussion.

2. Data and Methods

Zonal winds are obtained from daily reanalysis products from the European Center for Medium-Range Weather Forecasts (ECMWF) Interim Reanalysis (ERA-Interim) (Dee et al., 2011) and SST data are from the NOAA 1/4° daily Optimum Interpolation SST (Reynolds et al., 2007). Different reanalysis products, such as NCEP-DOE Reanalysis 2 (Kanamitsu et al., 2002) and ERA5 (Hersbach et al., 2023), are tested for the robustness of the following results. In addition, daily rainfall data are obtained from the Tropical Rainfall Measuring Mission (TRMM 3B42) (Huffman et al., 2007) from 1998 to 2019. All intraseasonal variabilities are obtained with a 20–100-day bandpass Butterworth filter. Forecasts from 11 coupled S2S models (Qin et al., 2020; Vitart et al., 2017) are used for the evaluation of prediction skill of our forecast system.

2.1. Retrieving the CIO Mode

The EOF analysis is conducted with a combined field of daily intraseasonal SSTs and daily intraseasonal zonal winds at 850 hPa. A Butterworth bandpass filter between 20 and 100 days is applied to obtain the intraseasonal

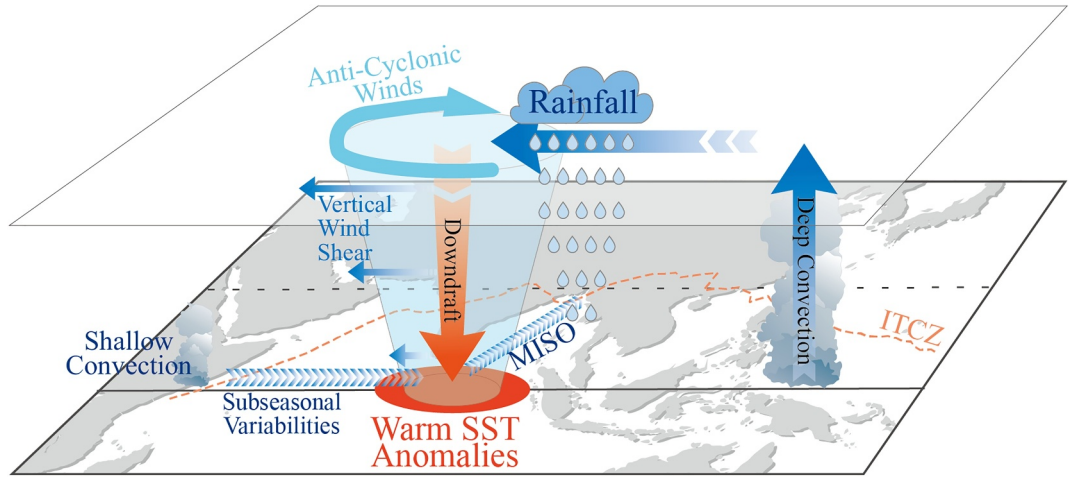

Figure 1. Schematic of the positive Central Indian Ocean mode. See the main text and Zhou et al. (2017a) for descriptions of the detailed processes.

component. Meanwhile, the high-frequency components (with a period shorter than 20 days) are removed. The domain extends from 40°E to 120°E and from 20°N to 20°S. Each field is normalized by subtracting its regional mean and dividing by its regional variance within the above domain before the EOF analysis. SSTs and winds can then be compared and combined. The CIO mode is the first combined EOF mode and its principal component is the CIO mode index. See Zhou et al. (2017a) for further technical details. The sketch of the positive CIO mode is shown in Figure 1.

2.2. Obtaining the CIO Mode Index Using a Projection Method

The CIO mode index defined in Section 2.1 requires bandpass filtering. As a result, the intraseasonal components are maintained, but its application to practical forecast is inhibited. Thus, a projected CIO mode index is obtained as below.

For each time t, the daily climatological mean SST and mean zonal winds at 850 hPa are subtracted from total SSTs and total zonal winds (without filtering; Input step in Figure 2) so that the low-frequency variabilities are largely removed. Then, the SST anomalies and the zonal wind anomalies are normalized by the specified mean and standard deviation obtained when computing the CIO mode. Next, the normalized SSTs and zonal winds at 850 hPa are combined in the same way to generate the CIO mode (Line 2 in Figure 2). Then the combined fields are projected onto the canonical CIO mode (Line 3 in Figure 2), and the projected CIO mode index is obtained at time t (CIO_t). Note that no filtering is used during the entire process of projection. Thus, the projected CIO mode index at any time can be obtained with the observed SSTs and zonal winds simultaneously. No data before or after that day are needed.

2.3. Machine Learning Algorithm of Long Short-Term Memory (LSTM)

A neural network is applied based on Long Short-Term Memory (LSTM) network (Hochreiter & Schmidhuber, 1997), which captures the long-term and short-term dependences on temporal and sequential data as designed. The LSTM is a type of Recurrent Neural Network. It has been widely used to simulate sequential forecast problems and has achieved great success in machine translation (Melamud et al., 2016; Wang et al., 2016), speech recognition (Graves et al., 2013), and spatial-temporal data mining (Alahi et al., 2016; Dai et al., 2021; Kong & Wu, 2018). The input of the LSTM cell is CIO_t, that is, the projected CIO mode index at time t (Figure 2). The hidden state h_t and the memory cell c_t are the outputs of the LSTM cell at time step t. The LSTM cell consists of three gates, that is, the forget gate, input gate, and output gate. The data feeding into an LSTM cell are the input at the current time step CIO_t and the hidden state of the previous time step h_{t-1}. The first gate is the *forget gate*, which decides what information is discarded from the cell. The formula for the forget gate is $f_t = \sigma\left(W_f \cdot [h_{t-1}, \mathrm{CIO}_t] + b_f\right)$, where W_f is the learnable weight parameter matrix; b_f is the bias parameter; [,] denotes concatenate operation; and σ is a sigmoid function, describing how much of each component should be let

19448007, 2024, 24, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024GL112308 by Lei Zhou - Shanghai Jiaotong University , Wiley Online Library on [11/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

10.1029/2024GL112308

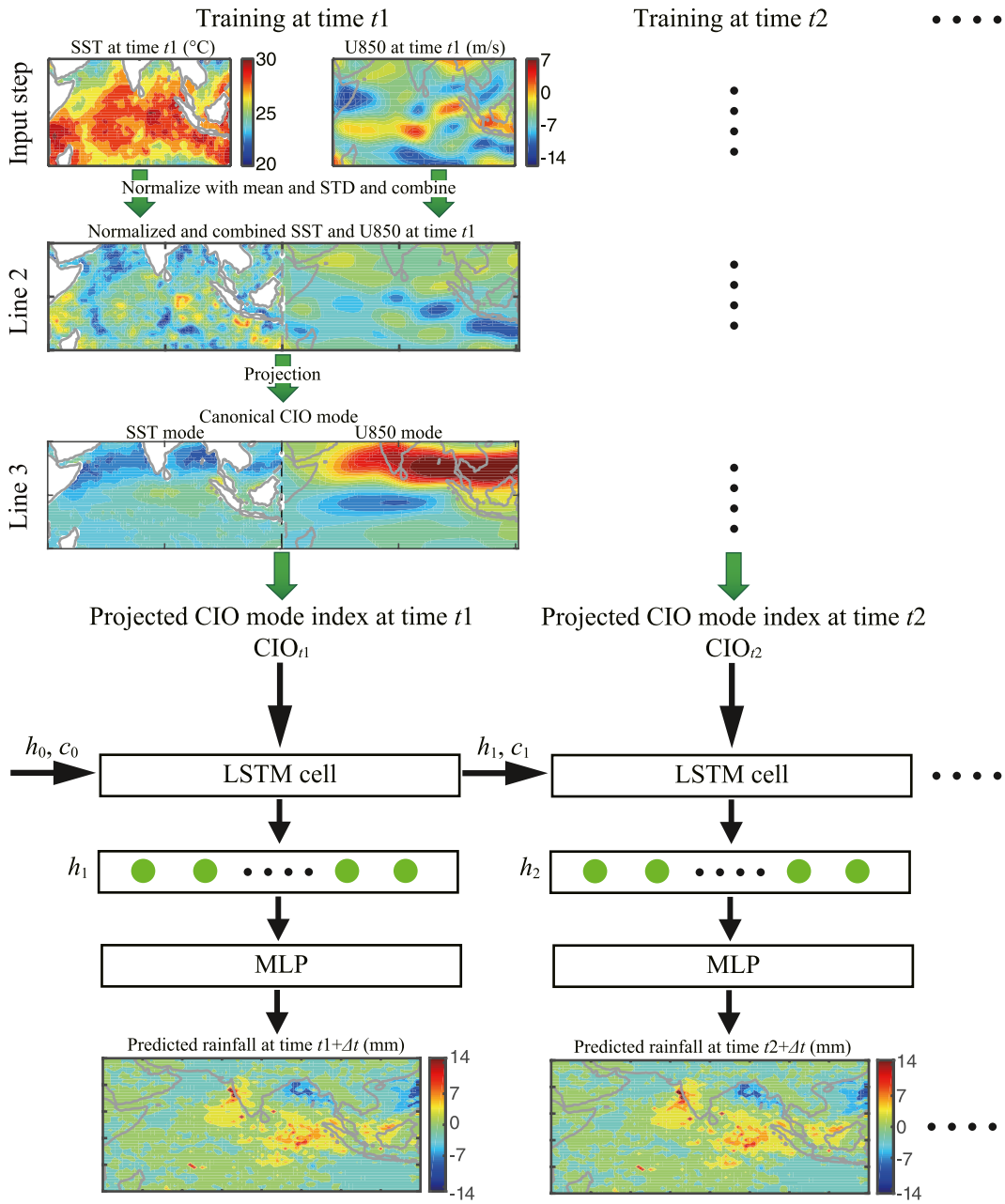

Figure 2. Structure of the forecast system. At time t, the input to the system are total sea surface temperature and total zonal winds at 850 hPa (U850; Input step). They are normalized by subtracting their mean and dividing by their standard deviation (STD; Line 2). The normalized variables are merged and projected onto the canonical Central Indian Ocean (CIO) mode (Line 3). The projected CIO mode index at time t (CIO_t) feeds into the Long Short-Term Memory cell (see Methods for details). After passing the multi-layer perceptron, the spatial distribution of rainfall at $t + \Delta t$ is predicted with a lead time of Δt. The process is repeated for each time step, and the precipitation is forecasted for subsequent time steps.

through. The second gate is the *input gate*, which decides when to read data into the cell. The third gate is needed to export the entries from the cell, denoted as the *output gate*. The formula for the input gate and the output gate are $i_t = \sigma(W_i \cdot [h_{t-1}, \mathrm{CIO}_t] + b_i)$ and $o_t = \sigma(W_o[h_{t-1}, CIO_t] + \mathrm{b}_o)$, separately, where W_i and W_o are the learnable weight parameter matrices, and b_i and b_o are the learnable bias parameters. Additionally, LSTM introduces a candidate memory cell $\tilde{c}_t$ as $\tanh(W_c \cdot [h_{t-1}, \mathrm{CIO}_t] + b_c)$, where W_c is the weight parameter matrix, b_c is the bias parameter, and they are all learnable parameters. Therefore, LSTM has two dedicated gates for controlling the maintenance and the loss of memories, that is, the input gate i_t governs how much new information

19448007, 2024, 24, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024GL112308 by Lei Zhou - Shanghai Jiaotong University , Wiley Online Library on [11/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19448007, 2024, 24, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024GL112308 by Lei Zhou - Shanghai Jiaotong University , Wiley Online Library on [11/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

is retained via $\tilde{c}_t$ and the forget gate f_t controls how much previous information is memorized via c_{t-1}, that is, $c_t = f_t * c_{t-1} + i_t * \tilde{c}_t$. Lastly, we need to decide what information to output as hidden state $h_t = o_t * \tanh(c_t)$. Therefore, whether all memory information is passed through to the predictor depends on the value of the output gate.

2.4. Flowchart of the Hybrid Forecast System

Our forecast system for intraseasonal precipitation is a hybrid approach using the projected CIO mode index and the LSTM network. For the training process, at any given time t, the projected daily CIO mode index (CIO_t) is obtained from total SSTs and zonal winds at 850 hPa, which is then input into the LSTM cell. Sequentially, a multi-layer perceptron (MLP) (Kong & Wu, 2018) is applied, which maps the outputs from the LSTM cell to a low-dimensional space to predict the intraseasonal rainfall at time $t + \Delta t$, that is, $\hat{p}_{t+\Delta t} = \mathrm{MLP}[\mathrm{LSTM}(CIO_t)]$. Empirically, we find that using the neural network of three layers for MLP yields the best compromise between the performance and the training time. In addition, based on the data volume and performance, the number of hidden layer neurons is set to 60, 30, and 15, respectively. As the number of layers increases, the number of neurons gradually decreases, and features are gradually synthesized and compressed, which helps the model to gradually extract useful information. This also simultaneously sets the rectified linear unit as the activation function to introduce nonlinear characteristics and alleviates the problem of gradient vanishing. Then, an objective function is applied to evaluate the forecast against observations, $\mathcal{L} = \sum_{\tau} \mathrm{MSE}(\hat{p}_\tau, p_\tau)$, where p_τ denotes observations, and $\mathrm{MSE}(\bullet, \bullet)$ is a criterion that measures the mean squared error. The training process produces the best estimates of the learnable variables (e.g., W_f, W_i, W_o, W_c, b_f, b_i, b_o, b_c), which are used for prediction. Each independent spatial point has been trained into an independent model. The output of MLP in each model represents a one-dimensional array of rainfall sequences at specific spatial points, and ultimately the one-dimensional output of each model together constitutes the spatial distribution of rainfall. The entire procedure is summarized in Figure 2. Since TRMM data are available for 22 years (from 1998 to 2019), the data length for training is limited. Thus, a cross-validation is adopted for testing. Take year 2003 for example, data for 21 years (1998–2002 and 2004–2019) are used for training and data for 2003 are used for testing. The results in this study are not sensitive to the training data. For example, when making forecasts of monsoon intraseasonal rainfall for 2014, we compared two kinds of training processes. In one method, the training data were obtained from 1998 to 2013 and 2015–2019, which was the same as above. In the other method, the training data were only obtained from 1998 to 2013. The forecasts using the parameters obtained from the two training methods were comparable (not shown).

3. Results

3.1. Dynamics of the Central Indian Ocean Mode

The CIO mode represents the mechanistic links among various dynamic and thermodynamic fields during the Indian summer monsoon. As sketched in Figure 1, for the positive CIO mode, warm intraseasonal SST anomalies occur over the central tropical Indian Ocean. An anticyclonic gyre in the lower troposphere occurs with an easterly branch around 5°S and a westerly branch around 10°N. The phase relation between the SST anomalies and the low-level wind anomalies at 850 hPa is a key feature of the CIO mode, which indicates the force and response relation between the ocean and the atmosphere (Peña et al., 2003; Xi et al., 2015). The CIO mode is fueled by barotropic energy transfer due to meridional shear of large-scale zonal winds. Vertically, there are easterly winds in the lower and the upper troposphere, but the latter is stronger than the former. As a result, easterly vertical wind shear occurs for the positive CIO mode, which creates a favorable environment for the northward propagating MISOs (Jiang et al., 2004; Kang et al., 2010). Therefore, the CIO mode acts like a T-junction and directs the eastward propagating ISVs from the western tropical Indian Ocean northward into the monsoon region, bringing ample momentum and moisture from the tropics into the subtropics. As a natural mode at intraseasonal timescales, the CIO mode is dynamically distinct from the large-scale climate system in the tropics such as ENSO and IODZM (Zhou et al., 2017b). Thus, the CIO mode can provide independent information for predictive understanding of the monsoon, although the interactions with other climate modes cannot be fully excluded.

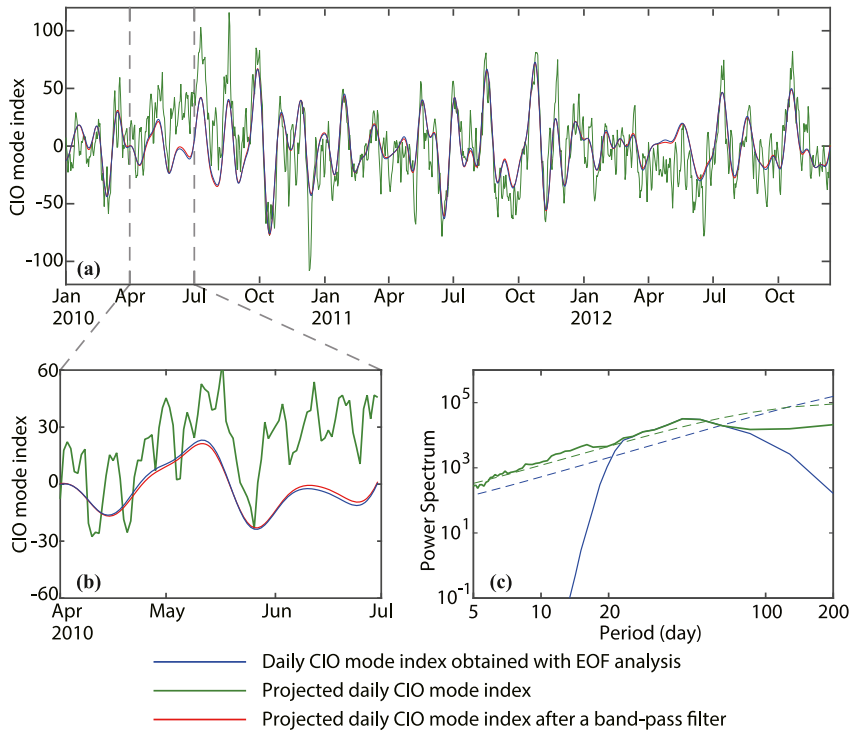

Figure 3. Projected Central Indian Ocean (CIO) mode index captures the intraseasonal properties of the actual CIO mode index. (a) Blue line: the actual CIO mode index obtained with the combined Empirical Orthogonal Function (EOF) analysis, that is, the principal component of the first combined EOF mode. Green line: The daily index obtained by projecting the total sea surface temperature and total zonal winds at 850 hPa onto the canonical CIO mode, without applying any filtering (Section 2.2). Red line: The projected index (green line) after a bandpass filtering between 20 and 100-day. Three years from 2010 to 2012 are randomly selected. The red line matches the blue line so well that their differences are hardly discernible. (b) Enlarged figure of (a) from April to June in 2010. (c) The power spectrum of the actual CIO mode index (solid blue line) and the power spectrum of the projected CIO mode index (solid green line). The blue and green dash lines mark the 95% confidence level of corresponding spectra.

3.2. Forecast System for Intraseasonal Rainfall Based on the Physical Mode and Machine Learning

The forecast system for intraseasonal rainfall over the monsoon region takes advantage of the dynamics of the CIO mode and the ML ability of an LSTM network (Hochreiter & Schmidhuber, 1997).

The CIO mode preserves the physical consistency between predictors and predictands. The CIO mode (Figure 1) and its index (Figure 3a) are not sensitive to the periods of data used for the EOF mode retrieval (Figure 4) (Qin et al., 2020). At intraseasonal timescales, the CIO mode index has a close correlation with the monsoon precipitation (Qin et al., 2020; Zhou et al., 2017a) (Figure S2a in Supporting Information S1), especially over the Bay of Bengal, where a large variance of intraseasonal precipitation occurs (Figure S2b in Supporting

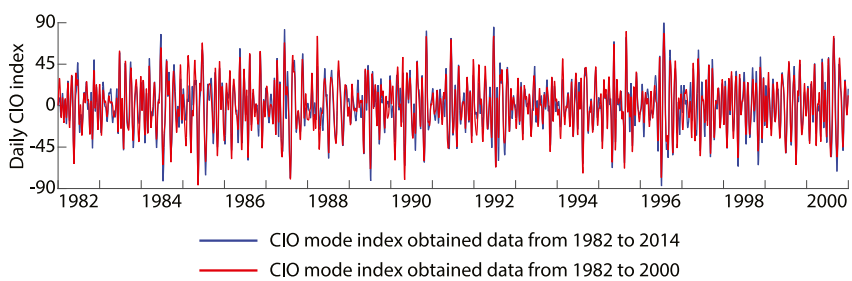

Figure 4. Daily Central Indian Ocean (CIO) mode index obtained with data for different time periods. The daily CIO mode index was obtained with data from 1982 to 2014 (blue line) and from 1982 to 2000 (red line), using the combined Empirical Orthogonal Function analysis.

19448007, 2024, 24, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024GL112308 by Lei Zhou - Shanghai Jiaotong University , Wiley Online Library on [11/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

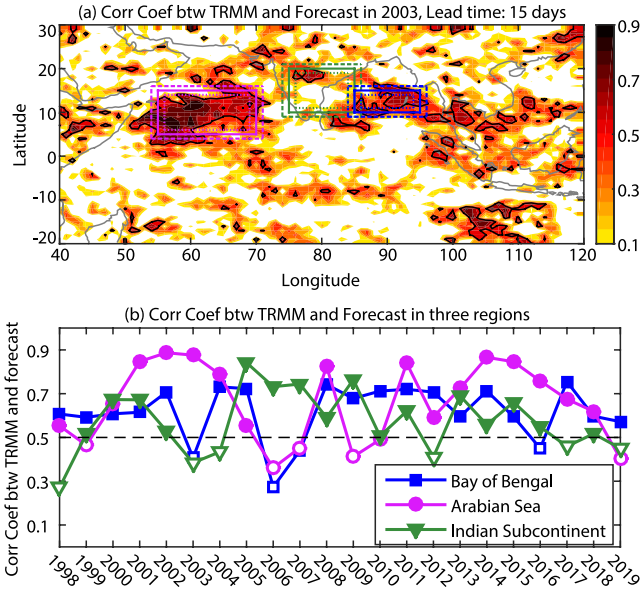

Figure 5. Prediction skill of the forecast system for the intraseasonal Indian monsoon rainfall. (a) The correlation coefficients of the predicted intraseasonal rainfall and observations (TRMM data with a spatial resolution of 1° longitude × 1° latitude) during the Indian summer monsoon (June–September) in 2003. The lead time is 15 days. The contours of 0.5 are marked in black. Three regions are boxed with solid lines, that is, the Bay of Bengal (10°N–15°N, 85°E–95°E), the Arabian Sea (5°N–15°N, 55°E–70°E), and the Indian subcontinent (10°N–20°N, 75°E–85°E). For sensitivity tests, all three regions are enlarged by 1° latitude and longitude (dashed boxes), and shrunk by 1° latitude and longitude (dotted boxes). (b) The correlation coefficients between predicted and observed intraseasonal monsoon rainfall over the three regions (solid boxes in a) for 22 years. The coefficients larger than 0.5 are represented with filled markers, while the ones smaller than 0.5 are represented with empty markers. The lead time is 15 days for the Bay of Bengal and the Arabian Sea, while it is 7 days for the Indian subcontinent. The coefficient of 0.5 is marked with a black dashed line.

Information S1). Large negative correlations over the tropical Indian Ocean also indicate strong relations between the CIO mode index and intraseasonal rainfall, which can facilitate the prediction via the ML algorithm. Intraseasonal precipitation, the most useful and yet the most challenging component of monsoon predictions, is the target of the empirical forecast system in this study. This is especially the case since both the high-frequency variabilities [such as low pressure systems (LPS)] and the long-term global warming are reported to project onto the duration, intensity, and frequency of the intraseasonal timescale active and break periods (Goswami et al., 2003; Saha et al., 2019; Singh et al., 2014, 2019). Hence, the predictand here is intraseasonal rainfall. In practice, the actual CIO mode index is not available for forecast since a time series, including future data, is needed to apply the bandpass filter. Therefore, a projection method is applied instead (Figure 3; Section 2.2). The projected CIO mode index (green lines in Figures 3a and 3b) closely resembles the actual index (blue lines in Figures 3a and 3b), and the intraseasonal components of the CIO mode index are fully captured by the projected index (red lines in Figures 3a and 3b). However, the projected CIO mode index includes high-frequency noises (Figure 3c), which greatly reduce the performance of the forecast system.

The ML algorithm, that is, the LSTM network (Section 2.3) (Hochreiter & Schmidhuber, 1997), is applied to suppress the high-frequency noise in the projected CIO mode index and to isolate intraseasonal variabilities that have predictability for monsoon intraseasonal rainfall. For the training procedure, at any time t, the projected CIO mode index (CIO_t) is fed into the LSTM cell. After passing an MLP, precipitation at time $t + \Delta t$ ($p_{t+\Delta t}$) is predicted. The procedure is summarized in Figure 2 and is described in Section 2.4. The process is repeated for each time step in the training set. The series of predicted precipitation is calibrated against the TRMM observations. The cross-validation is applied to increase the training period (Section 2.4). As a result, the best parameters for the forecast system are determined and applied in the testing procedure.

Taking year 2003 for demonstration, when the lead time is 15 days, daily rainfall forecast is made from Jun 1 to Sep 30 (122 days in all). The model achieved a significant correlation (>0.5) over the 122 days between forecasted and observed rainfall over the Bay of Bengal, Arabian Sea, and discernible regions over the Indian subcontinent (Figure 5a). Forecast skills in four more randomly selected years are presented in Figure S3 in Supporting Information S1. The results are roughly consistent with the spreading of high simultaneous correlations (both positive and negative) between the CIO mode index and monsoonal intraseasonal rainfall (Figure S2a in Supporting Information S1). Forecasts for 2003 with lead times of 10 and 5 days are shown in Figure S4 in Supporting Information S1. The corresponding results for other years are qualitatively the same, and hence not shown. The correlations between forecasts and observations are significantly higher than those in the S2S models (Figure S1 in Supporting Information S1), in which the correlations between the forecasts from S2S models and TRMM data are not larger than 0.5 in almost any regions when the lead time is 10 days or longer. Since the ML algorithm can capture more complex relations between the CIO mode index and MISO than the linear correlations, the areas with high prediction skills are wider than those with significant correlations, as shown in Figure S2a in Supporting Information S1.

The cross-validation is conducted for the monsoon seasons in 22 years from 1998 to 2019. The correlation coefficients between forecasts and observations are larger than 0.5 in most years (filled markers in Figure 5b) over three critical regions, that is, the Bay of Bengal (10°N–15°N, 85°E–95°E), the Arabian Sea (5°N–15°N, 55°E–70°E), and the Indian subcontinent (10°N–20°N, 75°E–85°E), when the lead times are 15 days for the first two regions and 7 days for the Indian subcontinent. The correlations are smaller than 0.5 in only 4 years, 6 years, and 6 years for the three regions (empty markers in Figure 5b), respectively. The evolutions of forecast performance along with the increase of lead time over the three regions are shown in Figure 6 (also taking the year of 2003 for example). The corresponding correlation coefficients from 11 S2S models are superimposed with thin lines, and

19448007, 2024, 24, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024GL112308 by Lei Zhou - Shanghai Jiaotong University , Wiley Online Library on [11/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

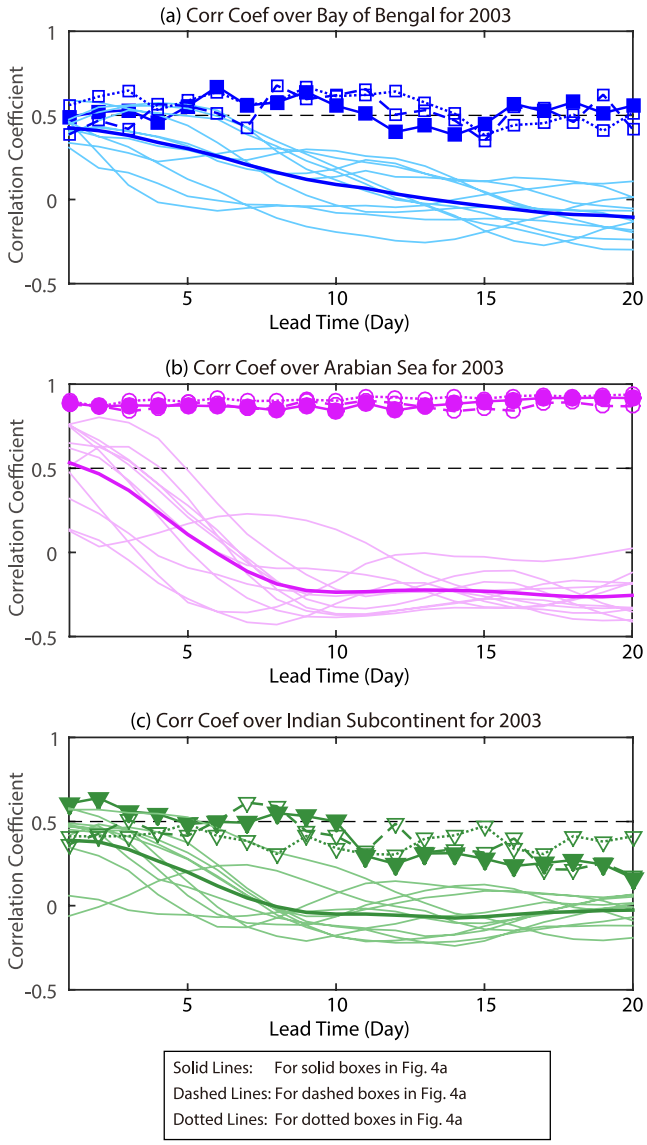

Figure 6. Prediction of monsoon rainfall along with lead times in three regions. Solid lines with markers show the correlation coefficients between predictions and TRMM data during the Indian summer monsoon (Jun–Sep) for 2003 over the Bay of Bengal (a), the Arabian Sea (b), and the Indian subcontinent (c). The three regions are marked in Figure 5a with solid boxes. The dashed (dotted) lines in each panel shows the results for the dashed (dotted) boxes in Figure 5a. The corresponding correlation coefficients from 11 of the S2S models are superimposed with thin lines. Each line is for an S2S model. The multi-model average is shown with a thick line with no marker. The coefficients of 0.5 are marked with black dashed lines.

the multi-model mean is shown with a thick line with no markers. Over the Bay of Bengal (Figure 6a), the prediction ability of our forecast system is retained up to a lead time of 20 days, and it outperforms the S2S models after the lead time of 5 days. For the Arabian Sea (Figure 6b), the correlation coefficients of our system are persistently higher than all S2S models and our prediction ability can be well beyond the lead time of 20 days. For the Indian subcontinent (Figure 6c), which is particularly important for the intraseasonal forecasts over the heavily populated regions, none of the S2S models can predict the intraseasonal monsoon rainfall beyond 4 days. Our system can maintain the prediction skill out to a lead time of 10 days. The limited skill over central Indian subcontinent is understandable, since monsoonal rainfall in that region is also subject to high-frequency transient impacts, such as LPS with a typical period of 3–5 days (Das et al., 2024; Goswami et al., 2003, 2014; Saha et al., 2019), which are not explicitly included in the intraseasonal CIO mode index.

Obviously, spatial heterogeneity exists in forecast skills, since the algorithm (Figure 2) is executed at each individual grid point. Thus, the sensitivity tests are performed by perturbing the aforementioned three regions. The three regions (solid boxes in Figure 5a) are enlarged by 1° longitude and latitude (dashed boxes in Figure 5a) and shrunk by 1° longitude and latitude (dotted boxes in Figure 5a), respectively. The correlation coefficients between the predicted and observed intraseasonal monsoon rainfall for the 22 years are presented in Figure S5 in Supporting Information S1, which display comparable skills as in Figure 5b, although there are quantitative differences due to the spatial heterogeneity (Figure 5a). Correspondingly, the evolutions of forecast performance along with the increase of lead time over the enlarged regions (dashed boxes in Figure 5a) and the shrunk regions (dotted regions in Figure 5a) are superimposed in Figure 6 with dashed lines and dotted lines, respectively. The solid, dashed, and dotted lines in Figure 6 are not clearly distinguishable, which indicates that the forecast skills are not sensitive to the domain sizes. In addition, the prediction abilities of our forecast system in all regions are significantly higher than those in the S2S models (thin lines in Figure 6).

4. Conclusions and Discussion

The CIO mode is an intrinsic ocean-atmosphere coupled mode, which captures the physical and dynamical basis for intraseasonal monsoon rainfall prediction (Zhou et al., 2017a). An empirical forecast system is constructed by employing the CIO mode physics with an ML algorithm. The combination of advances in physical understanding and big data techniques is being reported more and more to lead to practical improvements in various climate predictions (Gibson et al., 2021; Ham et al., 2019). The lead time is extended to ∼15 days over the ocean and to ∼10 days over the Indian subcontinent. Combined with the state-of-the-art prediction of MISO and cutting-edge big data techniques, such as the deep learning algorithms (Li et al., 2022; Zheng et al., 2020; Zhu et al., 2022), a much better forecast of the monsoon intraseasonal rainfall can be envisioned. Especially, predictions of heavy rainfall events, which are reported to be increasing and are critical (Goswami et al., 2006; Roxy et al., 2017) for managing devastating floods over the monsoon region, can be significantly improved. Needless to say that this prediction skill must be translated to an operational forecast system with the current climate prediction models and the available precipitation data from the Global Precipitation Mission and the various reanalysis products (Sahai et al., 2014; Wang et al., 2015).

The prediction skills of our system have spatial and interannual variation. In the complex monsoon system, the intraseasonal rainfall is subject to many factors. The CIO mode captures a major feature among all natural factors.

19448007, 2024, 24, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024GL112308 by Lei Zhou - Shanghai Jiaotong University , Wiley Online Library on [11/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19448007, 2024, 24, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024GL112308 by Lei Zhou - Shanghai Jiaotong University , Wiley Online Library on [11/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The dynamics of CIO mode and its relation to intraseasonal monsoonal rainfall have been diagnosed in previous studies (e.g., Meng et al., 2022; Qin et al., 2020, 2022; Zhou et al., 2017a, 2017b, 2018). We believe that the strong physical intimate link between the CIO mode and the intraseasonal monsoonal rainfall is the basis for the enhanced prediction skills of our system. However, this relation between CIO mode and intraseasonal rainfall varies in space and time. It is understandable that the skill is higher when and where the bond is stronger. As can be expected, the skill becomes lower when and where the association becomes weaker.

In addition, the forecast over the Indian subcontinent seems immanently harder than those over the Bay of Bengal and the Arabian Sea. The dynamical reasons require further dedicated explorations. Nevertheless, generally, LPS has an important impact on monsoon rainfall, which is not included in current system, since the ISVs are the focus of this study. The land processes and their interactions with the atmosphere have non-negligible impacts on rainfall, such as mineral dust (Das et al., 2015; Jin et al., 2021), vegetation (Budakoti et al., 2021; Lee et al., 2009), soil conditions (Paul et al., 2016; Robock et al., 2003; Saha et al., 2012), etc. In contrast, the CIO mode is only an ocean-atmosphere coupled mode, which does not include important land properties. We suggest that this could be the physical reason that our system has a relative low forecast skill over the Indian subcontinent. Therefore, a challenge for future studies will be to explore how to properly incorporate the land properties in the predictor or the dynamical forecast models (Bach et al., 2024).

Although the forecast skills in Figure 5a and Figure S3 in Supporting Information S1 are largely lower over the Indian subcontinent than they are over the ocean, the skill of our system over the Indian subcontinent is still significantly higher than the S2S models. As shown in Figure 6c, the predicted rainfall of our system and TRMM data have a correlation coefficient larger than 0.5 when the lead time is 10 days (green lines with markers). In contrast, most S2S modes drop below 0.5 when the lead time is 2–3 days (thin green lines); the composite of S2S models do not exceed the correlation coefficient of 0.5 even when the lead time is 1 day (thick green line with no markers). Therefore, our system presents meaningful improvement in monsoonal intraseasonal rainfall prediction, even over the Indian subcontinent.

Data Availability Statement

All data used in this study are available publicly. The ERA-interim reanalysis products are obtained from Dee et al. (2011). NCEP-DOE Reanalysis 2 products are from Kanamitsu et al. (2002). ERA5 products are from Hersbach et al. (2023) (registration may required by the website). Optimum Interpolation SST data are from Reynolds et al. (2007). TRMM rainfall data (3B42) are from Tropical Rainfall Measuring Mission (2011). The S2S products are available at https://apps.ecmwf.int/datasets/data/s2s/.

Acknowledgments
LZ is supported by grants from the National Natural Science Foundation of China (42125601), the National Key Research and Development Program of China (2023YFF0805300), Innovation Group Project of Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai) (311020004). YY is supported by National Natural Science Foundation of China (62176243). JQ is supported by the National Natural Science Foundation of China (42106003). JD is supported by National Natural Science Foundation of China (41927805) and the National Key Research and Development Program of China (2018AAA0100602). RM gratefully acknowledges the faculty position at the Indian Institute of Technology, Bombay.

References

Abram, N. J., Gagan, M. K., Cole, J. E., Hantoro, W. S., & Mudelsee, M. (2008). Recent intensification of tropical climate variability in the Indian Ocean. *Nature Geoscience*, *1*(12), 849–853. https://doi.org/10.1038/ngeo357

Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L., & Savarese, S. (2016). Social LSTM: Human trajectory prediction in crowded spaces. In *Paper presented at 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 27-30 June 2016*.

Ashok, K., Guan, Z., & Yamagata, T. (2001). Impact of the Indian Ocean dipole on the relationship between the Indian monsoon rainfall and ENSO. *Geophysical Research Letters*, *28*(23), 4499–4502. https://doi.org/10.1029/2001gl013294

Bach, E., Krishnamurthy, V., Mote, S., Shukla, J., Sharma, A. S., Kalnay, E., & Ghil, M. (2024). Improved subseasonal prediction of South Asian monsoon rainfall using data-driven forecasts of oscillatory modes. *Proceedings of the National Academy of Sciences*, *121*(15), e2312573121. https://doi.org/10.1073/pnas.2312573121

Budakoti, S., Chauhan, T., Murtugudde, R., Karmakar, S., & Ghosh, S. (2021). Feedback from vegetation to interannual variations of Indian summer monsoon rainfall. *Water Resources Research*, *57*(5), e2020WR028750. https://doi.org/10.1029/2020WR028750

Dai, S., Yu, Y., Fan, H., & Dong, J. (2021). Personalized POI recommendation: Spatio-temporal representation learning with social tie. In *Paper presented at database systems for advanced applications*. Springer International Publishing.

Das, S., Dey, S., Dash, S. K., Giuliani, G., & Solmon, F. (2015). Dust aerosol feedback on the Indian summer monsoon: Sensitivity to absorption property. *Journal of Geophysical Research: Atmospheres*, *120*(18), 9642–9652. https://doi.org/10.1002/2015JD023589

Das, S., Goswami, D. J., Mahanta, R., Saha, P., & Goswami, B. N. (2024). Dynamics of May 'onset' of Indian summer monsoon over Northeast India. *Quarterly Journal of the Royal Meteorological Society*, *150*(764), 4533–4549. https://doi.org/10.1002/qj.4828

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/qj.828

Gadgil, S., Joseph, P. V., & Joshi, N. V. (1984). Ocean-atmosphere coupling over monsoon regions. *Nature*, *312*(5990), 141–143. https://doi.org/10.1038/312141a0

Gadgil, S., Vinayachandran, P. N., Francis, P. A., & Gadgil, S. (2004). Extremes of the Indian summer monsoon rainfall, ENSO and equatorial Indian Ocean oscillation. *Geophysical Research Letters*, *31*(12). https://doi.org/10.1029/2004gl019733

Ghosh, S., Das, D., Kao, S.-C., & Ganguly, A. R. (2012). Lack of uniform trends but increasing spatial variability in observed Indian rainfall extremes. *Nature Climate Change*, *2*(2), 86–91. https://doi.org/10.1038/nclimate1327

Gibson, P. B., Chapman, W. E., Altinok, A., Delle Monache, L., DeFlorio, M. J., & Waliser, D. E. (2021). Training machine learning models on climate model output yields skillful interpretable seasonal precipitation forecasts. *Communications Earth & Environment*, *2*(1), 159. https://doi.org/10.1038/s43247-021-00225-4

Gill, E. C., Rajagopalan, B., & Molnar, P. (2015). Subseasonal variations in spatial signatures of ENSO on the Indian summer monsoon from 1901 to 2009. *Journal of Geophysical Research: Atmospheres*, *120*(16), 8165–8185. https://doi.org/10.1002/2015jd023184

Goswami, B. B., Deshpande, M., Mukhopadhyay, P., Saha, S. K., Rao, S. A., Murthugudde, R., & Goswami, B. N. (2014). Simulation of monsoon intraseasonal variability in NCEP CFSv2 and its role on systematic bias. *Climate Dynamics*, *43*(9), 2725–2745. https://doi.org/10.1007/s00382-014-2089-5

Goswami, B. N. (2005). South Asian monsoon. In *Intraseasonal variability in the atmosphere-ocean climate system* (pp. 19–61). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-27250-x_2

Goswami, B. N., Ajayamohan, R. S., Xavier, P. K., & Sengupta, D. (2003). Clustering of synoptic activity by Indian summer monsoon intraseasonal oscillations. *Geophysical Research Letters*, *30*(8). https://doi.org/10.1029/2002GL016734

Goswami, B. N., Venugopal, V., Sengupta, D., Madhusoodanan, M. S., & Xavier, P. K. (2006). Increasing trend of extreme rain events over India in a warming environment. *Science*, *314*(5804), 1442–1445. https://doi.org/10.1126/science.1132027

Goswami, B. N., & Xavier, P. K. (2003). Potential predictability and extended range prediction of Indian summer monsoon breaks. *Geophysical Research Letters*, *30*(18). https://doi.org/10.1029/2003gl017810

Graves, A., Jaitly, N., & Mohamed, A. (2013). Hybrid speech recognition with deep bidirectional LSTM. In *Paper presented at 2013 IEEE workshop on automatic speech recognition and understanding, 8-12 Dec. 2013*.

Ham, Y.-G., Kim, J.-H., & Luo, J.-J. (2019). Deep learning for multi-year ENSO forecasts. *Nature*, *573*(7775), 568–572. https://doi.org/10.1038/s41586-019-1559-7

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., et al. (2023). ERA5 hourly data on pressure levels from 1940 to present [Dataset]. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. https://doi.org/10.24381/cds.bd0915c6

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Gu, G., Hong, Y., et al. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, *8*(1), 38–55. https://doi.org/10.1175/jhm560.1

Jiang, X., Li, T., & Wang, B. (2004). Structures and mechanisms of the northward propagating boreal summer intraseasonal oscillation. *Journal of Climate*, *17*(5), 1022–1039. https://doi.org/10.1175/1520-0442(2004)017<1022:Samotn>2.0.Co;2

Jin, Q., Wei, J., Lau, W. K. M., Pu, B., & Wang, C. (2021). Interactions of Asian mineral dust with Indian summer monsoon: Recent advances and challenges. *Earth-Science Reviews*, *215*, 103562. https://doi.org/10.1016/j.earscirev.2021.103562

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S. K., Hnilo, J. J., Fiorino, M., & Potter, G. L. (2002). NCEP-DOE AMIP-II reanalysis (R-2). *Bulletin America Meteorology Social*, *83*(11), 1631–1643. https://doi.org/10.1175/bams-83-11-1631(2002)083<1631:nar>2.3.co;2

Kang, I.-S., Kim, D., & Kug, J.-S. (2010). Mechanism for northward propagation of boreal summer intraseasonal oscillation: Convective momentum transport. *Geophysical Research Letters*, *37*(24). https://doi.org/10.1029/2010GL045072

Kong, D., & Wu, F. (2018). HST-LSTM: A hierarchical spatial-temporal long-short term memory network for location prediction. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, 2341–2347. https://doi.org/10.24963/ijcai.2018/324

Kumar, K. K., Rajagopalan, B., Hoerling, M., Bates, G., & Cane, M. (2006). Unraveling the mystery of Indian monsoon failure during El Niño. *Science*, *314*(5796), 115–119. https://doi.org/10.1126/science.1131152

Lee, E., Chase, T. N., Rajagopalan, B., Barry, R. G., Biggs, T. W., & Lawrence, P. J. (2009). Effects of irrigation and vegetation activity on early Indian summer monsoon variability. *International Journal of Climatology*, *29*(4), 573–581. https://doi.org/10.1002/joc.1721

Li, X., Zhou, Y., & Wang, F. (2022). Advanced information mining from ocean remote sensing imagery with deep learning. *Journal of Remote Sensing*, *2022*, 9849645. https://doi.org/10.34133/2022/9849645

Li, Y., Xü, K., Wu, Z., Zhu, Z., & Wang, Q. J. (2023). A statistical–dynamical approach for probabilistic prediction of sub-seasonal precipitation anomalies over 17 hydroclimatic regions in China. *Hydrology and Earth System Sciences*, *27*(22), 4187–4203. https://doi.org/10.5194/hess-27-4187-2023

Lin, J.-L., Weickman, K. M., Kiladis, G. N., Mapes, B. E., Schubert, S. D., Suarez, M. J., et al. (2008). Subseasonal variability associated with Asian summer monsoon simulated by 14 IPCC AR4 coupled GCMs. *Journal of Climate*, *21*(18), 4541–4567. https://doi.org/10.1175/2008JCLI1816.1

Meehl, G. A., Arblaster, J. M., Caron, J. M., Annamalai, H., Jochum, M., Chakraborty, A., & Murtugudde, R. (2012). Monsoon regimes and processes in CCSM4. Part I: The Asian–Australian monsoon. *Journal of Climate*, *25*(8), 2583–2608. https://doi.org/10.1175/jcli-d-11-00184.1

Melamud, O., Goldberger, J., & Dagan, I. (2016). *context2vec: Learning generic context embedding with bidirectional LSTM*. Association for Computational Linguistics.

Meng, Z., Zhou, L., Murtugudde, R., Yang, Q., Pujiana, K., & Xi, J. (2022). Tropical oceanic intraseasonal variabilities associated with central Indian Ocean mode. *Climate Dynamics*, *58*(3), 1107–1126. https://doi.org/10.1007/s00382-021-05951-1

Paul, S., Ghosh, S., Oglesby, R., Pathak, A., Chandrasekharan, A., & Ramsankaran, R. (2016). Weakening of Indian summer monsoon rainfall due to changes in land use land cover. *Scientific Reports*, *6*(1), 32177. https://doi.org/10.1038/srep32177

Peña, M., Kalnay, E., & Cai, M. (2003). Statistics of locally coupled ocean and atmosphere intraseasonal anomalies in Reanalysis and AMIP data. *Nonlinear Processes in Geophysics*, *10*(3), 245–251. https://doi.org/10.5194/npg-10-245-2003

Qin, J., Zhou, L., Ding, R., & Li, B. (2022). Predictability limit of monsoon intraseasonal precipitation: An implication of central Indian Ocean Mode. *Frontiers in Marine Science*, *8*. https://doi.org/10.3389/fmars.2021.809798

Qin, J., Zhou, L., Li, B., & Murtugudde, R. (2020). Simulation of central Indian Ocean Mode in S2S models. *Journal of Geophysical Research: Atmospheres*, *125*(21), e2020JD033550. https://doi.org/10.1029/2020jd033550

Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., & Schlax, M. G. (2007). Daily high-resolution-blended analyses for sea surface temperature. *Journal of Climate*, *20*(22), 5473–5496. https://doi.org/10.1175/2007jcli1824.1

Robock, A., Mu, M., Vinnikov, K., & Robinson, D. (2003). Land surface conditions over Eurasia and Indian summer monsoon rainfall. *Journal of Geophysical Research*, *108*(D4). https://doi.org/10.1029/2002JD002286

Roxy, M. K., Ghosh, S., Pathak, A., Athulya, R., Mujumdar, M., Murtugudde, R., et al. (2017). A threefold rise in widespread extreme rain events over central India. *Nature Communications*, *8*(1), 708. https://doi.org/10.1038/s41467-017-00744-9

19448007, 2024, 24, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024GL112308 by Lei Zhou - Shanghai Jiaotong University , Wiley Online Library on [11/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Sabeerali, C. T., Rao, S. A., George, G., Rao, D. N., Mahapatra, S., Kulkarni, A., & Murtugudde, R. (2014). Modulation of monsoon intraseasonal oscillations in the recent warming period. *Journal of Geophysical Research: Atmospheres*, *119*(9), 2013JD021261. https://doi.org/10.1002/2013jd021261

Saha, S. K., Halder, S., Suryachandra Rao, A., & Goswami, B. N. (2012). Modulation of ISOs by land-atmosphere feedback and contribution to the interannual variability of Indian summer monsoon. *Journal of Geophysical Research*, *117*(D13). https://doi.org/10.1029/2011jd017291

Saha, S. K., Hazra, A., Pokhrel, S., Chaudhari, H. S., Sujith, K., Rai, A., et al. (2019). Unraveling the mystery of Indian summer monsoon prediction: Improved estimate of predictability limit. *Journal of Geophysical Research: Atmospheres*, *124*(4), 1962–1974. https://doi.org/10.1029/2018JD030082

Sahai, A. K., Abhilash, S., Chattopadhyay, R., Borah, N., Joseph, S., Sharmila, S., & Rajeevan, M. (2014). High-resolution operational monsoon forecasts: An objective assessment. *Climate Dynamics*, *44*(11–12), 3129–3140. https://doi.org/10.1007/s00382-014-2210-9

Shukla, R. P. (2014). The dominant intraseasonal mode of intraseasonal South Asian summer monsoon. *Journal of Geophysical Research: Atmospheres*, *119*(2), 635–651. https://doi.org/10.1002/2013jd020335

Singh, D., Ghosh, S., Roxy, M. K., & McDermid, S. (2019). Indian summer monsoon: Extreme events, historical changes, and role of anthropogenic forcings. *WIREs Climate Change*, *10*(2), e571. https://doi.org/10.1002/wcc.571

Singh, D., Tsiang, M., Rajaratnam, B., & Diffenbaugh, N. S. (2014). Observed changes in extreme wet and dry spells during the South Asian summer monsoon season. *Nature Climate Change*, *4*(6), 456–461. https://doi.org/10.1038/nclimate2208

Sperber, K. R., & Annamalai, H. (2008). Coupled model simulations of boreal summer intraseasonal (30–50 day) variability, Part 1: Systematic errors and caution on use of metrics. *Climate Dynamics*, *31*(2), 345–372. https://doi.org/10.1007/s00382-008-0367-9

Suhas, E., Neena, J. M., & Goswami, B. N. (2013). An Indian monsoon intraseasonal oscillations (MISO) index for real time monitoring and forecast verification. *Climate Dynamics*, *40*(11–12), 2605–2616. https://doi.org/10.1007/s00382-012-1462-5

Tropical Rainfall Measuring Mission (TRMM). (2011). TRMM (TMPA) rainfall estimate L3 3 hour 0.25 degree x 0.25 degree V7 [Dataset], Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC). https://doi.org/10.5067/TRMM/TMPA/3H/7

Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., et al. (2017). The Subseasonal to Seasonal (S2S) prediction project database. *Bulletin America Meteorology Social*, *98*(1), 163–173. https://doi.org/10.1175/bams-d-16-0017.1

Vittal, H., Ghosh, S., Karmakar, S., Pathak, A., & Murtugudde, R. (2016). Lack of dependence of Indian summer monsoon rainfall extremes on temperature. *An Observational Evidence*, *6*(1), 31039. https://doi.org/10.1038/srep31039

Waliser, D. E., Stern, W., Schubert, S., & Lau, K. M. (2003). Dynamic predictability of intraseasonal variability associated with the Asian summer monsoon. *Quarterly Journal of the Royal Meteorological Society*, *129*(594), 2897–2925. https://doi.org/10.1256/qj.02.51

Wang, B., Lee, J. Y., Kang, I. S., Shukla, J., Park, C. K., Kumar, A., et al. (2008). Advance and prospectus of seasonal prediction: Assessment of the APCC/CliPAS 14-model ensemble retrospective seasonal prediction (1980–2004). *Climate Dynamics*, *33*(1), 93–117. https://doi.org/10.1007/s00382-008-0460-0

Wang, B., Xiang, B., Li, J., Webster, P. J., Rajeevan, M. N., Liu, J., & Ha, K.-J. (2015a). Rethinking Indian monsoon rainfall prediction in the context of recent global warming. *Nature Communications*, *6*(1), 7154. https://doi.org/10.1038/ncomms8154

Wang, Y., Huang, M., Zhu, X., & Zhao, L. (2016). *Attention-based LSTM for aspect-level sentiment classification*. Association for Computational Linguistics.

Webster, P. J., Magana, V. O., Palmer, T. N., Shukla, J., Tomas, R. A., Yanai, M., & Yasunari, T. (1998). Monsoons: Processes, predictability, and the prospects for prediction. *Journal of Geophysical Research*, *103*(C7), 14451–14510. https://doi.org/10.1029/97jc02719

Xi, J., Zhou, L., Murtugudde, R., & Jiang, L. (2015). Impacts of intraseasonal SST anomalies on precipitation during Indian summer monsoon. *Journal of Climate*, *28*(11), 4561–4575. https://doi.org/10.1175/jcli-d-14-00096.1

Zhang, K., Li, J., Zhu, Z., & Li, T. (2021). Implications from subseasonal prediction skills of the prolonged heavy snow event over southern China in early 2008. *Advances in Atmospheric Sciences*, *38*(11), 1873–1888. https://doi.org/10.1007/s00376-021-0402-x

Zheng, G., Li, X., Zhang, R.-H., & Liu, B. (2020). Purely satellite data-driven deep learning forecast of complicated tropical instability waves. *Science Advances*, *6*(29), eaba1482. https://doi.org/10.1126/sciadv.aba1482

Zhou, L., Murtugudde, R., Chen, D., & Tang, Y. (2017a). A central Indian Ocean Mode and heavy precipitation during the Indian summer monsoon. *Journal of Climate*, *30*(6), 2055–2067. https://doi.org/10.1175/JCLI-D-16-0347.1

Zhou, L., Murtugudde, R., Chen, D., & Tang, Y. (2017b). Seasonal and interannual variabilities of the central Indian Ocean Mode. *Journal of Climate*, *30*(16), 6505–6520. https://doi.org/10.1175/JCLI-D-16-0616.1

Zhou, L., Murtugudde, R., Neale, R. B., & Jochum, M. (2018). Simulation of the Central Indian Ocean Mode in CESM: Implications for the Indian summer monsoon system. *Journal of Geophysical Research: Atmospheres*, *123*(1), 58–72. https://doi.org/10.1002/2017jd027171

Zhu, Y., Zhang, R.-H., Moum, J. N., Wang, F., Li, X., & Li, D. (2022). Physics-informed deep-learning parameterization of ocean vertical mixing improves climate simulations. *National Science Review*, *9*(8). https://doi.org/10.1093/nsr/nwac044

Zhu, Z., Li, T., Hsu, P.-c., & He, J. (2015). A spatial–temporal projection model for extended-range forecast in the tropics. *Climate Dynamics*, *45*(3), 1085–1098. https://doi.org/10.1007/s00382-014-2353-8

Zhu, Z., Zhou, Y., Jiang, W., Fu, S., & Hsu, P. (2023). Influence of compound zonal displacements of the South Asia high and the western Pacific subtropical high on Meiyu intraseasonal variation. *Climate Dynamics*, *61*(7), 3309–3325. https://doi.org/10.1007/s00382-023-06726-6

19448007, 2024, 24, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024GL112308 by Lei Zhou - Shanghai Jiaotong University , Wiley Online Library on [11/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License