Machine Learning Parameterization of the Multi-scale Kain-Fritsch (MSKF) Convection Scheme and stable simulation coupled in WRF using WRF-ML v1.0

Xiaohui Zhong[1,*], Xing Yu[2], and Hao Li[1,*]

[1]Fudan University, Shanghai, 200433, China
[2]Shenzhen Institute of Artificial Intellegence and Robotics for Society, Guangdong, 518000, China
[*]These authors contributed equally to this work.

Correspondence: Xing Yu (yuxing@cuhk.edu.cn)

Abstract. Warm-sector heavy rainfall along the South China coast poses significant forecasting challenges due to its localized nature and prolonged duration. To improve the prediction of such high-impact weather events, high-resolution numerical weather prediction (NWP) models are increasingly used to more accurately represent topographic effects. However, as these models' grid spacing approaches the scale of convective processes, they enter a "gray zone" where the models struggle to fully resolve the turbulent eddies within the atmospheric boundary layer, necessitating partial parameterization. The appropriateness of applying convection parameterization (CP) schemes within this gray zone remains controversial. To address this, scale-aware CP schemes have been developed to improve the representation of convective transport. Among these, the multi-scale Kain-Fritsch (MSKF) scheme enhances the traditional Kain-Fritsch (KF) scheme, incorporating modifications that facilitate its effective application at spatial resolutions as high as 2 km. In recent years, there has been an increasing application of machine learning (ML) models across various domains of atmospheric sciences, including efforts to replace conventional physical parameterizations with ML models. This work introduces a multi-output bidirectional long short-term memory (Bi-LSTM) model intended to replace the scale-aware MSKF CP scheme. Data for training and testing the model are generated using the Weather Research and Forecast (WRF) model over South China at a horizontal resolution of 5 km. Furthermore, this work evaluates the performance of the WRF model coupled with the ML-based CP scheme against simulations with traditional MSKF scheme. The results demonstrate that the Bi-LSTM model can achieve high accuracy, indicating the promising potential of ML models to substitute the MSKF scheme in the gray zone.

1 Introduction

Warm-sector heavy rainfall often occurs in South China during the pre-flood season, primarily influenced by the East Asian summer monsoon (Ding, 2004). These rainfall events are characterized by intense and localized precipitation over limited area. Despite their small scale, such unexpected and extreme warm-sector rainfall can cause significant damage, including flooding homes and vehicles, destroying crop fields, and endangering lives, leading to economic losses ranging from millions to even billions of dollars (Tao, 1981; Zhao et al., 2007; Zhong et al., 2015). Accurately predicting warm-sector heavy rainfall

with Numerical Weather Prediction (NWP) models is challenging due to the complex interaction of various factors, such as the low-level jet (LLJ), land-sea contrast, topography, and urban landscape (Zhong and Chen, 2017; Luo et al., 2017; Jian et al., 2002; Di et al., 2006; Xia and Zhao, 2009; Zhang and Ni, 2009). The complex terrain and heterogeneous land surface of South China region are crucial in promoting active convection. Previous studies (Giorgi et al., 2016; Mishra et al., 2018; Schumacher et al., 2020; Onishi et al., 2023) have demonstrated that higher spatial resolution improves the performance of convective rainfall forecasts by more accurately resolving topographic features. Acknowledging the importance of resolution in forecasting severe convective weather, both the Chinese government and the community have increasingly supported the development of high-resolution operational forecast models specifically designed for warm-sector rainstorms and sudden local rainstorms. In early 2017, the China Meteorological Administration (CMA) launched an initiative to develop a comprehensive framework for evaluating the forecast performance of all available models, including high-resolution regional models, and advancing key technologies for forecasting high-impact weather.

The increased computational resources has facilitated a shift towards the implementation of regional NWP models with increasingly finer grid spacings, typically within the range of 1 to 10 km. However, when the model grid spacing approaches the scale of convection, entering the so-called "gray zone" (Wyngaard, 2004; Hong and Dudhia, 2012), cumulus convection transitions from being completely unresolved to partially resolved. Theoretically, the accurate representation of the smallest turbulent scales, achievable only through Direct Numerical Simulation (DNS) at resolutions from millimeters to centimeters (Jeworrek et al., 2019), still requires the use of parameterization of turbulence or convection in weather modeling. There is ongoing debate regarding the efficacy of employing convection parameterization (CP) within the gray zone. Several studies (Chan et al., 2013; Johnson et al., 2013) have found that reducing horizontal grid spacing to below 4 km while using CP scheme, does not enhance and may even degrade, precipitation forecast performance. In contrast, other studies (Lean et al., 2008; Roberts and Lean, 2008; Clark et al., 2012) showed that forecasts with a horizontal grid spacing of 1 km, where convection is explicitly resolved, yielded more accurate spatial representation of accumulated rainfall over 48 hours compared to forecasts using 12 km and 4 km grid spacings. This discrepancy in research findings, with some indicating no benefit from finer grid spacing and others suggesting improved forecast accuracy, seems to stem from the application of the CP at scales beyond its originally intended operational range. Therefore, it remains unclear if utilizing any CP schemes in the gray zone is effective for predicting localized warm-sector heavy rainfall.

To enhance prediction accuracy in the gray zone, researchers have developed scale-aware CP schemes. These schemes dynamically parameterize convective processes based on the horizontal grid spacing, thus facilitating seamless transitions between different spatial scales. A pivotal study by Jeworrek et al. (2019) demonstrated that two specific scale-aware CP schemes, Grell-Freitas (Grell and Freitas, 2014) and multi-scale Kain-Fritsch (MSKF) (Zheng et al., 2016), surpassed conventional CP schemes in predicting both the timing and intensity of precipitation over the Southern Great Plains of the United States. Additionally, Ou et al. (2020) showed that the MSKF scheme outperformed other CP schemes, including Grell-3D Ensemble (Grell and Dévényi, 2002) and New Simplified Arakawa-Schubert (Han and Pan, 2011), in precipitation simulation. This was evidenced by its lower root mean squared error (RMSE) values when compared against in-situ observations and satellite data. Despite the increasing adoption of these scale-aware schemes due to their superior performance, it is crucial to acknowledge

that their efficacy also rely on various empirical parameters (Villalba-Pradas and Tapiador, 2022). Therefore, developing specialized CP schemes for the gray zone in NWP models continues to be a significant challenge.

In recent years, an increasing number of studies have investigated the use of machine learning (ML) models as alternatives to conventional physics-based CP schemes. These ML-based schemes have demonstrated potential for efficacy across various horizontal resolutions, benefiting from being trained on data from simulations that operate at varying grid resolutions (Yuval and O'Gorman, 2020). Unlike conventional CP schemes, which often rely on assumptions such as convective quasi-equilibrium (Arakawa, 2004), ML-based parameterization schemes do not require such assumptions. Notably, random forests (RFs) and fully-connected (FC) neural networks (NNs) have become predominant ML models for CP schemes in previous studies. RFs offer the advantages of inherently enforcing physical constraints, such as energy conservation and non-negative surface precipitation, essential for maintaining stable simulations. O'Gorman and Dwyer (2018) demonstrated RFs' capability to emulate moist convection in an aquaplanet general circulation model (GCM), maintaining stability and effectively reproducing key climate statistics. Furthermore, Yuval and O'Gorman (2020) employed the coarse-grained output from a high-resolution three-dimensional (3D) GCM model, simulated on an idealized equatorial beta plane, to train the RF parameterization. They showed that the RF parameterization is capable of reproducing the climate of the high-resolution simulation at coarser resolutions. However, FC NNs offer several advantages over RFs, such as the potential for higher accuracy and lower memory requirements. Krasnopolsky et al. (2013) introduced a stochastic CP scheme using an ensemble of 3-layer NNs, trained with data generated by a cloud-resolving model (CRM) during the TOGA-COARE [1] experiment, demonstrating its capacity for generating reasonable decadal climate simulations across a broader tropical Pacific region when incorporated into the National Center of Atmospheric Research (NCAR) Community Atmospheric Model (CAM). Similarly, Gentine et al. (2018) leveraged deep NN (DNN) trained on data from idealized and aquaplanet simulations performed using the SuperParameterized Community Atmosphere Model (SPCAM). The DNN predicts temperature and moisture tendencies due to convection and clouds, as well as the cloud liquid and ice water contents. Additionally, Rasp et al. (2018) successfully implemented an NN-based parameterization in a global GCM on an aquaplanet, conducting stable prognostic simulations over multiple years that accurately reproduced the climatology of SPCAM and capturing crucial aspects of variability, including extreme precipitation and realistic tropical waves. However, Rasp (2020) also found that minor changes to the configuration rapidly led to simulation instabilities, underscoring the need to address the robustness of NN parameterizations in GCMs. Yuval et al. (2021) developed a FC NN to that predicts subgrid fluxes instead of tendencies, incorporating physical constraints from coarse-grained high-resolution atmospheric simulation in an idealized domain. Brenowitz and Bretherton (2018, 2019) proposed a novel loss function designed to minimize accumulated prediction error over multiple time steps to enhance long-term stability and accuracy, by excluding upper atmospheric humidity and temperature from the input. Nonetheless, the approach of removing certain variables from the input is relatively rudimentary, demanding additional research to enhance the stability of NN-based parameterizations when integrated into the model.

[1] TOGA-COARE is an acronym for Tropical Ocean Global Atmospheres/Coupled Ocean Atmosphere Response Experiment. It is an international research program that investigates the interaction or coupling of the ocean and atmosphere in the western Pacific warm pool region from November 1992 to February 1993, encompassing 120 days of field experiments involving the deployment of oceanographic ships, moorings, drifters, and Doppler radars (ship, land, air).

Previous studies have predominantly used FC NNs to emulate convection, while more advanced NN structures have the potential to achieve higher accuracy. In a pioneering study, Han et al. (2020) explored the use of a deep residual convolutional NN (ResNet) (He et al., 2016) for the emulation of convection and cloud parameterization in the SPCAM model using a realistic configuration. They compared the performance of ResNet with various NN architectures, including a FC DNN, a DNN with skip connections, and a convolutional NN (CNN) without skip connections. The results revealed that ResNet and CNNs without skip connections outperformed FC NNs and DNNs with skip connections in accuracy, with ResNet and CNNs without skip connections showing comparable performance. This finding highlights the significant role of convolutions in enhancing accuracy. Furthermore, Yao et al. (2023) evaluated multiple ML model structures for simulating atmospheric radiative transfer processes, encompassing FC NNs, CNNs, bidirectional recurrent-based NNs (RNNs), transformer-based NNs (Vaswani et al., 2017), and Fourier Neural Operators (FNO (Li et al., 2020)). Their results indicated that models capable of preceiving global context of the entire atmospheric column significantly outperformed FC NNs and CNNs. Particularly, the bidirectional long short-term memory (Bi-LSTM) achieved the highest levels of accuracy. Similar to radiative transfer modeling, Han et al. (2020) also emphasized the importance of ML having a global perspective of the entire atmospheric column for ML models in convection modeling. They demonstrated that increasing the depths of CNNs from 4 to 22 layers significantly improved model accuracy, a benefit primarily attributed to the expansion of the receptive field in deeper CNN layers. Therefore, ML models that integrate both global and local perception capabilities are better suited for developing ML-based CP schemes.

Previous research have mostly focused on replacing CP schemes in GCM models with ML models for climate forecasting. The complexity of CP schemes in weather forecasting models surpasses that in GCMs (Arakawa, 2004). Generally, CP schemes in GCMs, whether in explicit or implicit form, assume that both the horizontal grid size and the temporal intervals for physics implementation are significantly larger and longer compared to the grid size and duration of individual moist-convective elements. In contrast, CP schemes in high-resolution models must account for dependencies on both the model's resolution and the time interval for implementing the physics (Arakawa, 2004). The ultimate goal is to develop ML models, based on data from super-parameterization or cloud-resolving models, to replace conventional CP schemes in weather forecasting models. This replacement seeks to reduce uncertainties and improve the efficacy of ML parameterizations. This study represents an initial effort to employ a ML model as an alternative to conventional CP schemes in weather forecasting models. For our dataset, we used the Weather and Research Forecasting (WRF) (Skamarock et al., 2021) model that covers the South China region, incorporating the scale-aware MSKF scheme employed as the CP scheme. The MSKF scheme, an improved version of the Kain-Fritsch (KF) scheme (Kain and Fritsch, 1990, 1993; Kain, 2004), aims to mitigate the overestimation of precipitation, address the premature convection trigger issue, particularly evident in overestimating precipitation during summer. To address these issues, the MSKF incorporates a scale-dependent capability, such as modifying the formulation of the convective adjustment timescale. This vital parameter, which determines the intensity and duration of convection, has been made dynamic and dependent on grid resolution (Zhang et al., 2021). Furthermore, we utilize a Bi-LSTM model to emulate the convective processes and couple it with the WRF model through the WRF-ML coupler developed by Zhong et al. (2023a). The performance of the ML-based CP scheme is evaluated in both offline and online settings.

The paper is structured as follows. Section 2 provides a description of the WRF model for data generation, as well as the input and output data of the ML model. In Section 3, the original and the ML-based MSKF scheme is introduced. The results for both offline and online testing of the ML-based MSKF scheme are presented in Section 4. Finally, Section 5 presents the summary and conclusion.

2 Data

2.1 Data generation

The dataset was generated by running the WRF model version 4.3 (Skamarock et al., 2019, 2021). The following subsections provide a comprehensive explanation of the WRF model configurations, as well as the input and output variables employed in the development of the ML-based CP scheme.

The WRF model is compiled using the GNU Fortran (gfortran version 7.5.0) compiler with the "dmpar" option. The WRF model is run using the domain configuration illustrated in Figure 1. The WRF model is configured with a single domain consisting of 44000 grid points, with a horizontal grid spacing of 5 km and dimensions of 220 $\times$ 200 grid points in the west-east and north-south directions. The model consists of 45 vertical levels (i.e., 44 vertical layers), with a model top at 50 hPa. Additionally, the WRF model is configured with physics schemes, including WSM 6-class graupel scheme (Hong and Lim, 2006) for microphysics, Bougeault-Lacarrère (BouLac) scheme (Bougeault and Lacarrère, 1989) for planetary boundary layer (PBL) mixing, the Monin-Obukhov (Janjic) surface layer scheme (Janjic, 1996), the Unified Noah model (Livneh et al., 2011) for land surface, RRTMG for both shortwave and longwave radiation (Iacono et al., 2008), and MSKF (Zheng et al., 2016) for cumulus. The time step used for all WRF simulations is set to 15 seconds.

The initial and boundary conditions for this work were derived from the ERA5 reanalysis dataset, which was provided by the European Centre for Medium-range Weather Forecast (ECMWF) (Hersbach et al., 2020). The ERA5 reanalysis dataset used in this study has a horizontal resolution of 0.25° and consists of 29 pressure levels below 50 hPa. To create a dataset for developing the ML model, the WRF simulations were initialized at 12 UTC and conducted 9 times every 2 days, specifically from May 20th, 2022 to June 5th, 2022. Throughout the simulations, the MSKF scheme was called every 5 model minutes, generating outputs at each call. The simulations ran for 36 hours each time, with the first 24 hours used for training and the last 12 hours for validation. Therefore, the total number of training samples is 114,444,000 [2] while the offline validation set contains 57,024,000 [3] samples.

Furthermore, given the possible discrepancy between offline performance, we conducted experiments that coupled the ML-based MSKF scheme with WRF model. This coupling aims at evaluating the online efficacy of the ML-based MSKF scheme by comparing it with the original WRF simulations. These simulations were performed 4 times every 2 days, with each simulation extending over a period of 168 hours (7 days). The initialization days spanned from June 12th, 2022 to June 18th, 2022.

[2] 114,444,000 = 44000 $\times$ 9 $\times$ (24 $\times$ 60 / 5 + 1)

[3] 57,024,000 = 44000 $\times$ 9 $\times$ 12 $\times$ 60 / 5

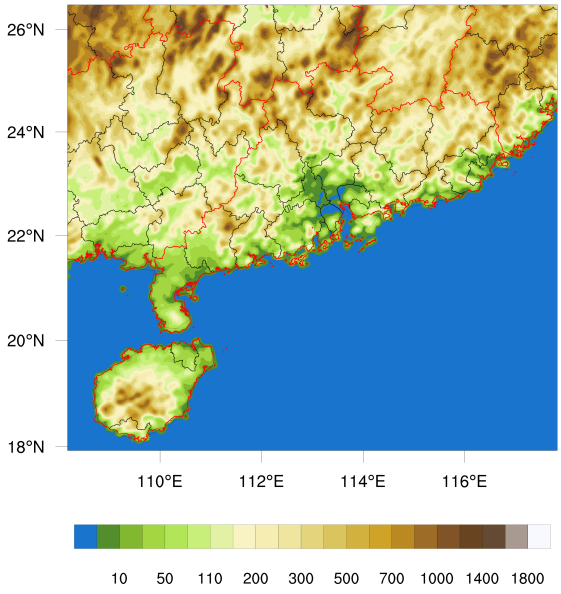

Figure 1. Digital evaluation data of the single WRF domain with horizontal resolution at 5°. Red lines are the province borderlines, and black lines are the city borderlines.

2.2 Input and output data

Table 1 presents a comprehensive list of the input and output variables used in this study, consistent with those utilized in the original MSKF scheme. There are 17 variables exclusively used as input, while 9 variables serve as both input and output. Specifically, the output variable "raincv", representing the time-step precipitation due to convection, is calculated through multiplying the precipitation rate by the model's time step. Among all the variables, 5 are two-dimensional (2D) surface variables, while the remaining ones are 3D variables characterized by 44-layer vertical profiles. Moreover, the ML model used in this study incorporates 4 derived variables as input. These variables consist of a 2D Boolean variable indicating convection triggering based on "nca" values, the pressure difference across adjacent vertical levels, the saturated water vapor mixing ratio, and relative humidity. Furthermore, the output "w0avg", which depends on vertical wind component (w) and input "w0avg", is also included as an input to model. In total, the ML model utilizes 27 input variables.

The variable "nca" represents the cloud relaxation time and must be an integer multiple of the model time step. For all WRF model simulations conducted in this study, a fixed time step of 15 seconds is used. Thus, "nca" is expected to be exactly divisible by 15. To eliminate dependence on the specific model time step, "nca" is divided by the model time step before normalization is applied during model training. Moreover, within the MSKF scheme, "nca" plays a crucial role in determining the triggering of convection. Convection is triggered when "nca" is equal to or exceeds half of the model time step.

Table 1. Definition of all the input and output variables, and whether they are surface or 3D variables and their corresponding units. There are 44 model layers.

Type	Variable name	Definition	type	Unit
Input	u	meridional wind component	3D	m/s
	v	zonal wind component	3D	m/s
	w	vertical wind component	3D	m/s
	t	temperature	3D	K
	qv	water vapor mixing ratio	3D	kg/kg
	p	pressure	3D	Pa
	th	potential temperature	3D	K
	dz8w	layer thickness	3D	m
	rho	air density	3D	kg/m^3
	pi	Exner function, which is dimensionless pressure and can be defined as: $(\frac{p}{p_0})^{R_d/c_p}$		
	hfx	upward heat flux at surface	surface	W/m^2
	ust	u$*$ in similarity theory	surface	W/m^2
	pblh	planetary boundary layer height	surface	m
Derived Input	p$_{diff}$	pressure difference between adjacent levels	3D	Pa
	qv$_{sat}$	saturated water vapor mixing ratio	3D	kg/kg
	rh	relative humidity	3D	-
	trigger	boolean trigger indicating convection triggering	surface	-
Input and Output	rthcuten	potential temperature tendency due to cumulus parameterization	3D	K/s
	rqvcuten	water vapor mixing ratio tendency due to cumulus parameterization	3D	kg/kg/s
	rqccuten	cloud water mixing ratio tendency due to cumulus parameterization	3D	kg/kg/s
	rqrcuten	rain water mixing ratio tendency due to cumulus parameterization	3D	kg/kg/s
	rqicuten	cloud ice mixing ratio tendency due to cumulus parameterization	3D	kg/kg/s
	rqscuten	snow mixing ratio tendency due to cumulus parameterization	3D	kg/kg/s
	w0avg	average vertical velocity	3D	m/s
	nca	counter of the cloud relaxation time	3D	s
	pratec	precipitation rate due to cumulus parameterization	surface	mm/s
Output	raincv	precipitation due to cumulus paramterization	surface	mm

To ensure consistency with the dimensions of the 3D variables, the surface variables are padded by duplicating the values of the surface layer for all layers before feeding them into the model. Prior to utilizing the variables in the Bi-LSTM model for training and validation, normalization is applied to ensure uniformity in the magnitudes of all the variables. Each variable is divided by the maximum absolute value in the atmospheric column (for 3D variables) or at the surface (for surface variables).

3 Method

This section describes the flow chart of original MSKF scheme for determining convection trigger, ML model structures and training, and the evaluation methods.

3.1 Description of original MSKF module

The MSKF scheme is a scale-aware adaptation of the KF CP scheme, initially developed by Kain and Fritsch (1990, 1993) and further refined by Kain (2004). Figure 2 illustrates the convection trigger process within the MSKF scheme. At the beginning of each simulation step, the scheme evaluates the variable "nca" to ascertain whether it equals or surpasses a threshold, defined as half the model's time step (dt). Should "nca" equal or exceed the half of dt, there is no need to update convective tendencies or precipitation rates due to ongoing convection. In contrast, a "nca" value below this threshold triggers the MSKF scheme to employ a one-dimensional cloud model. This model calculates a set of variables related to cloud characteristics to evaluate the potential of convection triggering. Essential variables include the lifting condensation level (LCL), convective available potential energy (CAPE), cloud top and base heights, and entrainment rates. The LCL is crucial for determining the emergence of potential convective activities, with a lower LCL favoring more intense convection. CAPE quantifies the buoyant energy available to an air parcel for the formation of deep convective clouds upon reaching its Level of Free Convection (LFC) above the LCL, with higher CAPE values signifying a greater potential for intense convection. The cloud base is generally at the LCL, whereas the cloud top is defined at the altitude where buoyancy becomes negligible. Meanwhile, the vertical extent between the cloud base and top affect the cloud's growth and precipitation potential. The MSKF scheme requires surpassing a specific CAPE threshold to trigger convection. Furthermore, it assesses entrainment rates to measure the impact of ambient air on the evolution of convective system. At grid points where convection is triggered, the MSKF scheme calculates both convective tendencies and precipitation rates; otherwise these values are set to zero. However, the variable "w0avg" is consistently updated, regardless of convection status. Active convection leads to a decrement in "nca" by one model time step for each iteration within WRF model cycle.

3.2 Description of ML-based MSKF scheme

In the original MSKF scheme, atmospheric columns are processed sequentially, one at a time, until all horizontal grid points within the domain have been processed. In contrast, the ML-based MSKF scheme processes data in batches, as indicated by "B" in Figure 3, consisting of 27 features across 44 vertical layers. As a result, the input data has dimensions of B $\times$ 27 $\times$ 44. Before being fed into the ML model, the input data undergoes pre-processing through a module incorporating a 1-dimensional (1D) convolutional layer. This module expands the feature dimension from 27 to 64. The following sections provide a comprehensive description of the structures of the ML model.

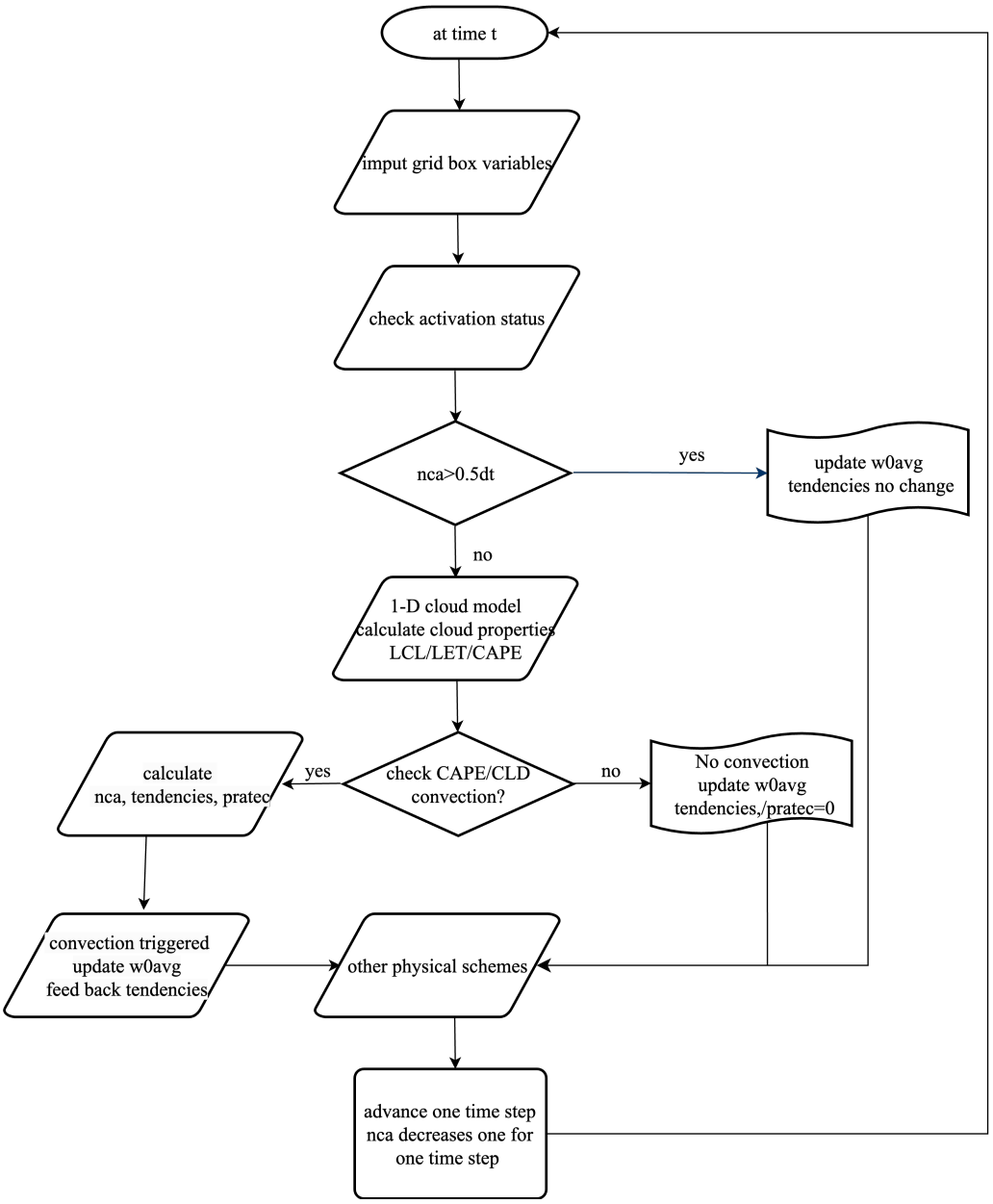

Figure 2. A flow chart outlining convection trigger process in the original MSKF scheme.

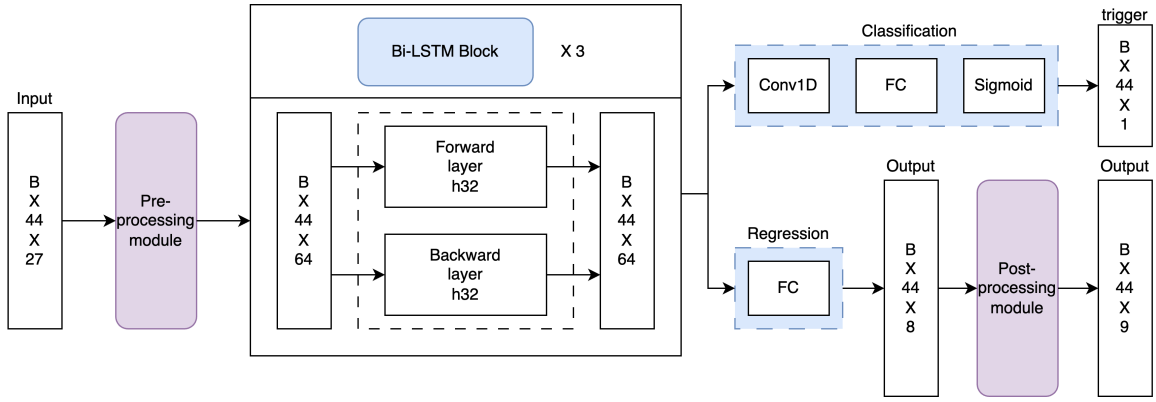

Figure 3. The architecture of the multi-output Bi-LSTM model for combined classification and regression predictions.

3.2.1 ML model structure

Predicting whether convection is triggered as well as modeling convective tendencies and precipitation rates are two main tasks in conventional CP schemes. Regression-based models alone may result in inconsistent tendencies, leading to conflicting signals for triggering convection at specific grid points. Similarly, models solely dependent on classification lack the capability to generate essential tendencies for an effective CP scheme. Therefore, the development of a ML-based CP scheme necessitates the integration of both a binary classification model for the prediction of convection trigger and a regression model for convective tendencies. To address this, we propose a multi-output Bi-LSTM model capable of concurrently conducting regression and classification predictions (Figure 3). Our proposed model consists of a shared Bi-LSTM layer for learning features, a classification subnetwork, and a regression subnetwork. The shared Bi-LSTM layer includes three repeated Bi-LSTM blocks, with each block containing a forward and a backward layer that have a feature dimension of 32. The classification subnetwork is composed of a 1×1 1D convolutional layer, a FC layer, and a Sigmoid activation layer. The output of the Sigmoid layer represents the probability distribution of the convection trigger. The binary cross-entropy loss function is employed as the cost function for this classification task. Meanwhile, the regression subnetwork incorporates a FC layer to output precipitation rate, "nca", and convective tendencies. Finally, outputs from both subnetworks are the processed through a post-processing module to ensure their physical consistency, with further details provided in the subsequent subsection.

3.2.2 Post-processing module

The post-processing module is designed to ensure physical consistency of all variables. To achieve this, the following rules are applied: 1) At grid points where the input "nca" is equal to or greater than half the value of dt, all other variables remain unchanged as they are still within the convection lifetime. 2) The output "nca" must be an integer. 3) At grid points where convection is predicted to be inactive, all corresponding output variables are default to zero. In addition, the calculation of time-step convective precipitation (raincv) follows the methodology outlined in the previous section 2.2.

3.2.3 Model training

As our model incorporates both classification and regression tasks, we optimize its performance by minimizing a multi-task loss function (Ren et al., 2016). The loss function is defined as the sum of the binary cross entropy loss for the convection trigger and a weighted combination of the $L1$ loss for all output variables from the regression subnetwork. The specific formulation of the loss function is as follows:

$$L = \frac{1}{N_{cls}} \sum_{i,j} L_{cls}(p_{i,j}, p_{i,j}^{gt}) + \sum_{c} \lambda_c \frac{1}{N_{reg}} \sum_{i,j} p_{i,j}^{gt} L1_c \tag{1}$$

Here, i and j denote the grid points in the domain. $p_{i,j}$ represents the probability of convection being triggered. The ground-truth label $p_{i,j}^{gt}$ takes a value of 1 if convection is triggered and 0 otherwise. The classification loss L_{cls} is calculated using the binary cross entropy loss. For the regression loss of different variables c, λ_c functions as a weight that balances the output variables by considering their respective magnitudes. The term $p_{i,j}^{gt} L1_c$ indicates that the $L1$ regression loss is activated only for triggered grid points ($p_{i,j}^{gt} = 1$) and is disabled otherwise ($p_{i,j}^{gt} = 0$). Both loss terms are normalized by N_{cls} and N_{reg}, which correspond to the total number of grid points and the number of triggered grid points, respectively.

Adam optimizer (Kingma and Ba, 2014) is used with an initial learning rate of 0.003 update the parameters of the model. Furthermore, the plateau scheduler is implemented to decrease the learning rate by a factor of 0.5 when the loss fails to decrease for five epochs. The model is trained for 150 epochs using a batch size of 44000.

3.3 Evaluation methods

The ML-based MSKF scheme is evaluated in both offline and online settings. The offline performance of the ML-based MSKF scheme is evaluated by comparing it against the outputs of the original MSKF scheme using validation dataset, including rthcuten, rqvcuten, rqccuten, rqrcuten, nca, and pratec. The overall model performance metrics include RMSE and correlation coefficient. The mean absolute error (MAE) and mean bias error (MBE) per vertical layer were were calculated using the equation below:

$$MAE_l = \frac{1}{N} \sum_{i=1}^{N} |Y_{ML}(i,l) - Y(i,l)| \tag{2}$$

$$MBE_l = \frac{1}{N} \sum_{i=1}^{N} Y_{ML}(i,l) - Y(i,l) \tag{3}$$

where $Y(i,l)$ and $Y_{ML}(i,l)$ represent the outputs from the original MSKF scheme and ML-based MSKF scheme, respectively. Here, i denotes the horizontal grid point of a vertical profile, N is the number of the horizontal grid points in the domain, l represents the vertical layer index.

4 Results

4.1 Offline validation of the ML-based MSKF scheme

The offline validation was conducted using data that was not used during the training process. Figure 4 compares the cloud relaxation time (nca), precipitation rate (pratec), and convective tendencies predicted by both the original MSKF scheme and the ML-based MSKF scheme, respectively.To facilitate the comparison, the units of precipitation rate and temperature tendencies were converted to mm·d^{-1} and K·d^{-1} from mm·s^{-1} and K·s^{-1}, respectively, by applying a conversion factor of 86,400 (24 $\times$ 3600). Similarly, the water vapor mixing ratio (rqvcuten), cloud water mixing ratio (rqccuten), and rain water mixing ratio (rqrcuten) due to convection were multiplied by 86,400,000 (24 $\times$ 3600 $\times$ 1000) to convert from kg·kg^{-1}·s^{-1} to g·kg^{-1}·d^{-1}. Among the output variables listed in Table 1, the variable w0avg, is excluded as it is calculated using an equation with the ground truth as input in this offline validation. Hence, evaluating w0avg in the offline evaluation is unnecessary.

Among all variables illustrated in Figure 4, the variable "nca" exhibits a significantly higher RMSE of 4.32, with data points widely dispersed across a wide range of values. This suggests that accurately predicting convection poses a considerable challenge. To eliminate the dependency on time steps, "nca" is divided by the model's time step of 15 seconds before proceeding with plotting and statistical evaluations. The precipitation rate demonstrates the highest correlation coefficient and minimal variability, as most data points cluster closely around the 1:1 line. While temperature and the four moisture tendencies exhibit some degree of variability, the majority of data points align closely to the 1:1 line. The correlation coefficient of convection trigger is 0.91, not shown in Figure 4. Overall, the ML-based MSKF scheme shows a strong correlation with the original MSKF scheme for all examined variables, with correlation coefficients consistently higher than 0.91. This indicates that the the ML-based MSKF scheme has the potential to replace the original scheme.

To obtain a comprehensive understanding of the vertical distribution of errors, Figure 5 presents the vertical profiles of error statistics associated with convective tendencies. The solid and dotted lines in the figure represent the MAE and MBE of tendencies at each vertical layer, respectively. Additionally, the shaded area corresponds to the 5th and 95th percentiles of differences between tendencies predicted by the ML-based MSKF predicted scheme and the original MSKF scheme, respectively. The distribution of vertical errors in all tendencies exhibits a notable uniformity, with higher variance observed within the pressure layers between 800 and 1,000 hPa. These pressure layers correspond to the atmospheric layer where convection occurs most frequently. Due to the significantly lower cloud and rain content compared to water vapor in the atmosphere, the error magnitudes for rqccuten and rqrcuten are considerably lower than those observed for rqvcuten.

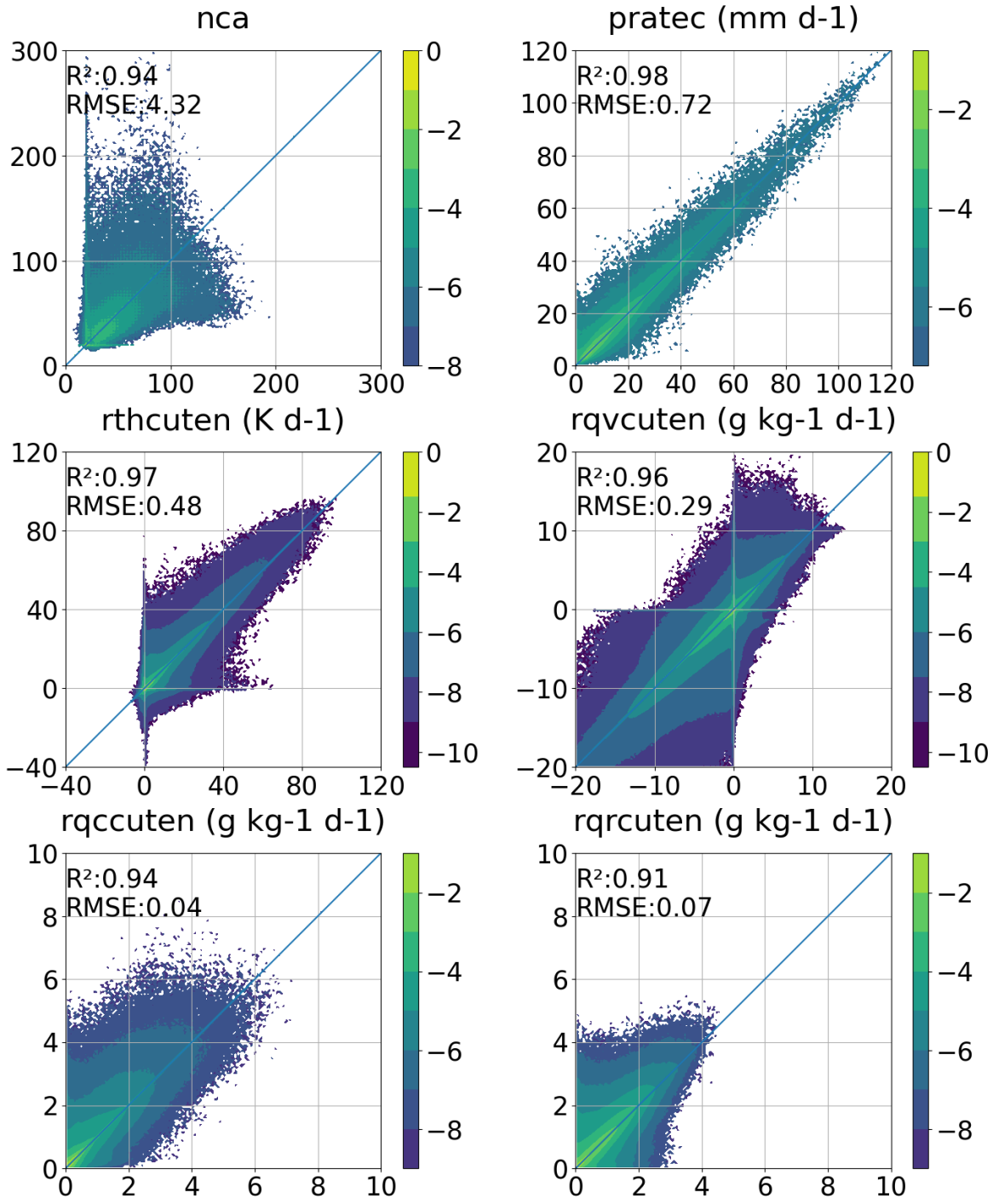

Figure 4. Comparison of the predicted (y axis) and true (x axis) nca, pratec, rthcuten (first column), rqvcuten (second column), rqccuten (third column), and rqrcuten for using validation data in the offline setting. Colors indicate the proportion of samples across the entire testing dataset, with values on the colorbar normalized through the application of a logarithm base 10.

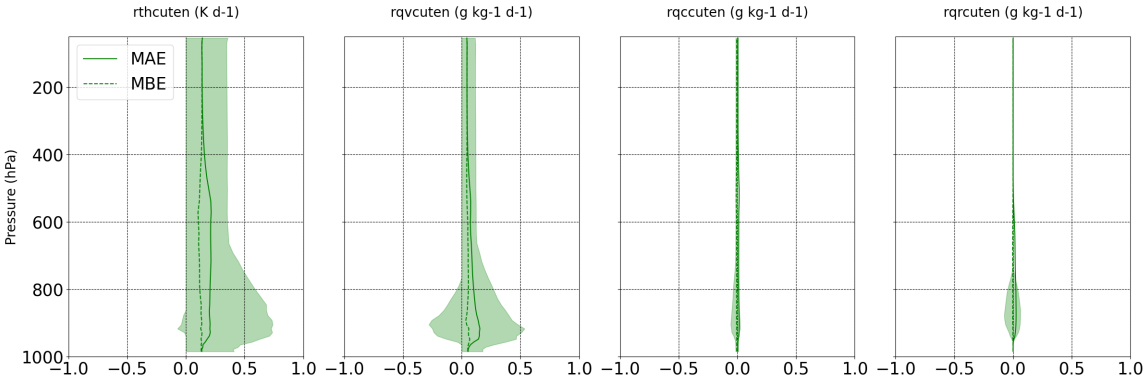

Figure 5. Vertical profiles of the statistics in rthcuten (first column), rqvcuten (second column), rqccuten (third column), and rqrcuten (fourth column) using validation data in the offline setting data using ML-based emulators. The solid and dotted lines show the MAE and MBE profile, respectively, and the shaded area indicates the 5th and 95th percentile of differences (prediction—target) at each layer.

4.2 Prognostic validation

This subsection presents the performance of the ML-based MSKF scheme in the online setting.

The ML-based MSKF scheme was integrated into the WRF model as a substitute for the original MSKF scheme to simulate convective processes. Utilizing the WRF-ML coupler (Zhong et al., 2023a), this novel ML-based MSKF scheme was seamlessly incorporated into the WRF framework. A comparative analysis was conducted by initializing both the modified WRF model, which incorporates the ML-based scheme, and the original WRF model on June 12, 14, 16, and 18, 2022, for simulations extending over 168 hours. It is worth mentioning that these simulations were performed independently of the training dataset, ensuring the evaluation of the scheme's generalization capability.

Figure 6 presents the averaged spatial forecasts for predictions generated by the original WRF model. These forecast results include the accumulations of both convective precipitation ($RAINC$) and non-convective precipitation ($RAINNC$) over a 12-hour period, along with the 2-meter temperature ($T2M$) at 24, 72, 120, and 168 hours. The figure also demonstrates the mean absolute difference (MAD) between WRF simulations coupled with the ML-based MSKF scheme and those utilizing the original MSKF scheme. Within the spatial forecasts, red and blue patterns signify the magnitudes of the forecasted values, whereas in the spatial differences, these colors denote the positive and negative biases in the ML-based simulations, respectively. Green patterns suggest minimal deviation from the original WRF simulations. Furthermore, we calculate a domain-averaged MAD to evaluate the overall performance of the ML-based scheme in prognostic simulations. Generally, the differences are small, indicating good agreement between WRF simulations coupled with the ML-based MSKF scheme and the original WRF simulations. Notably, the differences do not increase with the progression of simulation time, as evidenced by a comparable domain-averaged MAD at 168 forecast hours compared to that at 24 forecast hours. These findings suggest that the ML-based MSKF scheme achieves stable prognostic simulations.

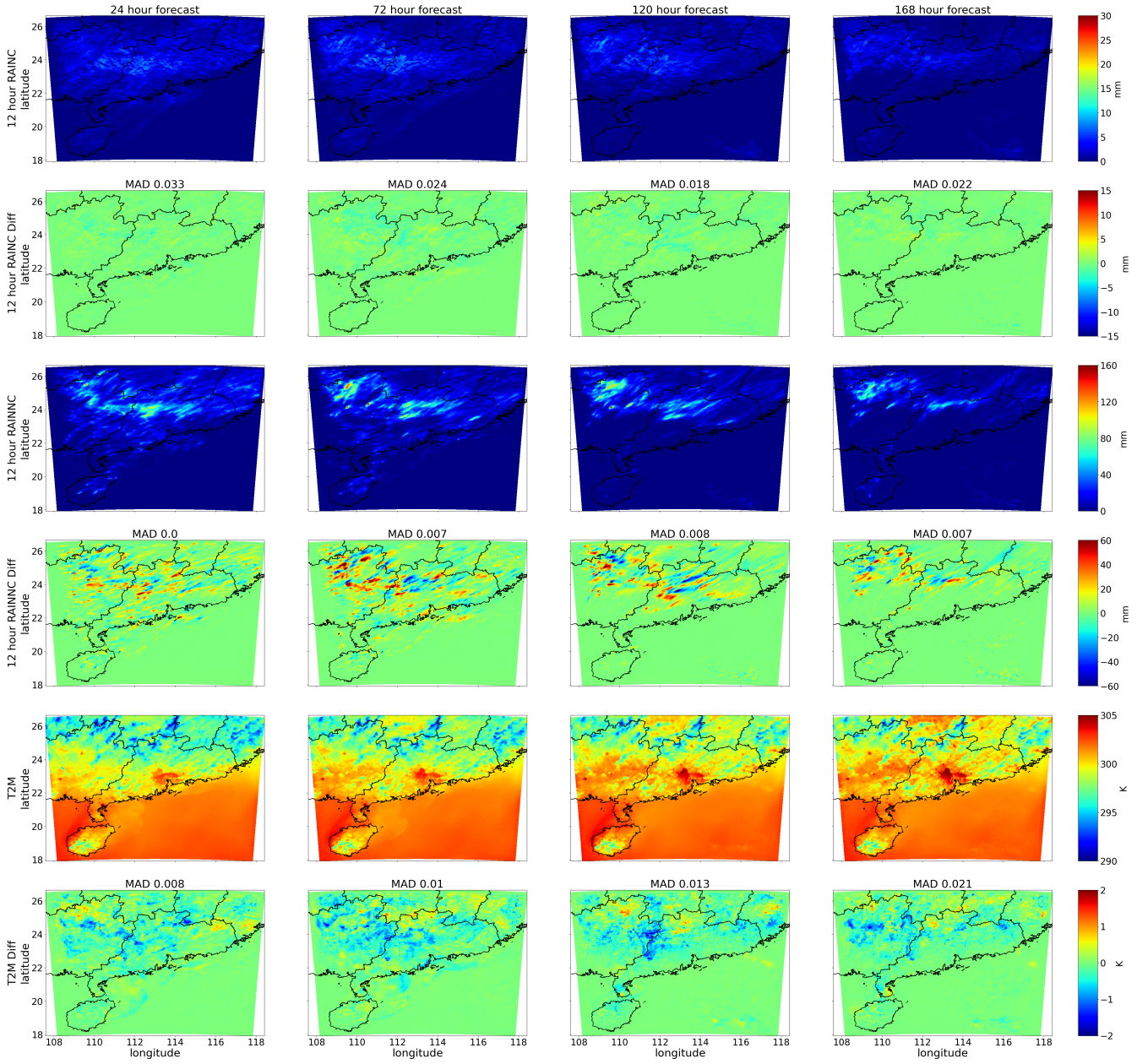

Figure 6. Spatial map of the average WRF simulations using the original MSKF scheme (in the first, third, and fifth rows) along with the average MAD between WRF simulations coupled with the ML-based MSKF scheme and WRF simulation with the original MSKF scheme (in the second, fourth and sixth rows). The simulations are shown for the 12-hour accumulated convective precipitation ($RAINC$) in the first and second rows, the 12-hour accumulated non-convective precipitation ($RAINNC$) in the third and fourth rows, and the 2-meter temperature ($T2M$) at forecast lead times of 24 hours (first column), 72 hours (second column), 120 hours (third column), and 168 hours (fourth column).

Figure 7 provides a comparative analysis of domain-averaged time series forecasts from both the original WRF simulations and WRF simulations coupled with the ML-based MSKF scheme. This comparison includes 6-hour accumulations of $RAINC$ and $RAINNC$, as well as $T2M$ forecasts. The results demonstrate that WRF simulations coupled with the ML-based MSKF schemes are in close alignment with the original WRF simulations, particularly in capturing the diurnal variations of $RAINC$, $RAINNC$, and $T2M$. Notably, the $T2M$ forecasts demonstrate remarkable consistency, underscoring the efficacy of the ML-based MSKF scheme in maintaining the predictive accuracy of the original scheme.

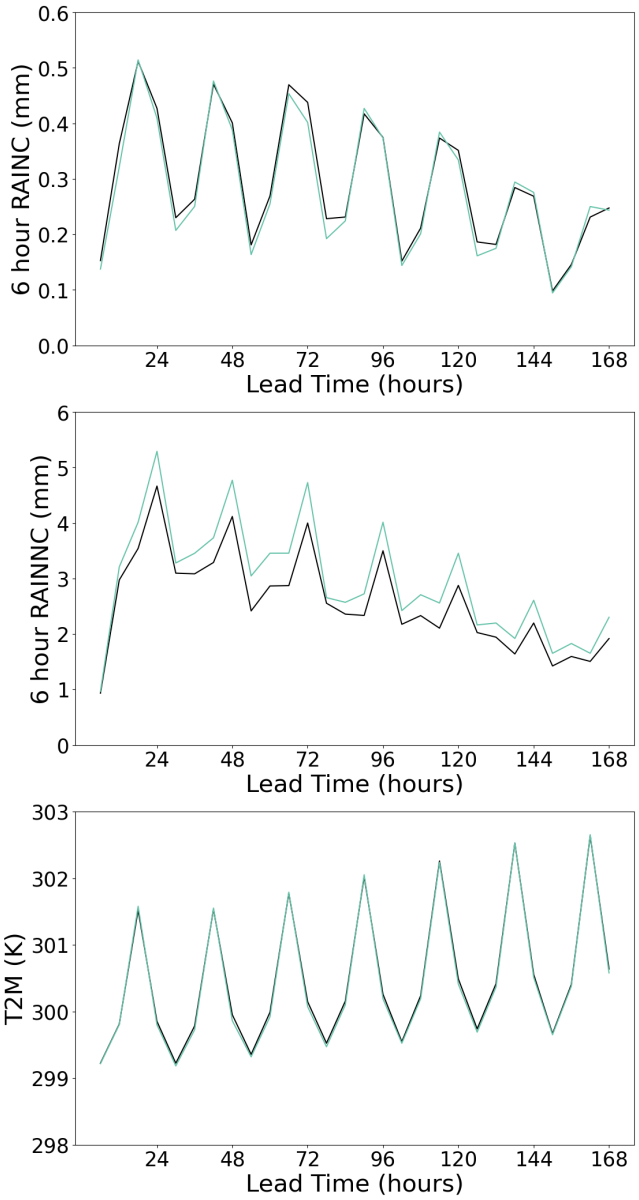

Figure 7. Comparison of domain-averaged forecasts derived from the original WRF simulations (black lines) and WRF simulations coupled with the ML-based MSKF scheme (light green lines) of 6-hour accumulated $RAINC$ (first row) and $RAINNC$ (second row), along with $T2M$ (third row).

5 Conclusions

In this paper, we proposed a multi-output Bi-LSTM model to develop a ML-based MSKF scheme for predicting convection trigger and reproducing the convective process in the gray zone. The model is trained on data generated by the WRF simulations at a spatial resolution of 5 km, covering the South China region. The output variables of the ML-based MSKF scheme are identical to those of the original MSKF scheme, encompassing cloud relaxation time ("nca"), precipitation rate ("pratec"), time-step convective precipitation ("raincv"), and convective tendencies. This ML-based scheme ensures physical consistency among all output variables by incorporating a post-processing module to refine the output from the Bi-LSTM model. Offline validation demonstrates the excellent performance of the ML-based MSKF scheme. Furthermore, the ML-based MSKF scheme is coupled with the WRF model using WRF-ML coupler. The WRF simulations coupled with the ML-based MSKF scheme is compared against the WRF simulation with the original MSKF scheme. Results shows that the ML-based scheme can generate forecasts similar to the original ML scheme in online settings, showing the potential substitution of the MSKF scheme by ML models in gray-zone.

This study demonstrates the feasibility of employing ML models as substitutes for conventional CP scheme within the high-resolution weather forecasting model. Future efforts will focus on the development of ML models, based on data generated by super-parameterization or cloud-resolving models, to replace conventional CP schemes in weather forecasting models. The objective of this substitution is to reduce uncertainties and improve performance of weather forecast models.

Code and data availability. The source code for the WRF model version 4.3 used in this study is available at https://doi.org/10.5281/zenodo.10039053 (Skamarock et al., 2023). The source code and data used in this are available at https://doi.org/10.5281/zenodo.10032404 (Zhong et al., 2023b).

Author contributions. X.Z. trained the deep learning models and calculate the statistics of model performance. X.Y. and X.Z. conducted the WRF simulations to provide dataset for training and evaluation, and offered valuable suggestions on the model training and paper revision. X.Z. and X.Y. wrote, reviewed and edited the original draft; X.Z., X.Y., and H.L. supervised and supported this research and gave important opinions. All of the authors have contributed to and agreed to the published version of the manuscript.

Competing interests. The authors declare no conflict of interest.

Acknowledgements. This work was supported by Basic and Applied Basic Research Foundation of Guangdong Province, under Grant No. 2021A1515012582. We are thankful for Mesoscale and Microscale Meteorology Laboratory (MMM) at NCAR for developing and sharing the WRF source codes.

References

Arakawa, A.: The Cumulus Parameterization Problem: Past, Present, and Future, Journal of Climate, 17, 2493 – 2525, https://doi.org/https://doi.org/10.1175/1520-0442(2004)017<2493:RATCPP>2.0.CO;2, 2004.

Bougeault, P. and Lacarrère, P.: Parameterization of orographic induced turbulence in a mesobeta scale model, Monthly Weather Review, 117, 1872–1890, https://doi.org/10.1175/1520-0493(1989)117<1872:POOITI>2.0.CO;2, 1989.

Brenowitz, N. D. and Bretherton, C. S.: Prognostic Validation of a Neural Network Unified Physics Parameterization, Geophysical Research Letters, 45, 6289–6298, https://doi.org/https://doi.org/10.1029/2018GL078510, 2018.

Brenowitz, N. D. and Bretherton, C. S.: Spatially Extended Tests of a Neural Network Parametrization Trained by Coarse-Graining, Journal of Advances in Modeling Earth Systems, 11, 2728–2744, https://doi.org/https://doi.org/10.1029/2019MS001711, 2019.

Chan, S. C., Kendon, E. J., Fowler, H. J., Blenkinsop, S., Ferro, C. A. T., and Stephenson, D. B.: Does increasing the spatial resolution of a regional climate model improve the simulated daily precipitation?, Climate Dynamics, 41, 1475–1495, https://doi.org/10.1007/s00382-012-1568-9, 2013.

Clark, A. J. et al.: An Overview of the 2010 Hazardous Weather Testbed Experimental Forecast Program Spring Experiment, Bulletin of the American Meteorological Society, 93, 55 – 74, https://doi.org/10.1175/BAMS-D-11-00040.1, 2012.

Di, X. R., Xiong, Z. S., and Beijing: A Study of Circumstances of Meso-β-Scale Systems of Strong Heavy Rainfall in Warm Sector Ahead of Fronts in South China, Chinese Journal of Atmospheric Sciences/Daqii kexue, 2006.

Ding, Y.: Seasonal march of the East-Asian summer monsoon, World Scientific, 2004.

Gentine, P., Pritchard, M., Rasp, S., Reinaudi, G., and Yacalis, G.: Could Machine Learning Break the Convection Parameterization Deadlock?, Geophysical Research Letters, 45, 5742–5751, https://doi.org/https://doi.org/10.1029/2018GL078202, 2018.

Giorgi, F., Torma, C., Coppola, E., Ban, N., Schär, C., and Somot, S.: Enhanced summer convective rainfall at Alpine high elevations in response to climate warming, Nature Geoscience, 9, 584–589, https://doi.org/10.1038/ngeo2761, 2016.

Grell, G. A. and Dévényi, D.: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques, Geophysical Research Letters, 29, 38–1, 2002.

Grell, G. A. and Freitas, S. R.: A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling, Atmospheric Chemistry and Physics, 14, 5233–5250, https://doi.org/10.5194/acp-14-5233-2014, 2014.

Han, J. and Pan, H.-L.: Revision of convection and vertical diffusion schemes in the NCEP Global Forecast System, Weather and Forecasting, 26, 520–533, 2011.

Han, Y., Zhang, G. J., Huang, X., and Wang, Y.: A Moist Physics Parameterization Based on Deep Learning, Journal of Advances in Modeling Earth Systems, 12, e2020MS002 076, https://doi.org/https://doi.org/10.1029/2020MS002076, e2020MS002076 2020MS002076, 2020.

He, K., Zhang, X., Ren, S., and Sun, J.: Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778, 2016.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, 2020.

Hong, S.-Y. and Dudhia, J.: Next-Generation Numerical Weather Prediction: Bridging Parameterization, Explicit Clouds, and Large Eddies, Bulletin of the American Meteorological Society, 93, ES6 – ES9, https://doi.org/https://doi.org/10.1175/2011BAMS3224.1, 2012.

Hong, S.-Y. and Lim, J.-O. J.: The WRF single-moment 6-class microphysics scheme (WSM6), Asia-Pacific Journal of Atmospheric Sciences, 42, 129–151, 2006.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, Journal of Geophysical Research: Atmospheres, 113, 2008.

Janjic, Z.: The surface layer parameterization in the NCEP Eta Model, World Meteorological Organization-Publications-WMO TD, pp. 4–16, 1996.

Jeworrek, J., West, G., and Stull, R.: Evaluation of Cumulus and Microphysics Parameterizations in WRF across the Convective Gray Zone, Weather and Forecasting, 34, 1097 – 1115, https://doi.org/https://doi.org/10.1175/WAF-D-18-0178.1, 2019.

Jian, S., Ping, Z., and Xiuji, Z.: THE MESOSCALE STRUCTURE OF A SOUTH CHINA RAINSTORM AND THE INFLUENCE OF COMPLEX TOPOGRAPHY (in Chinese), Acta Meteorologica Sinica, pp. 333–342, https://doi.org/10.11676/qxxb2002.040, 2002.

Johnson, A., Wang, X., Kong, F., and Xue, M.: Object-Based Evaluation of the Impact of Horizontal Grid Spacing on Convection-Allowing Forecasts, Monthly Weather Review, 141, 3413 – 3425, https://doi.org/https://doi.org/10.1175/MWR-D-13-00027.1, 2013.

Kain, J. S.: The Kain–Fritsch Convective Parameterization: An Update, Journal of Applied Meteorology, 43, 170 – 181, https://doi.org/https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2, 2004.

Kain, J. S. and Fritsch, J. M.: A One-Dimensional Entraining/Detraining Plume Model and Its Application in Convective Parameterization, Journal of Atmospheric Sciences, 47, 2784 – 2802, https://doi.org/https://doi.org/10.1175/1520-0469(1990)047<2784:AODEPM>2.0.CO;2, 1990.

Kain, J. S. and Fritsch, J. M.: Convective Parameterization for Mesoscale Models: The Kain-Fritsch Scheme, pp. 165–170, American Meteorological Society, Boston, MA, https://doi.org/10.1007/978-1-935704-13-3_16, 1993.

Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, 2014.

Krasnopolsky, V. M., Fox-Rabinovitz, M. S., and Belochitski, A. A.: Using ensemble of neural networks to learn stochastic convection parameterizations for climate and numerical weather prediction models from data simulated by a cloud resolving model, Advances in Artificial Neural Systems, 2013, 5–5, 2013.

Lean, H. W., Clark, P. A., Dixon, M., Roberts, N. M., Fitch, A., Forbes, R., and Halliwell, C.: Characteristics of High-Resolution Versions of the Met Office Unified Model for Forecasting Convection over the United Kingdom, Monthly Weather Review, 136, 3408 – 3424, https://doi.org/10.1175/2008MWR2332.1, 2008.

Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., and Anandkumar, A.: Fourier neural operator for parametric partial differential equations, arXiv preprint arXiv:2010.08895, 2020.

Livneh, B., Restrepo, P. J., and Lettenmaier, D. P.: Development of a unified land model for prediction of surface hydrology and land–atmosphere interactions, Journal of Hydrometeorology, 12, 1299–1320, 2011.

Luo, Y., Zhang, R., Wan, Q., Wang, B., Wong, W. K., Hu, Z., Jou, B. J.-D., Lin, Y., Johnson, R. H., Chang, C.-P., Zhu, Y., Zhang, X., Wang, H., Xia, R., Ma, J., Zhang, D.-L., Gao, M., Zhang, Y., Liu, X., Chen, Y., Huang, H., Bao, X., Ruan, Z., Cui, Z., Meng, Z., Sun, J., Wu, M., Wang, H., Peng, X., Qian, W., Zhao, K., and Xiao, Y.: The Southern China Monsoon Rainfall Experiment (SCMREX), Bulletin of the American Meteorological Society, 98, 999 – 1013, https://doi.org/https://doi.org/10.1175/BAMS-D-15-00235.1, 2017.

Mishra, S. K., Anand, A., Fasullo, J., and Bhagat, S.: Importance of the Resolution of Surface Topography in Indian Monsoon Simulation, Journal of Climate, 31, 4879 – 4898, https://doi.org/https://doi.org/10.1175/JCLI-D-17-0324.1, 2018.

O'Gorman, P. A. and Dwyer, J. G.: Using Machine Learning to Parameterize Moist Convection: Potential for Modeling of Climate, Climate Change, and Extreme Events, Journal of Advances in Modeling Earth Systems, 10, 2548–2563, https://doi.org/https://doi.org/10.1029/2018MS001351, 2018.

Onishi, R., Hirai, J., Kolomenskiy, D., and Yasuda, Y.: Real-Time High-Resolution Prediction of Orographic Rainfall for Early Warning of Landslides, pp. 237–248, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-031-16898-7_17, 2023.

Ou, T., Chen, D., Chen, X., Lin, C., Yang, K., Lai, H.-W., and Zhang, F.: Simulation of summer precipitation diurnal cycles over the Tibetan Plateau at the gray-zone grid spacing for cumulus parameterization, Climate Dynamics, 54, 3525–3539, 2020.

Rasp, S.: Coupled online learning as a way to tackle instabilities and biases in neural network parameterizations: general algorithms and Lorenz 96 case study (v1.0), Geoscientific Model Development, 13, 2185–2196, https://doi.org/10.5194/gmd-13-2185-2020, 2020.

Rasp, S., Pritchard, M. S., and Gentine, P.: Deep learning to represent subgrid processes in climate models, Proceedings of the National Academy of Sciences, 115, 9684–9689, https://doi.org/10.1073/PNAS.1810286115, 2018.

Ren, S., He, K., Girshick, R., and Sun, J.: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, 2016.

Roberts, N. M. and Lean, H. W.: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events, Monthly Weather Review, 136, 78–97, 2008.

Schumacher, V., Fernández, A., Justino, F., and Comin, A.: WRF high resolution dynamical downscaling of precipitation for the Central Andes of Chile and Argentina, Frontiers in Earth Science, 8, 328, 2020.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J., Duda, M., Barker, D., et al.: A description of the advanced research WRF version 4, NCAR tech. note ncar/tn-556+ str, 145, https://doi.org/doi:10.5065/1dfh-6p97, 2019.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D. M., and Xiang-yu, H.: A Description of the Advanced Research WRF Model Version 4.3, National Center for Atmospheric Research: Boulder, CO, USA, https://doi.org/doi:10.5065/1dfh-6p97, 2021.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J., Duda, M., Barker, D., et al.: The source codes of WRF-V4.3 and WPS-V4.3 [Software]. Zenodo. https://doi.org/10.5281/zenodo.10039053, 2023.

Tao, S. Y.: Storm Rainfall in China (in Chinese)., Science Press, Beijing, 1981.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I.: Attention is all you need, Advances in neural information processing systems, 30, 2017.

Villalba-Pradas, A. and Tapiador, F. J.: Empirical values and assumptions in the convection schemes of numerical models, Geoscientific Model Development, 15, 3447–3518, https://doi.org/10.5194/gmd-15-3447-2022, 2022.

Wyngaard, J. C.: Toward Numerical Modeling in the "Terra Incognita", Journal of the Atmospheric Sciences, 61, 1816 – 1826, https://doi.org/https://doi.org/10.1175/1520-0469(2004)061<1816:TNMITT>2.0.CO;2, 2004.

Xia, R. and Zhao, S.: Diagnosis and modeling of meso–scale systems of heavy rainfall in warm sector ahead of front in South China (middle part of Guangdong province) in June 2005, Chinese journal of atmospheric sciences/Daqi Kexue, 33, 2009.

Yao, Y., Zhong, X., Zheng, Y., and Wang, Z.: A Physics-Incorporated Deep Learning Framework for Parameterization of Atmospheric Radiative Transfer, Journal of Advances in Modeling Earth Systems, 15, e2022MS003 445, https://doi.org/https://doi.org/10.1029/2022MS003445, e2022MS003445 2022MS003445, 2023.

Yuval, J. and O'Gorman, P. A.: Stable machine-learning parameterization of subgrid processes for climate modeling at a range of resolutions, Nature communications, 11, 3295, 2020.

Yuval, J., O'Gorman, P. A., and Hill, C. N.: Use of neural networks for stable, accurate and physically consistent parameterization of subgrid atmospheric processes with good performance at reduced precision, Geophysical Research Letters, 48, e2020GL091 363, 2021.

Zhang, X. and Ni, Y.: A comparative study of a frontal and a non-frontal convective systems, Acta Meteorologica Sinica, pp. 108–121, https://doi.org/10.11676/qxxb2009.012, 2009.

Zhang, X., Bao, J.-W., Chen, B., and Huang, W.: Evaluation and Comparison of Two Deep Convection Parameterization Schemes at Convection-Permitting Resolution, Monthly Weather Review, 149, 3419 – 3432, https://doi.org/https://doi.org/10.1175/MWR-D-21-0016.1, 2021.

Zhao, P., Zhang, R., Liu, J., Zhou, X., and He, J.: Onset of southwesterly wind over eastern China and associated atmospheric circulation and rainfall, Climate Dynamics, 28, 797–811, 2007.

Zheng, Y., Alapaty, K., Herwehe, J. A., Genio, A. D. D., and Niyogi, D.: Improving High-Resolution Weather Forecasts Using the Weather Research and Forecasting (WRF) Model with an Updated Kain–Fritsch Scheme, Monthly Weather Review, 144, 833 – 860, https://doi.org/https://doi.org/10.1175/MWR-D-15-0005.1, 2016.

Zhong, L., Mu, R., Zhang, D., Zhao, P., Zhang, Z., and Wang, N.: An observational analysis of warm-sector rainfall characteristics associated with the 21 July 2012 Beijing extreme rainfall event, Journal of Geophysical Research, 120, 3274–3291, 2015.

Zhong, S. and Chen, Z.: The Impacts of Atmospheric Moisture Transportation on Warm Sector Torrential Rains over South China, Atmosphere, 8, https://doi.org/10.3390/atmos8070116, 2017.

Zhong, X., Ma, Z., Yao, Y., Xu, L., Wu, Y., and Wang, Z.: WRF–ML v1.0: a bridge between WRF v4.3 and machine learning parameterizations and its application to atmospheric radiative transfer, Geoscientific Model Development, 16, 199–209, https://doi.org/10.5194/gmd-16-199-2023, 2023a.

Zhong, X., Yu, X., and Li, H.: Machine Learning Parameterization of the Multi-scale Kain-Fritsch (MSKF) Convection Scheme and stable simulation coupled in WRF using WRF-ML v1.0 (Version 1.0) [Dataset] [Software]. Zenodo. https://doi.org/10.5281/zenodo.10032404, 2023b.