

Geophysical Research Letters^{*}

RESEARCH LETTER

10.1029/2023GL106676

Key Points:

- Traditional deep learning (DL) models often grapple with heavy rainfall forecasting
- New physical-informed DL models are proposed to address the problem of precipitation forecasting
- The new ω-GNN and ω-EGNN models outperform the NWP and other DL models, especially in forecasting heavy rainfall

Supporting Information:

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

Correspondence to:

Y. Wang and G. Huang, wangya@mail.iap.ac.cn; hg@mail.iap.ac.cn

Citation:

Chen, Y., Wang, Y., Huang, G., & Tian, Q. (2024). Coupling physical factors for precipitation forecast in China with graph neural network. *Geophysical Research Letters*, *51*, e2023GL106676. https://doi. org/10.1029/2023GL106676

Received 4 OCT 2023 Accepted 7 JAN 2024

© 2024. The Authors. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Coupling Physical Factors for Precipitation Forecast in China With Graph Neural Network

Yutong Chen^{1,2,3} , Ya Wang¹, Gang Huang^{1,2,3}, and Qun Tian⁴

¹State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China, ²Laboratory for Regional Oceanography and Numerical Modeling, Qingdao National Laboratory for Marine Science and Technology, Qingdao, China, ³College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing, China, ⁴Guangzhou Institute of Tropical and Marine Meteorology/Guangdong Provincial Key Laboratory of Regional Numerical Weather Prediction, CMA, Guangzhou, China

Abstract Precipitation exerts far-reaching impacts on both socio-economic fabric and individual well-being, necessitating concerted efforts in accurate forecasting. Deep learning (DL) models have increasingly demonstrated their prowess in forecasting meteorological elements. However, traditional DL prediction models often grapple with heavy rainfall forecasting. In this study, we propose physics-informed localized graph neural network models called ω -GNN and ω -EGNN, constrained by the coupling of physical variables and climatological background to predict precipitation in China. These models exhibit notable and robust improvements in identifying heavy rainfall while maintaining excellent performance in forecasting light rain by comparing to numerical weather prediction (NWP) and other DL models with multiple perturbation experiments in different data sets. Surprisingly, within a certain range, even when a DL model utilizes more input variables, GNN can still maintain its advantage. The methods to fuse physics into DL model demonstrated in this study may be promising and call for future studies.

Plain Language Summary We need accurate precipitation forecast for the preparedness of daily life, business and natural disaster and so on. In recent years, deep learning (DL) has been a hot topic in many sectors, including meteorological and climatological forecasting. Here we propose two DL models called ω -GNN and ω -EGNN, respectively, which integrate physical knowledge into DL methods, to predict precipitation in China. The two models show better behavior than today's operational numerical weather prediction (NWP) and other proven effective DL architectures in critical metrics, even with less input information. This indicates that our methods of physics-informed DL models are practicable.

1. Introduction

Precipitation has significant impacts on socio-economy and our daily life (Konapala et al., 2020; Lesk et al., 2016; Trenberth et al., 2003). Extremely heavy rainfall usually leads to disasters such as floods, landslides, and mud-rock flow, which threaten people's lives (O'Gorman, 2015; Westra et al., 2014). As a result, accurate precipitation forecast is urgently needed (Clark et al., 2016; Ding et al., 2020). However, the mainstream numerical weather prediction (NWP) models are not precise enough due to their inherent defects in the approximation of dynamic and physical processes (Bauer et al., 2015). Statistical modeling approaches, which can either post-process NWP output or directly forecast precipitation, are often used to avoid these drawbacks (Wilks, 2009).

In recent times, there has been an exponential growth in deep learning (DL) technology, a data-driven method that has proven successful in meteorological and climatological forecasting and prediction (e.g., Bi et al., 2022; Ham et al., 2019; Liu et al., 2023; Pathak et al., 2022; Ravuri et al., 2021; Zhang et al., 2023). In contrast to NWP, DL models do not explicitly include atmospheric dynamics, which may impact their performance and application prospects (Reichstein et al., 2019). Notably, DL models may struggle with the predictions for heavy rainfall events. The use of conditional generative models is an effective approach for improving heavy rainfall predictions, especially in nowcasting (Hess et al., 2022; Ravuri et al., 2021; Zhang et al., 2023). Furthermore, DL models may not comply with important physical coupling (Han et al., 2020) and hinder precipitation prediction. In such situations, DL models informed by physical priors may prove beneficial (Karniadakis et al., 2021; Kashinath et al., 2021).

Emerging Graph neural network (GNN) algorithms have shown promise in addressing these challenges (Kipf & Welling, 2017; Scarselli et al., 2009; Veličković et al., 2018). GNNs are designed to process graph-structured data and can effectively learn the evolution of a node of interest (NoI) by considering its neighboring nodes, making them a powerful tool for modeling complex interactions and imposing constraints on DL models. For example, Wu et al. (2021, 2022) utilized GNN to construct the spatial relation between rain gauges and forecast the rainfall in the East of China. Li et al. (2023) demonstrated the spatial interaction of observation stations in heatwave events. Other than spatial structure, GNNs can also be applied to represent the coupling of physical variables such as air-sea interaction (e.g., Mu et al., 2021).

In this study, we propose a novel approach called the GNN-coupled (ω -GNN) model that forecasts precipitation, which considers the coupling relationship of meteorological elements based on the omega equation and moist equation. Through comparison with several other DL models trained and validated by the same data set under identical conditions, we found that the ω -GNN model considering physical couplings significantly outperforms the models that do not incorporate physical processes, leading to an improved performance in precipitation forecasting.

This study is structured as follows: Section 2 introduces the data used in this study. Section 3 is a methodology, including the construction of the model. In Section 4, we show our experimental results. Finally, we give a conclusion and discussion in Section 5.

2. Data

2.1. Data Sets

The ECMWF-IFS prediction serves as input data for this study. We select the following six variables considering the omega equation and moist equation: total precipitation, temperature, u-wind, y-wind, geopotential height, and relative humidity. The latter five parameters are at 500 hPa. Total precipitation has been converted to 6h-accumulated precipitation and graded into five levels based on the following thresholds: 0.1 mm/6h, 2.5 mm/6h, 10 mm/6h, and 20 mm/6h. In this study, we utilize the ECMWF Reanalysis v5 (ERA-5; Hersbach et al., 2023) data set and the merged precipitation observations, known as Climate Prediction Center Morphing Technique and gauge observations (CMPA, Version 2.0; Shen et al., 2014) as the precipitation labels. Our primary focus lies in the comparative analysis of various DL models, with a specific emphasis on validating the superiority of a physics-informed DL model. Despite the known bias in China (e.g., Qin et al., 2021), we opt to employ the ERA-5 precipitation data due to its advantageous attributes of high temporal and spatial resolution, as well as extensive spatial coverage. To assess the model's generalization and its applicability to observational data, additional testing conducted using CMPA data. To embed the climatological impacts on the synoptic-scale transient eddies. Monthly Oceanic Niño Index (ONI) from the Climate Prediction Center (CPC) is also employed to represent the ENSO phase. The spatial domain selected for our analysis encompasses the region between 15°N and 54.75°N latitude and 70°E to 134.75°E longitude, which specifically targets precipitation data over the land area of China (atmospheric conditions above oceans are also included in the experiments, but not presented). The spatial resolution is interpolated to $0.25^{\circ} \times 0.25^{\circ}$, resulting in a 160×260 image for each atmospheric variable.

The experiments are implemented on the data during 2018/01–2022/05, containing 16,032 instances. For training and testing our ω -GNN model, the data set is divided into three parts: training data set (2018/01–2020/11) with 8364 instances; validation data set (2020/11–2021/07) with 3,648 instances; and testing data set (2021/07–2022/05) with 4,020 instances.

3. Methods

3.1. GNN Model

3.1.1. Graph Formulation

Graph is a data structure composed of nodes and edges which connect node pairs. In practice, each node can represent an entity, and edges reflect some kind of relationship between these entities. A graph can be concretized and pictured like Figure S1a in Supporting Information S1 and represented as an adjacency matrix in calculation. To capture the coupling between physical variables through GNNs, we need to formulate a graph based on the



structure of physics. Precipitation is influenced by two critical physical factors: ascending motion and relative humidity of the troposphere. Ascending motion can be described by ω -equation (Dai & Nie, 2020; Holton & Hakim, 2013):

$$\left(\partial_{\rm pp} + \frac{\sigma}{f^2} \nabla^2\right) \omega = -\frac{1}{f} \partial_p ({\rm Adv}_{\zeta}) - \frac{R}{pf^2} \nabla^2 {\rm Adv}_T.$$
(1)

This equation implies that ascending motion, horizontal wind, air temperature, and pressure influence and determine each other in little air mass (Hu et al., 2018; Nie & Sobel, 2016). Then ascending motion will directly make for precipitation. At the same time, only saturated vapor can condense into rainfall. We first select variables following omega equation and moist equation and capture relationships present at these equations by a graph. Thus, Figure S1b in Supporting Information S1 shows the structure of the physical process in precipitation. This graph contains six atmospheric elements (physical variables) represented as nodes, with edges reflecting a fuzzy kind of relationship distilled from physical laws. The use of this graph ensures the exchange of information between the elements at either end of an edge in the algorithm of GNNs.

3.1.2. Model Architecture

The ω -GNN model is an end-to-end 5-classification model, directly taking atmospheric elements as input and producing a predictive precipitation level as output. Figure 1a sketches the model architecture, which comprises three consecutive parts: an encoder, a ChebNet coupler, and a decoder (Grönquist et al., 2021; Ronneberger et al., 2015). For each atmospheric variable, the encoder has several Conv2D and MaxPooling layers that extract the spatial features at different levels and compress the 160 × 260 (Height, Width) fields into 8 × 13 × 128 (Height, Width, Channel) arrays (Figure 1b). These arrays are concatenated to form a 6 × (8 × 13 × 128) matrix, where each row vector corresponds to an atmospheric element. This matrix will be sent to the ChebNet coupler, which aggregates the elements following the graph shown in Figure S1b in Supporting Information S1 in spectral domain. It maps a graph (spatial domain) to spectral domain and filters the graph. The filter is a learnable parameter, which reflects the mode of node interaction. In the end, the decoder will take out the corresponding row vector of precipitation from the output matrix of the ChebNet coupler and raise its dimension to a 160 × 260 predictive field through the super-resolution technique of PixelShuffle (Figure 1d) (Shi et al., 2016). This can be described mathematically as a function:

$$P_{\text{pred}} = F(E_{\text{NWP}}; G(\varepsilon); \Theta)$$
(2)

where *F* is the ω -GNN model, ε represents the variables involved in precipitation process and the graph relationship between them, *G* represents graph neural network (GNN), and G(ε) reflects the non-linear aggregation of the variables through GNN according to their graph relationship. Θ is the set of trainable and non-trainable params, and *E*_{*NWP*} is the input of model *F*, which contains six components:

$$E_{\text{NWP}} = (P_{\text{NWP}}, T_{\text{NWP}}, u_{\text{NWP}}, v_{\text{NWP}}, h_{\text{NWP}}, \text{rh}_{\text{NWP}}).$$
(3)

precipitation, temperature, *u*-wind, *v*-wind, geopotential height, and relative humidity obtained from the ECMWF-IFS. For sophisticated decoding, skip-connections are employed to provide a hierarchy of detailed information from encoder to decoder (Han et al., 2021; He et al., 2015).

The ChebNet coupler is the critical part to realize the coupling of atmospheric elements. Essentially its algorithm is

$$Y = \sum_{k=0}^{K} T_k(\tilde{L}) X W_k \tag{4}$$

where $X, Y \in \mathbb{R}^{6\times(8\times13\times128)}$ and $W_k \in \mathbb{R}^{(8\times13\times128)\times(8\times13\times128)}$ in this study. T_k is the *k*th-order term of the Chebyshev polynomial, and $\tilde{L} \in \mathbb{R}^{6\times6}$ is the normalized Laplace matrix of the graph used in the ChebNet. Noticing that every 128 elements in each row of *X* and *Y* can be traced back to the same region (see Figure 1c), we make a hypothesis that the coupling of variables takes place locally:

$$Y = \sum_{k=0}^{K} T_k (\tilde{L}) (x^1 w_k^1 \| \dots \| x^{8 \times 13} w_k^{8 \times 13})$$
(5)

where $x^n \in \mathbb{R}^{6\times 128}$, $w_k^n \in \mathbb{R}^{128\times 128}$ and \parallel means concatenation. It is reasonable because there is few impact at a distance in physics. This hypothesis helps us reduce the size of the model by $O(10^2)$ and the computational complexity by $O(10^4)$ while retain a high model performance.





Figure 1. (a) A general diagram of the ω -GNN model. (b) The details of the encoder and the embedding for each atmospheric element; n is the dimension of the embedding vector for the climatological background. (c) The details of the locality hypothesis, take one element for example, (in reality there are 6). (d) The details of the decoder. Red represents the corresponding row vector of precipitation.

The ChebNet theory (and other spectral-based GNN theories) assumes an undirected graph (Chen et al., 2020; Wang & Zhang, 2022). However, the weight on each edge is difficult to determine from the complex relationship of ω -equation. Therefore, we assume that the weights on the edges are all equal to 1, and instead, we design a self-attention mechanism in our model, which multiplies each element by a weight before it is sent into the coupler (Vaswani et al., 2017). The weights are calculated as:

$$\alpha^{n} = \operatorname{softmax}\left((\operatorname{tanh}((x^{n} \odot A \cdot x^{n})w^{n} + b^{n}))w^{\prime n} + b^{\prime n}\right).$$
(6)

According to the coupler, the weight $a^n \in \mathbb{R}^6$ is local too, where *n* means the *n*th region. *A* is the adjacency matrix of the graph, so $A \cdot x^n$ is the sum of all the connected nodes toward a NoI. \odot is Hadamard product. This set of weights guarantees that the weight for each atmospheric element has considered all its relationships with other elements. The data put into the coupler are:

 x^n

$$= \alpha^n \odot x^n. \tag{7}$$



3.2. Climatological Background Embedding

In addition to considering the physical process of precipitation, incorporating the climatological background state as prior knowledge is crucial in DL models due to its impacts on synoptic-scale transient eddies (Barsugli et al., 1999; Lau et al., 2005; Wang et al., 2021, 2022). In this study, we place particular emphasis on seasonality and ENSO phases since the data set covers $4\sim5$ years (Diaz et al., 2001; Rasmusson et al., 1990). We employ entity embedding to prompt the model to distinguish climatological and other sparse factors including lead time and valid time (Figure 1b). The embedding model is referred to as the ω -EGNN model, to discriminate from the ω -GNN model without embedding.

3.3. Experiment Setting and Training

To assess the performance of the GNN model, we initially compare it with NWP (the ECMWF-IFS). Subsequently, to analyze the feature fusion effects of GNN, we compare it with other deep learning models, including a U-NET model and a 3D-CNN model (CNN-3D), which performs feature fusion through convolution across different variables. We employ a 5-depth U-Net, which is now popular in weather forecasts, and the CNN-3D model is the same as the ω -GNN model except for a Conv3D coupler. Additionally, to assess the effectiveness of embedding climatological information, we compared ω -GNN and ω -EGNN.

All the DL models are trained for 100 epochs, using dice loss, a widely used loss function for imbalanced data (Jadon, 2020; Sudre et al., 2017):

Dice =
$$\frac{2 \times TP + \varepsilon}{2 \times TP + FP + FN + \varepsilon}$$
(8)

where TP means True Positive, FP means False Positive, FN means False Negative and ε is a small number. Since dice loss is based on binary classification, we see the 5-classification results of our experiments as 4 binary classifications and calculate their average. We repeat training 10 times for each DL model as perturbation experiments. The optimization algorithm used for the model is Adam, with a batch size of 16, and each model a trained for 100 epochs.

Threat score (TS), false alarm rate (FAR), probability of detection (POD), and bias score (BS) are adopted to evaluate predictive performance:

$$TS = \frac{IP}{TP + FP + FN}$$

$$FAR = \frac{FP}{TP + FP}$$

$$POD = \frac{TP}{TP + FN}$$

$$BS = \frac{TP + FP}{TP + FN}$$
(9)

among which TS and POD are the higher the better, FAR is the low the better and BS is the closer to 1 the better.

4. Results

4.1. Model Performance

Table S1 and Figures 2a–2d present the scores of each model. From a TS perspective, all four DL models show significant improvement over the NWP. Among the DL models, the GNN family of models (ω -GNN and ω -EGNN) exhibits notable superiority over others. It is worth noting that at lower thresholds (0.1 mm/6h and 2.5 mm/6h), the advantage of the GNN family over other DL models is relatively small. However, as the threshold increases, particularly for heavy precipitation events, the advantage of the GNN family becomes more pronounced. For instance, at a threshold of 20 mm/6h, ω -GNN outperforms U-Net and CNN-3D by 7.64% and 13.53%, respectively, while ω -EGNN shows even higher improvements of 9.44% and 15.43% over these models. These results demonstrate the superior performance of the GNN family of models, especially in capturing heavy precipitation events.



Geophysical Research Letters



Figure 2. Comparison of NWP, U-Net, CNN-3D, ω -GNN and ω -EGNN in (a) TS, the higher the better. (b) FAR, the low the better. (c) POD, the higher the better. (d) BS, the closer to 1 (the dashed gray line) the better. The colored bars display the mean of multiple trainings while the error bars show the standard deviation. *X* axis unit: mm/6h. And the distribution of TS difference between (e) ω -GNN and NWP, (f) ω -GNN and U-Net, (g) ω -GNN and CNN-3D, (h) ω -EGNN and NWP, (i) ω -EGNN and U-Net and (j) ω -EGNN and CNN-3D on the threshold of 20 mm/6h. The blank areas derive from the denominator of zero when calculating the TS, due to all the samples on the grids being TN (Ture Negative, which means no precipitation at this level).

In addition to the TS metric, other quantities have been calculated for thorough evaluation. Similar patterns as that in TS are observed across these metrics. Specifically, at thresholds of 0.1 mm/6h and 2.5 mm/6h, ω -GNN and ω -EGNN show slightly better POD scores compared to U-Net and CNN-3D. At higher thresholds of 10 mm/6h and 20 mm/6h, ω -GNN exhibits improvements of 15.07% and 20.33%, respectively, over U-Net and CNN-3D, while ω -EGNN shows improvements of 13.19% and 18.36%. Similar results could be found in FAR and BS scores. These results highlight the stronger capabilities of the GNN family models, especially in accurately identifying heavy precipitation events. Furthermore, the results of ω -GNN and ω -EGNN have a smaller uncertainty in almost all circumstances, implying that these models are more robust compared to other DL models.

The above results naturally raise a question: Can the GNN family of models maintain their aforementioned advantages when facing unequal modeling conditions, such as scenarios where other models have access to additional inputs? To try to answer this question, another U-Net model (hereafter U-Net+), using more variables additional to before, including t2m, surface pressure, surface relative humidity, land-sea mask, lake cover, high and low vegetation cover, and altitude, is trained for comparison. Even though U-Net + absorbs more data than ω -GNN and ω -EGNN, the findings still hold (Table S1 and Figure S2 in Supporting Information S1). This suggests that even when compared to models using more input data, GNN models may still have certain advantages. The extent to which this advantage can be maintained remains unknown.

Figures 2e-2j present the distributions of TS difference with the threshold of 20 mm/6h between GNN family models and other models. For most areas in the northeast, east, and middle of China, TS gets improved. Conversely, a reduction in TS is noted in the southern regions of China when compared to alternative models. Additional depictions of TS differences at various thresholds are included in the in Supporting Information S1 as Figures S3–S5. When precipitation is small, the forecasting performance of GNN models closely aligns with that of other DL models, and they outperform NWP almost nationwide. However, as precipitation intensifies (e.g., 10 mm/6h), the regions where GNN models exhibit a relative improvement over NWP models closely resemble the results with a 20 mm/6h threshold. Figure S6 in Supporting Information S1 displays the TS scores at various thresholds with increasing lead times. Across all lead times and precipitation ranges, DL models consistently outperform the NWP, albeit with a diminishing advantage as the lead time increases. However, as the threshold surpasses 20 mm/6h, the performance gap between DL models and the NWP narrows, and most DL models' advantage becomes less significant around lead times of approximately 20 hr. In contrast, the GNN family of models continues to exhibit superior and robust TS scores, significantly outperforming the NWP model, corroborating the results presented in Table S1 in Supporting Information S1. It is essential to acknowledge an increase in TS scores observed from 6 to 12 hr, a phenomenon also documented in previous studies (e.g., Tjernström et al., 2021). This phenomenon may be linked to the inherent complexity of precipitation, and the underlying causes remain yet to be fully understood.

Case studies are conducted to verify the practical prediction performance. Figure 3 and Figure S7 in Supporting Information S1 show three cases of 12hr forecast of 2021/08/29/00:00 (Figures 3a), 18hr forecast of 2021/08/26/18:00 (Figure 3b), and 30hr forecast of 2021/07/30/06:00 (Figure S7 in Supporting Information S1), respectively. In the first case, the heavy rainfall (red area) has a shape of oblong ellipse from the northeast to the southwest in the middle of the box. The result of NWP has a false rainfall center to the southwest of the true heavy rainfall, while U-Net and CNN-3D predict a much smaller area of heavy rainfall than ground truth. These errors are corrected by ω -GNN and ω -EGNN, as they predict a heavy rainfall area with a similar shape and size to ground truth and no false heavy rainfall center. In the second case, there are two rainfall centers around (93.0°E, 27.5°N) and (94.5°E, 28.0°N). U-Net and CNN-3D only predict one rainfall center. NWP captures both the two centers, but the west center is much smaller than the ground truth. The results of the GNN family are close to NWP but improve the size of the west center. These two cases demonstrate the ability of GNN family models to identify and predict heavy rain, which corresponds to their high TS in Section 4.1. In the last case, there are three third-level rainfall centers in the middle, south, and east of the box, respectively. However, NWP, U-Net, and CNN-3D overestimate the intensity of these centers as fourth-level. The GNN family models, especially the ω -EGNN, can eliminate these false alarms, which reduces the FAR of them.

4.2. Effect of Climatological Background Embedding

Guo and Berkhahn (2016) claimed that entity-embedding vectors possess realistic meaning in terms of distance and direction. To investigate the relationship between synoptic-scale precipitation and climatological background,



Geophysical Research Letters



Figure 3. NWP and DL model predictions for (a) case 2021/08/22/00:00, 12hr-lead time, (b) 2021/08/26/18:00, 18hr-lead time. Shading represents the predicted precipitation levels. In the case of DL models, the shading corresponds to the highest probability precipitation level among the 10 ensembles. The numerical values indicate the count of forecasts predicting intense precipitation (level 4) among the 10 ensembles. Specifically, 'A' denotes that all 10 ensembles forecast intense precipitation.

we reduce the dimension of the learned embedding vectors of seasonality and ENSO phase to 2D using principal component analysis, as shown in Figure S11 in Supporting Information S1. These vectors have been appropriately normalized for analysis.

Figures S11a and S11b in Supporting Information S1 depict two typical structures of seasonal embedding obtained from multiple sensitivity experiments. The first structure demonstrates a clear separation of the four seasons into distinct quadrants, while the second structure shows closer proximity between two seasons, such as winter, spring, or autumn. Overall, all winter, spring, and autumn precipitation patterns remain distinct from those of summer. These findings support the notion of seasonal differences from a predictive perspective and are consistent with different prediction errors between summer and winter (Figures S10c and S10d in Supporting Information S1).

Moreover, the embedding vectors representing different ENSO phases exhibit a more distinctive pattern (Figures S11e and S11f in Supporting Information S1), consistently forming a shape approximating an equilateral triangle in multiple training instances. This suggests that various ENSO phases have a profound impact on synoptic-scale precipitation over China, consequently resulting in systematic variations in forecast errors across different ENSO phases (e.g., Figure S11g and S11h in Supporting Information S1).

4.3. Generalization Capability Across Observational Data Set

The preceding research has highlighted the advantages of GNN models when applied to reanalysis data sets. This raises an intriguing question: how does this model perform when applied to observational data? To address this question, we conducted a tesing and analysis using CMPA data set. Due to variations in data sources between



Geophysical Research Letters



Figure 4. (a) Comparison of NWP, U-Net, CNN-3D, ω -GNN and ω -EGNN in TS, the higher the better. The colored bars display the mean of multiple trainings while the error bars show the standard deviation. *X* axis unit: mm/6h. And the distribution of TS difference between (b) ω -GNN and NWP, (c) ω -GNN and U-Net, (d) ω -GNN and CNN-3D, (e) ω -EGNN and NWP, (f) ω -EGNN and U-Net and (g) ω -EGNN and CNN-3D on the threshold of 20 mm/6h. The blank areas derive from the denominator of zero when calculating the TS, due to all the samples on the grids being TN (True Negative, which means no precipitation at this level).

domestic and international regions within the CMPA data set, our modeling and validation efforts are primarily focused on inland regions within 30°E–40°E, 105°N-115°N. The temporal domain and division of training, validation and test data sets are the same as Section 2.1.

As illustrated in Figure 4a, models within the GNN family consistently demonstrate significant superiority across all precipitation categories when compared to NWP. Of particular note is the performance of the GNN model, which performs comparably to U-NET for precipitation levels below 20 mm while significantly outperforming the 3D-CNN model. Most notably, for precipitation levels exceeding 20 mm, the GNN family of models exhibits a substantial advantage over other models. This underscores the exceptional performance of the designed GNN model when applied to observational data, particularly for intense precipitation events exceeding 20 mm. The difference between TS distributions of two models on 20 mm/6h have also been presented in Figures 4b–4g. In comparison to other models, the GNN models exhibit a noticeable enhancement in most regions. When compared to different models, there are some variations in the distribution of TS differences. This suggests that different models may capture distinct features, leading to differences in TS distribution.

5. Conclusion

In this study, we propose GNN-based ω -GNN and ω -EGNN models to forecast the precipitation of China. The ω -GNN model carries a constraint of atmospheric coupling inspired by omega and water vapor equations, and the ω -EGNN model additionally contains sparse factors like climatological background based on the ω -GNN model. To enhance computational efficiency, we transition from global to localized calculations in the GNN considering the characteristics of precipitation. This adjustment significantly reduces computational complexity while maintaining comparable performance.

These two models show significant improvements in the forecast skill of heavy rainfall compared to NWP and other non-physics-informed DL models under identical circumstances. Specifically, their prediction attains pronounced higher TS than other models, especially for thresholds greater than 10 mm/6h, and this advantage is maintained at every lead time within 72 hr, which means a better identification of heavy rainfall. The false alarm rate of the GNN family models is also lower than other models, indicating that they can reduce overestimating rainfall intensity. Furthermore, via multiple perturbation experiments, it becomes evident that the predictions generated by the GNN family models demonstrate reduced uncertainty when juxtaposed with those from other deep learning models. This implies that the GNN family models are more reliable and easier to train. These results imply that the two physics-informed GNN models can mitigate the negative effects of imbalanced data,

19448007, 2024, 2, Downloaded from https

agupubs.onii

doi/10.1029/2023GL106676 by CochraneChina,

Wiley Online Library on [18/01/2024]. See the Terms

and Conditions (https://onli

such as precipitation, on traditional DL models. Additionally, by introducing entity embedding, ω -EGNN excels in short-term predictions, while ω -GNN demonstrates greater proficiency in long-term forecasting.

The interpretability of the ω -GNN and ω -EGNN models are also analyzed. The embedding vectors of season display that all winter, spring, and autumn precipitation patterns remain distinct from those of summer, unveiling the seasonal difference of prediction. The embedding vectors also exhibit distinct variations in precipitation between different ENSO phases. These findings are consistent with the different forecast errors and highlight the importance of considering climatology background in synoptic prediction and correction. The physics-informed GNN models can enhance not only the performance but also the interpretability.

Our analysis is conducted using both ERA5 data and CMPA data as ground truth. ERA5 data exhibits better physical consistency with ECMWF forecasts, boasts higher spatial and temporal resolutions, and is more readily accessible. These characteristics simplify the training of DL models and enable direct model comparisons. CMPA data provides an excellent representation of hourly precipitation in China, enabling us to evaluate forecasting capability for observed precipitation. Our GNN-family models demonstrate superior performance on both data sets, highlighting their excellent generalization capabilities across different data sets and their suitability for observed precipitation. The construction of graphs based on the omega equation, the localized modification of ChebNet, and the utilization of embeddings, including their integration, dimensionality reduction, and comparison with climate diagnostics, can provide valuable insights for related research. Subsequent research endeavors should involve the utilization of observational data, such as station data, and more related variables, such as topography to conduct further evaluations of these models and to delve into potential enhancements in precipitation forecasting.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The data that supports the findings of this study is openly available at Hersbach et al. (2023).

References

- Barsugli, J. J., Whitaker, J. S., Loughe, A. F., Sardeshmukh, P. D., & Toth, Z. (1999). The effect of the 1997/98 EI nino on individual large-scale weather events. Bulletin of the American Meteorological Society.
- Bauer, P., Thorpe, A., & Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature*, 525(7567), 47–55. https://doi.org/10.1038/nature14956
- Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., & Tian, Q. (2022). Pangu-weather: A 3D high-resolution model for fast and accurate global weather forecast. https://doi.org/10.48550/arXiv.2211.02556

Chen, Z., Chen, F., Zhang, L., Ji, T., Fu, K., Zhao, L., et al. (2020). Bridging the gap between spatial and spectral domains: A survey on graph neural networks (Version 4). arXiv. https://doi.org/10.48550/ARXIV.2002.11867

Clark, P., Roberts, N., Lean, H., Ballard, S. P., & Charlton-Perez, C. (2016). Convection-permitting models: A step-change in rainfall forecasting. *Meteorological Applications*, 23(2), 165–181. https://doi.org/10.1002/met.1538

Dai, P., & Nie, J. (2020). A global quasigeostrophic diagnosis of extratropical extreme precipitation. Journal of Climate, 33(22), 9629–9642. https://doi.org/10.1175/JCLI-D-20-0146.1

- Diaz, H. F., Hoerling, M. P., & Eischeid, J. K. (2001). ENSO variability, teleconnections and climate change. International Journal of Climatology, 21(15), 1845–1862. https://doi.org/10.1002/joc.631
- Ding, Y., Liang, P., Liu, Y., & Zhang, Y. (2020). Multiscale variability of Meiyu and its prediction: A new review. Journal of Geophysical Research: Atmospheres, 125(7), e2019JD031496. https://doi.org/10.1029/2019jd031496
- Grönquist, P., Yao, C., Ben-Nun, T., Dryden, N., Dueben, P., Li, S., & Hoefler, T. (2021). Deep learning for post-processing ensemble weather forecasts. *Philosophical Transactions of the Royal Society A*, 379(2194), 20200092. https://doi.org/10.1098/rsta.2020.0092
 Guo, C., & Berkhahn, F. (2016). Entity embeddings of categorical variables.

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

Han, L., Chen, M., Chen, K., Chen, H., Zhang, Y., Lu, B., et al. (2021). A deep learning method for bias correction of ECMWF 24–240 h forecasts. Advances in Atmospheric Sciences, 38(9), 1444–1459. https://doi.org/10.1007/s00376-021-0215-y

- Han, Y., Zhang, G. J., Huang, X., & Wang, Y. (2020). A moist physics parameterization based on deep learning. Journal of Advances in Modeling Earth Systems, 12(9). https://doi.org/10.1029/2020MS002076
- He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition. Citing the data.
- Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., et al. (2023). ERA5 hourly data on single levels from 1940 to present [Dataset]. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.adbb2d47

Hess, P., Drüke, M., Petri, S., Strnad, F. M., & Boers, N. (2022). Physically constrained generative adversarial networks for improving precipitation fields from Earth system models. *Nature Machine Intelligence*, 4(10), 828–839. https://doi.org/10.1038/s42256-022-00540-1

Acknowledgments

The study is jointly supported by National Natural Science Foundation of China (42261144687 and 42141019) and the Second Tibetan Plateau Scientific Expedition and Research (STEP) program (Grant 2019QZKK0102). The authors would like to thank the National Key Scientific and Technological Infrastructure project "Earth System Science Numerical Simulator Facility" (EarthLab) and Beijing Paratera Technology Co., Ltd. for computational support. The authors would like to thank the Center for Artificial Intelligence in Atmospheric Science for all the support provided.

rned by the applicable Creative Commons

19448007, 2024, 2, Downloaded

from https

agupub

wiley.com/doi/10.1029/2023GL106676 by

Wiley Online Library on [18/01/2024]

See

Online Library

for rules

; OA

rned by the

applicable Creati

Holton, J. R., & Hakim, G. J. (2013). An introduction to dynamic meteorology (5th ed., p. 532). Academic Press.

- Hu, K., Huang, G., Wu, R., & Wang, L. (2018). Structure and dynamics of a wave train along the wintertime Asian jet and its impact on East Asian climate. Climate Dynamics, 51(11-12), 4123-4137. https://doi.org/10.1007/s00382-017-3674-1
- Jadon, S. (2020). A survey of loss functions for semantic segmentation. In 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), Via del Mar. Chile (Vol. 1-7), IEEE,
- Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. Nat Rev Phys, 3(6), 422-440. https://doi.org/10.1038/s42254-021-00314-5
- Kashinath, K., Mustafa, M., Albert, A., Wu, J. L., Jiang, C., Esmaeilzadeh, S., et al. (2021). Physics-informed machine learning: Case studies for weather and climate modelling. Philosophical Transactions of the Royal Society A, 379(2194), 20200093. https://doi.org/10.1098/ rsta.2020.0093

Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks.

- Konapala, G., Mishra, A. K., Wada, Y., & Mann, M. E. (2020). Climate change will affect global water availability through compounding changes in seasonal precipitation and evaporation. Nature Communications, 11(1), 3044. https://doi.org/10.1038/s41467-020-16757-
- Lau, N. C., Leetmaa, A., Nath, M. J., & Wang, H. L. (2005). Influences of ENSO-induced Indo-Western Pacific SST anomalies on extratropical atmospheric variability during the boreal summer. Journal of Climate, 18(15), 2922-2942. https://doi.org/10.1175/jcli3445.1
- Lesk, C., Rowhani, P., & Ramankutty, N. (2016). Influence of extreme weather disasters on global crop production. Nature, 529(7584), 84-87. https://doi.org/10.1038/nature16467
- Li, P., Yu, Y., Huang, D., Wang, Z., & Sharma, A. (2023). Regional heatwave prediction using graph neural network and weather station data. Geophysical Research Letters, 50(7), e2023GL103405. https://doi.org/10.1029/2023GL103405
- Liu, Q., Lou, X., Yan, Z., Qi, Y., Jin, Y., Yu, S., et al. (2023). Deep-learning post-processing of short-term station precipitation based on NWP forecasts. Atmospheric Research, 295, 107032. https://doi.org/10.1016/j.atmosres.2023.107032
- Mu, B., Qin, B., & Yuan, S. (2021). ENSO-ASC 1.0.0: ENSO deep learning forecast model with a multivariate air-sea coupler. Geoscientific Model Development, 14(11), 6977-6999. https://doi.org/10.5194/gmd-14-6977-2021
- Nie, J., & Sobel, A. H. (2016). Modeling the interaction between quasigeostrophic vertical motion and convection in a single column. Journal of the Atmospheric Sciences, 73(3), 1101-1117. https://doi.org/10.1175/JAS-D-15-0205.1
- O'Gorman, P. A. (2015), Precipitation extremes under climate change, Current Climate Change Reports, 1(2), 49-59, https://doi.org/10.1007/ s40641-015-0009-3
- Pathak, J., Subramanian, S., Harrington, P., Raja, S., Chattopadhyay, A., Mardani, M., et al. (2022). FourCastNet: A global data-driven high-resolution weather model using adaptive fourier neural operators. https://doi.org/10.48550/arXiv.2202.11214
- Oin, S., Wang, K., Wu, G., & Ma, Z. (2021). Variability of hourly precipitation during the warm season over eastern China using gauge observations and era5. Atmospheric Research, 264, 105872. https://doi.org/10.1016/j.atmosres.2021.105872
- Rasmusson, E. M., Wang, X., & Ropelewski, C. F. (1990). The biennial component of ENSO variability. Journal of Marine Systems, 1(1-2), 71-96. https://doi.org/10.1016/0924-7963(90)90153-2
- Ravuri, S., Lenc, K., Willson, M., Kangin, D., Lam, R., Mirowski, P., et al. (2021). Skilful precipitation nowcasting using deep generative models of radar. Nature, 597(7878), 672-677. https://doi.org/10.1038/s41586-021-03854-z
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat (2019). Deep learning and process understanding for data-driven Earth system science. Nature, 566(7743), 195-204. https://doi.org/10.1038/s41586-019-0912-1
- Ronneberger, O., Fischer, P., & Brox, T. (2015), U-Net: Convolutional networks for biomedical image segmentation. Medical image computing and computer-assisted Intervention-MICCAI 2015. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), Lecture notes in computer science (pp. 234–241). Springer International Publishing.
- Scarselli, F., Gori, M., Chung Tsoi, A., Hagenbuchner, M., & Monfardini, G. (2009). The graph neural network model. IEEE Transactions on Neural Networks, 20(1), 61-80. https://doi.org/10.1109/TNN.2008.2005605
- Shen, Y., Zhao, P., Pan, Y., & Yu, J. (2014). A high spatiotemporal gauge-satellite merged precipitation analysis over China. Journal of Geophysical Research: Atmospheres, 119(6), 3063-3075. https://doi.org/10.1002/2013jd020686
- Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, A. P., Bishop, R., et al. (2016). Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network.
- Sudre, C. H., Li, W., Vercauteren, T., Ourselin, S., & Cardoso, M. J. (2017). Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In M. J. Cardoso, et al. (Eds.), Deep learning in medical image analysis and multimodal learning for clinical decision support (Vol. 10553, pp. 240-248. of Lecture Notes in Computer Science, Springer International Publishing.
- Tiernström, M., Svensson, G., Magnusson, L., Brooks, I. M., Prytherch, J., Vüllers, J., & Young, G. (2021), Central arctic weather forecasting: Confronting the ECMWF IFS with observations from the Arctic Ocean 2018 expedition. Quarterly Journal of the Royal Meteorological Societv, 147(735), 1278-1299, https://doi.org/10.1002/gi.3971
- Trenberth, K. E., Dai, A., Rasmussen, R. M., & Parsons, D. B. (2003). The changing character of precipitation. Bulletin of the American Meteorological Society, 84(9), 1205-1218. https://doi.org/10.1175/BAMS-84-9-1205
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention is all you need.
- Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., & Bengio, Y. (2018). Graph attention networks.
- Wang, X., & Zhang, M. (2022). How powerful are spectral graph neural networks.
- Wang, Y., Hu, K., Huang, G., & Tao, W. (2021). Asymmetric impacts of El Nino and La Nina on the Pacific-North American teleconnection pattern: The role of subtropical jet stream, Environmental Research Letters, 16(11), 114040, https://doi.org/10.1088/1748-9326/ac31ed
- Wang, Y., Huang, G., Hu, K., Tao, W., Li, X., Gong, H., et al. (2022). Asymmetric impacts of El Niño and La Niña on the Pacific-south America teleconnection pattern. Journal of Climate, 35(6), 1825-1838. https://doi.org/10.1175/jcli-d-21-0285.1
- Westra, S., Fowler, H. J., Evans, J. P., Alexander, L. V., Berg, P., Johnson, F., et al. (2014). Future changes to the intensity and frequency of short-duration extreme rainfall: Future intensity of sub-daily rainfall. Reviews of Geophysics, 52(3), 522-555. https://doi. org/10.1002/2014RG000464
- Wilks, D. S. (2009). Extending logistic regression to provide full-probability-distribution MOS forecasts. Meteorological Applications, 16(3), 361-368, https://doi.org/10.1002/met.134
- Wu, Y., Tang, Y., Yang, X., Zhang, W., & Zhang, G. (2021). Graph convolutional regression networks for quantitative precipitation estimation. IEEE Geoscience and Remote Sensing Letters, 18(7), 1124–1128. https://doi.org/10.1109/LGRS.2020.2994087



Wu, Y., Yang, X., Tang, Y., Zhang, C., Zhang, G., & Zhang, W. (2022). Inductive spatiotemporal graph convolutional networks for short-term quantitative precipitation forecasting. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–18. https://doi.org/10.1109/tgrs.2022.3159530
 Zhang, Y., Long, M., Chen, K., Xing, L., Jin, R., Jordan, M. I., & Wang, J. (2023). Skilful nowcasting of extreme precipitation with NowcastNet. *Nature*, 619(7970), 526–532. https://doi.org/10.1038/s41586-023-06184-4