Using deep learning to predict the East Asian summer monsoon

Yuheng Tang\textsuperscript{1,2} and Anmin Duan\textsuperscript{1,2,∗}

\textsuperscript{1} State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, People’s Republic of China
\textsuperscript{2} College of Earth Science, University of Chinese Academy of Sciences, Beijing, People’s Republic of China

∗ Author to whom any correspondence should be addressed.

E-mail: amduan@lasg.iap.ac.cn

Keywords: machine learning, heatmap, East Asian summer monsoon

Supplementary material for this article is available online

Abstract

Accurate prediction of the East Asian summer monsoon (EASM) is beneficial to billions of people’s production and lives. Here, a convolutional neural network (CNN) and transfer learning are used to predict the EASM. The results of the constructed CNN regression model show that the prediction of the CNN regression model is highly consistent with the reanalysis dataset, with a correlation coefficient of 0.78, which is higher than that of each of the current state-of-the-art dynamic models. The heatmap method indicates that the robust precursor signals in the CNN regression model agree well with previous theoretical studies and can provide the quantitative contribution of different signals for EASM prediction. The CNN regression model can predict the EASM one year ahead with a confidence level above 95%. The above method can not only improve the prediction of the EASM but also help to identify the involved physical predictors.

1. Introduction

The East Asian summer monsoon (EASM) is a significant component of Asian monsoons and has a very large impact on precipitation and temperature in East Asia. Predicting the EASM in advance is urgently needed for the country and individuals. However, due to the complex physical mechanism of the EASM (Wang 2006), an in-depth understanding and prediction of the EASM is still a major challenge.

The existing precursor predictors of the EASM focus mainly on sea surface temperature (SST) in adjacent and remote oceans, such as the Indian Ocean (Yang et al 2007, Li et al 2008, Cao et al 2016), Atlantic (Chen et al 2018, Jin and Huo 2018, Choi and Ahn 2019) and Pacific (Huang and Wu 1989, Huang and Sun 1992, Fan et al 2013). In addition, the preceding snow cover/depth (Ye and Gao 1979, Wu and Qian 2003) and thermal conditions (Ye and Gao 1979, Wu et al 2007, Wang et al 2013) over the Tibetan Plateau (TP), together with snow depth/cover in the Eurasian continent (Zuo et al 2015), also influence the EASM to a certain degree. Some atmospheric internal modes, such as the Arctic Oscillation (AO) and Southern Hemisphere Annular Mode (SAM), can modulate the variation in the EASM through a variety of mechanisms (Ju et al 2005, Nan et al 2009, Gong et al 2011, He et al 2017). However, coupling effects always exist among different precursor factors due to land–air–sea interaction processes (Feng et al 2014, Cui et al 2015).

Traditionally, dynamic models, statistical models and statistical dynamic models have been utilized to predict the EASM (Wu et al 2009, Fan et al 2012, Wu and Yu 2015, Nie and Guo 2019, Zhou et al 2020). The predictions of dynamic models refer to the use of various state-of-the-art models to predict the EASM and highly depend on the performance of the model simulation. The predictions of statistical models rely on the input predictors and the model choice. Statistical dynamic models are mostly postprocessed based on dynamic model output to improve model prediction skills. To date, existing statistical models have failed to simultaneously consider the influence of the three oceans (Indian Ocean, Pacific and Atlantic), TP and polar regions on the EASM, and almost all of the models cannot accurately identify the specific key regions that significantly influence the EASM.
With the development of artificial intelligence, deep learning methods have been applied in weather and climate, such as the forecasting of precipitation (Xue et al 2020), air temperature (Shin et al 2020), El Niño/Southern Oscillation (ENSO; Ham et al 2019, Pal et al 2020) and weather patterns (Chattopadhyay et al 2020a, 2020b), representing subgrid processes in climate models (Rasp et al 2018), and revising the output of numerical weather prediction models (Li et al 2019a). The advantages of deep learning over traditional methods and the prospects for applications in Earth System Science have been thoroughly studied (Reichstein et al 2019). Because of the exponential increase in the amount of available data in Earth System Science, as well as the superiority of deep learning methods in image and sequence processing, Earth System Science data, which mainly exist in spatial-temporal fields and sequences, can be used to mine information through deep learning. In this paper, a deep learning method named convolutional neural network (CNN) is employed. It exhibits excellent performance in solving not only classification problems (Xue et al 2020, Chattopadhyay et al 2020a, 2020b) such as weather patterns and precipitation classification but also regression problems (Ham et al 2019, Xue et al 2020) such as precipitation and ENSO index prediction. Deep learning may hold great promise to address the challenges of EASM prediction. It employs sophisticated networks to learn complicated and nonlinear relationships between input and output data. In addition, as a deep learning method, CNN need to be fed large datasets and labels to learn the correct features. Because of the good performance of shallow CNN (Lecun et al 1998, Krizhevsky et al 2017), using a CNN with fewer layers and transfer learning techniques (Yosinski et al 2014) can be beneficial for training the CNN model with relatively few data. Because deep learning methods lack physical support (only data-driven methods), visualization methods will be more conducive to the interpretation of these methods (Ham et al 2019, Davenport and Diffenbaugh 2021).

In this study, we train a CNN regression model through the World Climate Research Programme Coupled Model Intercomparison Project Phase 6 (CMIP6; table S1 available online at stacks.iop.org/ERL/16/124006/mmedia)) and reanalysis dataset. Our aim is to use the deep learning method to improve the prediction skill of the EASM, especially long-term prediction, and to use a visualization method (Zhou et al 2016) to further understand the sources of the high prediction skill. The results suggest that the novel deep learning method has remarkable predictive ability for the EASM, and the visualization method shows that the significant sources of the prediction skill of the EASM have good agreement with previous theoretical studies. These results, therefore, can provide a reference for operational forecasts and further improvement of forecasting skills.

2. Materials and methods

2.1. Data

In order to train the CNN regression model, we use the multimodel outputs from CMIP6 as the training dataset and guarantee sufficient CMIP6 samples for training the CNN regression model (figure S1). The description of all 39 models is provided in table S1. Surface temperature (TS), surface upward sensible heat flux (HFSS) and zonal wind (UA) are used. All multimodel outputs are monthly data of the historical experiment from 1861 to 2004. To facilitate the analysis, all datasets are interpolated to $2.5^\circ \times 2.5^\circ$ grids with the bilinear interpolation method. Thus, the spatial fields over $0^\circ$–$360^\circ$ E, $87.5^\circ$–$90^\circ$ N of all simulation and reanalysis datasets are 72 grid points in latitude and 144 grid points in longitude.

UA is used to calculate the EASM index (Wang and Fan 1999). It is defined as the regional average difference in the zonal wind anomaly at 850 hPa between $5^\circ$–$15^\circ$ N, $90^\circ$–$130^\circ$ E and $22.5^\circ$–$32.5^\circ$ N, $110^\circ$–$140^\circ$ E. Among many different EASM indices, this simple index is best correlated with the leading principal component of the EASM (Wang et al 2008), and the positive phase of this mode usually shows a north–south dipole precipitation pattern of dry anomalies in the northern South China Sea and Philippine Sea and wet anomalies along the Yangtze River valley to southern Japan. The correlation between the EASM index and regional average precipitation index ($26^\circ$–$37^\circ$ N, $105^\circ$–$122^\circ$ E; Wang et al 2008) in four reanalysis datasets are from $-0.49$ to $-0.70$ and all passed the 99% significance test (figure S2). As shown in figure S2, these four reanalysis datasets can effectively describe the variations in the EASM index, including the European Centre for Medium-Range Weather Forecasts (ECWMF) fifth-generation reanalysis (ERA5) product, Japanese 55 year reanalysis (JRA55), second Modern-Era Retrospective analysis for Research and Applications (MERRA2) and the University of Colorado Boulder’s Cooperative Institute for Research in Environmental Sciences (CIRES) and the National Oceanic and Atmospheric Administration (NOAA) 20th century reanalysis version 3 (NOAA 20CR3; Slivinski et al 2019). NOAA 20CR3 is used due to its long time range (from 1836 to 2015). In addition, the variables used are the same as CMIP6 data. We use the bilinear interpolation method to interpolate the data from $1^\circ \times 1^\circ$ grids to $2.5^\circ \times 2.5^\circ$ grids. A time period from 1837 to 1980 is used for the NOAA 20CR3 data. We use the ERA5 monthly data from 1986 to 2019 as the test dataset in order to test the performance of the CNN regression model. The variables are TS, HFSS and eastward wind (the same as UA). All ERA5 surface sensible heat flux data are divided by $-86400$ to match the unit of HFSS in the CMIP6 and NOAA 20CR3 data. Additionally ERA5 variables are interpolated from $0.25^\circ \times 0.25^\circ$ grids to $2.5^\circ \times 2.5^\circ$ grids by the
2.2. The CNN regression model

As shown in Figure 1, we construct two CNN structures following the AlexNet architecture (Krizhevsky et al. 2017), and the CNN regression model consists of these two state-of-the-art CNN structures and works better than either of them. All of the CNN code is written by using the TensorFlow/Keras library. Inspired by Ham et al. (2019), the first CNN model (at the top of Figure 1) has four convolutional layers with 64 filters per layer, one fully connected (FC) layer, and one output layer. The kernel size of each filter is $5 \times 7$. We use the zero padding in the borders of the feature map in each convolutional layer, and the activation function is the Leaky Rectified Linear Unit (LReLU, with fixed parameter 0.3). The batch normalization layer is also used in each convolutional layer. Following the first, second and fourth convolutional layers, there is a max-pooling layer with a kernel size of $2 \times 2$. The second CNN model is similar to the first one, but the number of filters in each convolutional layer is 16, 32, 48 and 64. The second CNN model has two FC layers, with 128 and 32 neurons. On the right and bottom of each feature map, there are length and width dimensions of the feature map. The output of each CNN model is averaged to obtain the final output.

Figure 1. The structure of the CNN regression model, which includes the first (top) and second (bottom) CNN models. The first CNN model includes four convolutional layers, with 64 filters each, and one FC layer, with 128 neurons and one output layer. The number of convolutional layers is in front of the @ symbol. After the @ symbol, the kernel size is $5 \times 7$. Following the first, second and fourth convolutional layers, there is a max-pooling layer with a kernel size of $2 \times 2$. The second CNN model is similar to the first one, but the number of filters in each convolutional layer is 16, 32, 48 and 64. The second CNN model has two FC layers, with 128 and 32 neurons. On the right and bottom of each feature map, there are length and width dimensions of the feature map. The output of each CNN model is averaged to obtain the final output.
(Wang et al. 2013, Duan et al. 2017, Sun et al. 2019). The prediction skills of using both TS and HFSS were higher than those of using only TS or HFSS (table S2). Therefore, the TS and HFSS anomalies over 0°–360° E, 87.5° S–90° N for three consecutive months are used as the input predictors. The correlation analysis shows a robust relationship between the two predictors and the EASM (figure S3).

Because the CNN regression model is a deep learning method, a major limitation is to prepare proper climate data to feed into the CNN regression model for learning the correct features and predicting well. Following the methodology of Ham et al. (2019), due to the lack of sufficient observational data, 39 CMIP6 historical simulations are used to train the CNN regression model. In addition, NOAA 20CR3 data from 1837 to 1980 are utilized to fine-tune the CNN regression model. We reduce the learning rate and batch size to train the CNN regression model (train only the last layer) by the NOAA 20CR3 dataset to improve prediction skills, and this method is also called the transfer learning method (Yosinski et al. 2014). The ERA5 reanalysis test dataset for evaluating prediction skills are used from 1986 to 2019 to leave a five-year gap between training data and validation data to remove the influence of oceanic long term memory. To ensure the robustness and reproducibility of the results, both CNN models are run ten times with random initialization and averaged collectively.

2.3. The heatmap method
To improve the interpretability of the deep learning method, a visualization method (Zhou et al. 2016) is used. Ham et al. (2019) developed a heat map method based on the class activation map of Zhou et al. (2016), and we adapted the heat map method to our CNN regression model as follows:

\[
v(x, y)_{F_i, l, m} = \text{LReLU} \left[ \sum_{i=1}^{L_C} \theta(x, y)_{F_i, l, m} v(x, y)_{l, i, m} + b_{F_i, m} \right] / X_C Y_C, \quad (1)
\]

\[
h(x, y)_{\text{First}} = \sum_{m=1}^{M} v(x, y)_{F_i, l, m} \theta_{O, m} + b_0 / X_C Y_C. \quad (2)
\]

\[
h(x, y)_{\text{Second}} = \sum_{n=1}^{N} \left\{ \text{LReLU} \left[ \sum_{m=1}^{M} v(x, y)_{F_i, l, m} \theta_{F_i, m, n} \right] + b_{F_i, n} / X_C Y_C \right\} + b_0 / X_C Y_C; \quad (3)
\]

\[
\text{LReLU}(x) = \begin{cases} 
x & x > 0 \\
0.3x & x \leq 0
\end{cases}, \quad (4)
\]

where \(v(x, y)_{F_i, l, m}\) denotes the \(m\)th neuron in the FC layer \(F_i\). \(v(x, y)_{l, i, m}\) is the output of the last convolutional layer \(C\), which still maintains the spatial structure of the data. \(L_C\) is the number of feature maps in layer \(C\), and \(\theta(x, y)_{F_i, l, m}\) is the weight of the FC layer between the \(l\)th feature map and the \(m\)th neuron in the \(F_i\) layer at grid point \((x, y)\). \(b_{F_i, m}\) is the bias of the \(m\)th neuron in the FC layer \(F_i\), while \(X_C\) and \(Y_C\) are the dimensions of the feature map in layer \(C\). \(\text{LReLU}\) is the activation, which is defined by equation (4). For the first CNN model, the output of the heat map is \(h(x, y)_{\text{First}}\), and \(b_0\) and \(\theta_{O, m}\) denote the bias and weight between the \(F_1\) layer and the output layer, respectively. For the second CNN model, \(h(x, y)_{\text{Second}}\) is the output of the heat map. FC layer \(F_2\) has \(N\) neurons behind FC layer \(F_1\), \(\theta_{F_2, m, n}\) is the weight between the \(m\)th neuron of layer \(F_1\) and \(n\)th neuron of layer \(F_2\), and \(b_{F_i, n}\) is the bias of the \(n\)th neuron of layer \(F_2\). \(\theta_{O, n}\) and \(b_0\) are the weight and bias between layer \(F_2\) and the output layer, respectively. The final heat map is the sum of the first and second outputs of the heat map.

To improve the prediction skill, we mask the significant regions in the heat map, which means that we use only data in the shaded regions to train the CNN regression model. Other regions are filled with useless data such as zero. In this way, we can prevent some irrelevant signals from reducing the prediction accuracy and highlight precursor signals of the EASM.

For the skill metric, the Pearson correlation coefficient is used as follows:

\[
corr = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}}. \quad (5)
\]

where \(x_i\) and \(y_i\) represent the predicted and observed EASM indices, respectively, and \(x\) and \(y\) are the time means of \(x_i\) and \(y_i\), respectively.

3. Results

3.1. Heat map analysis
By using the TS and HFSS anomalies in boreal spring (March–May) to train the CNN regression model, it is indicated that this model can accurately predict the EASM, and the correlation coefficient between the EASM index sequences of the CNN regression model predictions and reanalysis data reaches 0.776 (figure S4). Especially in the relative trend (increasing or decreasing relative to the previous year) of the EASM index sequence, the CNN regression model performs well in all 34 years except the seven years of 1988, 1996, 2002, 2007, 2013, 2016 and 2017. Based on the threshold (0.75 times the standard deviation) of the sequence predicted by the CNN regression model, we select seven anomalously strong EASM years (1986, 1990, 1994, 1997, 2002, 2014 and 2018) and six anomalously weak EASM years (1987, 1988, 1998, 2005, 2008 and 2010).

To visually analyze the deep learning results, we first need to diagnose the uncertainty of CMIP6 historical simulations. As shown in figure S5, the standard deviation between CMIP6 model historical
simulations increases with latitude. In boreal spring, there are larger model deviations in the Antarctic than in the tropics and subtropics. However, the regions with large standard deviations of HFSS are mainly in the marginal seas and part of the land areas, such as high-elevation areas and northeastern South America. For TS or HFSS, the models have less uncertainty for the simulation on the ocean than the simulation on the land. In general, with the exception of HFSS during boreal winter in marginal seas and part of the Antarctic, model simulations in other regions have good consistency.

Through the heatmap method (Zhou et al. 2016) of deep learning visualization, we can analyze the contributions from different regions to the predicted EASM index. Owing to the characteristics of the heat map method, the positive or negative value of the shaded region does not mean a positive or negative correlation but represents only the relative magnitude of the contribution to the predicted EASM index. Here, TS and HFSS anomalies in boreal spring are used as predictors, so the contributions of different regions mainly indicate the effects of TS or HFSS anomalies. As mentioned above, the composite analysis method of the heat map is used in the anomalous years (seven anomalously high EASM index years and six anomalously low EASM index years; figure 2).

The CNN regression model is an entirely data-driven statistical model, but it can be linked with some physical processes by using this heat map method. For example, there are two remarkable regions in the tropical central and western Pacific and tropical Atlantic (figure 2), and this robust relationship between the two tropical oceans and the EASM can be detected in figures S3(e)–(g) and (l)–(n). Previous studies have also indicated that during boreal spring, the significant signals in the tropical Pacific (Huang and Wu 1989, Huang and Sun 1992, Fan et al. 2013) and tropical Atlantic (Chen et al. 2018, Jin and Huo 2018, Choi and Ahn 2019) greatly contribute to the subsequent EASM. Since the CNN regression model is trained by TS and HFSS anomalies year by year, the interannual variabilities appear more frequently than the inter decadal variabilities. Due to the seasonal phase lock of ENSO, the impacts of ENSO on the EASM are mainly exerted by the tropical western Pacific anticyclone and the Indian Ocean during boreal spring (Wang et al. 2000, Xie et al. 2009). In addition, these signals match well in figures 2 and 3. In fact, in the subtropics of the Atlantic and Pacific, the significant shaded region may reflect the influence of inter decadal variabilities, such as the Atlantic Multidecadal Oscillation (Lu et al. 2006) and Pacific Decadal Oscillation/Interdecadal Pacific Oscillation (Chen et al. 2013), on the subsequent EASM. In the tropical Indian Ocean, the precursor TS and HFSS show a notable correlation with the EASM (figure S3). The remarkable shaded region in the heat map is consistent with previous studies about the Indian Ocean precursor signal of the EASM (Yang et al. 2007, Liu and Duan 2017). Moreover, the connection between precipitation in North Africa and East Asia has been
Figure 3. The observed and predicted EASM index results. The predicted results are the outputs of the CNN regression model by using boreal spring TS and HFSS anomalies of ERA5 data as the predictors from 1986 to 2019. The black bold line is the observed sequence calculated by ERA5 data, and the parameter $r$ is the correlation coefficient between different predicted results and the observed sequence. The CNN regression model result is obtained by using boreal spring mask input variables. MME is the ensemble mean of the four Copernicus Climate Change Service (C3S) seasonal forecast systems (ECMWF-SEAS5, CMCC-SPS3.5, DWD-GCFS2.1 and Météo-France-System7).

As mentioned above, the shaded regions in figure 2 contribute greatly to predicting the EASM, so we mask all the shaded regions (figure S7, section 2) of the input variables to train the CNN regression model to improve the prediction skill of the EASM. We compare our results with eight state-of-the-art dynamic seasonal prediction models (figure 3). The correlation of the CNN regression model (0.783) is better than that of any other model in this paper, even slightly better than those of the ECMWF-SEAS5 (0.775) and DWD-GCFS2.1 (0.764) models. In particular, for the sign (positive or negative) of the EASM index, the prediction sequence of the CNN regression model matches well with the observed sequence, except for seven years (1992, 1999, 2000, 2011, 2015, 2017 and 2019). In the above seven years, there are only three years (1999, 2015 and 2017) in which the difference from the observed values is greater than one standard deviation, which indicates that for the prediction of the EASM index sign, only three of the 34 years are out of the error range. The ECMWF-SEAS5 and DWD-GCFS2.1 models are the second and third-ranked seasonal prediction systems. If we focus on the prediction of the last four years, the prediction of the CNN regression model is most similar to the observed sequence.

3.2. Comparison with multimodel results

As mentioned above, the CNN regression model can accurately predict the EASM three months in advance. To obtain longer lead-time prediction, we use TS and HFSS anomalies at different initial times to train the CNN regression model (figure 4). Because ECMWF-SEAS5, CMCC-SPS3.5, DWD-GCFS2.1 and Météo-France-System7 do not have long-term prediction data, we use only CanCM4, FGOALS-f2, CanCM4i and CanSIPv2 dynamic seasonal to interannual prediction results. Compared...
with other models, the CNN regression model can give a valid prediction back to December(−1)–January–February (D(−1)JF), which exceeds the 99.9% confidence level. Before D(−1)JF, the correlation of the CNN regression model is also obviously better than that of any other model in this paper. The standard deviation of the CNN regression model ensemble members is also smaller than that of the four dynamic models. The second-best model is CanSIPv2, which exceeds the 99.9% confidence level one month later than the CNN regression model. Using JFM data as the initial field, the result of CanSIPv2 is slightly better than that of the CNN regression model, while for other initial fields, the prediction of CanSIPv2 is always worse than that of the CNN regression model. The FGOALS-f2 model gives the second best correlation by using the boreal spring initial field, but with increasing lead time, the correlation drops rapidly. It is strange that for the initial field before November(−1)–December(−1)–January (ND(−1)J), the correlation of FGOALS-f2 is even slightly higher than that for the initial field of ND(−1)J. If we use the initial field before ND(−1)J, the correlation of CanCM4i is negative, and it is the only negative correlation model in all five models. If we focus on the lead times in which the correlation between the prediction results and the observation exceeds the moderate 95% confidence level, the CNN regression model can robustly predict the EASM even one year ahead (June(−1)–July(−1)–August(−1); JJA(−1)) and is better than any other prediction model. A previous study also employed dynamic model predictions to reach the one-year lead-time (Takaya et al 2021), and the short-term prediction (approximately three months ahead) performance was still significantly lower than our result.

To further explain the source of the high prediction skills, we take the EASM in 1998 as an example for analysis. As shown in figures S8(c) and (d) the CNN regression model can effectively capture the negative EASM index in 1998, although the strength of the EASM in 1998 is underestimated by the CNN regression model. The strong El Niño during 1997/98 exerts a great influence on the subsequent EASM (Huang et al 2000). Although ENSO is an important atmospheric variability in the tropics, Elliott et al (2001) have shown that the anomalous variability in Atlantic SST in 1997/98 is not only affected by ENSO but also independent of ENSO. Only after considering the changes in Atlantic SST can the simulated circulation anomalies in the Indian Ocean and Northwest Pacific in the summer of 1998 be more consistent with the observations (Lu and Dong 2005). That is, although the signal from the equatorial Pacific has a very significant impact on the subsequent EASM, the role of SST in the Atlantic cannot be ignored. Such a strong El Niño signal can also affect the SST of the tropical Indian Ocean through Walker circulation and ocean dynamic processes (Yu and Rienecker 1999, Fischer et al 2005). The heat maps in figure S8 indicate that the high prediction skills of the CNN regression model for one year lead time mainly come from the signals in the tropical eastern Pacific, tropical western Atlantic. Therefore, we speculate that the reason for such high prediction skills of EASM is the effective capture of ENSO and its associated variabilities in the CNN regression model. The conclusions are similar to the conclusions of Takaya et al (2021).
4. Conclusions and discussions

In this work, we construct a CNN regression model to predict the EASM and analyze the relative contributions from different regions for predicting the EASM. The CNN regression model performs well in predicting the EASM, and the prediction result can exceed the 95% confidence level even one year in advance. The correlation between the observed EASM index sequence and the CNN regression model is obviously higher than that between the observed EASM index sequence and any other dynamic model and system, except at the initial time of JFM (the CNN regression model is only slightly worse than CanSIPv2). Constructing the CNN regression model with boreal spring TS and HFSS anomalies, the correlation skill of the EASM index can reach 0.783, which is larger than the correlation skill of the EASM index of any other model reported in this paper. By using composite analysis in a heat map, we find that the regions with significant contributions to the subsequent EASM match well with the key regions identified by previous theoretical studies and are consistent with the correlation analysis (figures S3 and S6; Nan et al 2009, Gong et al 2011, Fan et al 2013, Hu and Duan 2015, Choi and Ahn 2019). This helps us understand the relative contributions of different precursor predictors to the prediction of the EASM and is beneficial for analyzing the physical mechanism between the predictors and the subsequent EASM.

We should keep in mind that the prediction results of the CNN regression model explain only less than 2/3 of the variance. Since the original data used have not been subject to filtering and other operation instead of directly training the CNN regression model, the data will automatically emphasize the interannual signals (the interannual signals appear more frequently than the interdecadal signals). There is a basic correspondence between the information displayed by the heat maps and the observational theoretical analysis, and more detailed differences need to be considered in the future research.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

The text ends with an acknowledgment section and statement that includes:

- The authors declare that they have no competing interests.
- The CMIP6, NOAA 20CR3, ERA5, and HadISST data used in this paper are from these websites (https://esgf-node.llnl.gov/search/cmip6/, https://psl.noaa.gov/data/)
- The work was supported by the National Natural Science Foundation of China (41725018) and National Natural Science Foundation of China (42030602).

ORCID iDs

Yuheng Tang [https://orcid.org/0000-0002-9639-3958]
Anmin Duan [https://orcid.org/0000-0002-3022-2582]

References

Armitage T W K, Bacon S and Kwok R 2018 Arctic sea level and surface circulation response to the Arctic Oscillation Geophys. Res. Lett. 45 6576–84
Bao Q, Wu X, Li J, Wang J, He B, Wang X, Liu Y and Wu G 2019 Outlook for El Niño and the Indian Ocean Dipole in autumn-winter 2018–2019 Chin. Sci. Bull. 64 73–78
Cao J, Gui S, Su Q and Yang Y 2016 The variability of the Indian–East Asian summer monsoon interface in relation to the spring seesaw mode between the Indian Ocean and the Central-Western Pacific J. Clim. 29 5027–40
Chattopadhyay A, Hassanzadeh P and Pasha S 2020a Predicting clustered weather patterns: a test case for applications of convolutional neural networks to spatio-temporal climate data Sci. Rep. 10 1317
Chattopadhyay A, Nabizadeh E and Hassanzadeh P 2020b Analog forecasting of extreme-causing weather patterns using deep learning J. Adv. Model. Earth Syst. 12 e2019MS001958
Chen S, Wu R and Chen W 2018 Modulation of spring northern tropical Atlantic sea surface temperature on the El Niño-Southern Oscillation-East Asian summer monsoon connection Int. J. Climatol. 38 5020–9
Chen W, Feng J and Wu R 2013 Roles of ENSO and PDO in the link of the East Asian winter monsoon to the following summer monsoon J. Clim. 26 622–35
Choi Y-W and Ahn J-B 2019 Possible mechanisms for the coupling between late spring sea surface temperature anomalies over tropical Atlantic and East Asian summer monsoon Clim. Dyn. 53 6995–7009
Cui Y, Duan A, Liu Y and Wu G 2015 Interannual variability of the spring atmospheric heat source over the Tibetan Plateau forced by the North Atlantic SSTA Clim. Dyn. 45 1617–34
Davenport F V and Diffenbaugh N S 2021 Using machine learning to analyze physical causes of climate change: a case study of US Midwest extreme precipitation Geophys. Res. Lett. 48 e2021GL093787
Duan A, Sun R and He J 2017 Impact of surface sensible heating over the Tibetan Plateau on the western Pacific subtropical high: a land–air–sea interaction perspective Adv. Atmos. Sci. 34 157–68
Wu Z et al 2009 An empirical seasonal prediction model of the east Asian summer monsoon using ENSO and NAO J. Geophys. Res. Atmos. 114
Wu Z and Yu L 2015 Seasonal prediction of the East Asian summer monsoon with a partial-least square model Clim. Dyn. 46 3067–78
Xie S-P, Hu K, Hafner J, Tokinaga H, Du Y, Huang G and Sampe T 2009 Indian Ocean capacitor effect on Indo–Western Pacific climate during the summer following El Niño J. Clim. 22 730–47
Xue M et al 2020 CNN-based near–real-time precipitation estimation from Fengyun-2 satellite over Xinjiang, China Atmos. Res. 250 105337
Yang J, Liu Q, Xie S-P, Liu Z and Wu L 2007 Impact of the Indian Ocean SST basin mode on the Asian summer monsoon Geophys. Res. Lett. 34 L02708
Ye D and Gao Y 1979 Meteorology of the Qinghai–Xizang (Tibet) Plateau (Beijing: Science) in Chinese

Yosinski J et al 2014 How transferable are features in deep neural networks? Adv. Neural Inf. Process. Syst. 27 3320–28
Yu L and Rienecker M M 1999 Mechanisms for the Indian Ocean warming during the 1997–98 El Niño Geophys. Res. Lett. 26 735–8
Zhou B et al 2016 Learning deep features for discriminative localization 2016 IEEE Conf. on Computer Vision and Pattern Recognition pp 2921–9
Zhou F, Ren H-L, Hu Z-Z, Liu M-H, Wu J and Liu C-Z 2020 Seasonal predictability of primary East Asian summer circulation patterns by three operational climate prediction models Q. J. R. Meteorol. Soc. 146 629–46
Zuo Z, Yang S, Zhang R, Xiao D, Guo D and Ma L 2015 Response of summer rainfall over China to spring snow anomalies over Siberia in the NCEP CFSv2 reforecast Q. J. R. Meteorol. Soc. 141 939–44