Improving flood forecasts capability of Taihang Piedmont basin by optimizing WRF parameter combination and coupling with HEC-HMS

Ting Zhang[1] · Ya Gao[1] · Ping Yu[2] · Jianzhu Li[1] · Ping Feng[1] · Huixin Ma[2]

Abstract

Based on numerical weather prediction model Weather Research and Forecasting (WRF) and Hydrologic Modeling System (HEC-HMS), a coupling model is constructed in Taihang Piedmont basin. The WRF model parameter scheme combinations composed of microphysics, planetary boundary layers, and cumulus parameterizations suitable for the study area are optimized. In both time and space, we tested the effects of the WRF model by a multi-index evaluation system composed of relative error, root meantime square error, probability of detection, false alarm ratio, and critical success index and established this system in two stages. A multi-attribute decision-making model based on Technique for Order Preference by Similarity to an Ideal Solution and grey correlation degree is proposed to optimize each parameter scheme. Among 18 parameter scheme combinations, Mellor-Yamada-Janjic, Grell-Devinji, Purdue-Lin, Betts-Miller-Janjić, and Single-Moment6 are ideal choices according to the simulation performance in both time and space. Using the unidirectional coupling method, the rolling rainfall forecast results of the WRF model in the 24 h and 48 h forecast periods are input to HEC-HMS hydrological model to simulate three typical floods. The coupling simulation results are better than the traditional forecast method, and it prolongs the flood forecast period of the Taihang Piedmont basin.

1 Introduction

Taihang Mountain is located between Shanxi Province and North China Plain, with an altitude above 1200 m. In the east of Taihang Mountain, the precipitation system is greatly affected by the terrain and weather system, which is easy to cause rainstorms and floods. Effectively improving the accuracy of flood forecasts and prolonging the forecast period are important measures for flood control and disaster reduction. A coupled atmospheric-hydrological flood forecasting model is constructed by combining numerical weather forecasts with a hydrological model. Different from traditional flood forecasting, which can only be carried out after the beginning of rainfall, the coupled model can effectively prolong the forecast period. With the progress of computer technology, numerical weather forecasting has developed rapidly, which provides strong support for the work of the atmospheric-hydrological coupling model.

There are two ways of coupling atmospheric and hydrological models: unidirectional coupling and bidirectional coupling. For unidirectional coupling, the hydrological model is driven with the output rainfall results of the climate model to predict flood and runoff (Vincendon et al. 2009), while bidirectional coupling fully considers the feedback mechanism between atmosphere and land. However, the bidirectional coupling is not flexible and the debugging is difficult. Jasper and Kaufmann (2003) used the numerical atmospheric prediction model Swiss model and Canadian

✉ Ping Yu
    yuping@tjau.edu.cn

    Ting Zhang
    zhangting_hydro@tju.edu.cn

    Ya Gao
    ya970525@163.com

    Jianzhu Li
    lijianzhu@tju.edu.cn

    Ping Feng
    fengping@tju.edu.cn

    Huixin Ma
    17622739262@163.com

1   State Key Laboratory of Hydraulic Engineering Simulation and Safety, Tianjin University, Tianjin 300072, China

2   College of Water Conservancy Engineering, Tianjin Agricultural University, Tianjin 300384, China

Mesoscale Compressible Community model (MC2) to drive the hydrological model WaSiM-ETH in high altitude area of the Southern Alps. Some studies have shown that high-resolution precipitation data is needed to accurately simulate the hydrological process in a large watershed. Yu et al. (1999) established unidirectional coupling between Penn State-NCAR Mesoscale Meteorological Model and HEC-HMS and proved this conclusion by simulating three typical storm rainfall events. Lin et al. (2006) constructed the coupling of MC2 and Xinanjiang model in Huaihe watershed to simulate the heavy rainfall and flood between 1998 and 2003. The results show that MC2 tends to overestimate the average precipitation of this area. However, the overall flood simulation is satisfactory, which shows the feasibility of applying this method to flood forecasting research. In recent years, numerical weather forecasting technology has been greatly developed. Weather Research and Forecasting (WRF) model has been widely used in land–atmosphere coupling for hydrological forecasting. Flesch et al. (2012) simulated two typical flood events using WRF model in Alberta, Canada, which proved that WRF model is sensitive to terrain. With the decrease of mountain altitude, the maximum precipitation simulated by WRF in mountainous and hilly areas decreased by 50%. Li et al. (2017) put the precipitation data of different forecast periods simulated by WRF quantitative precipitation forecast into the distributed hydrological Liuxi River model. From the results, it can be seen that with the increase of forecast period, the accuracy of flood forecasting becomes lower. The precipitation data output by the climate model has a certain systematic error, but the simulation performance can be significantly improved through some correction methods. Tang et al. (2014) proved that the coupling of WRF model and VIC model can simulate the evaporation, soil water content, and runoff well, and the coupling model can effectively restore the daily and monthly data. There is a certain gap between the simulation and actual runoff, and it is mainly due to the precipitation simulation of WRF model.

With the increasing demand of improving rainfall simulation accuracy in mesoscale climate models, many pieces of research have been carried out on the sensitivity of different parameter schemes of the numerical weather forecast. William et al. (2005) compared the sensitivity of different dynamic cores, physical parameter schemes, and initial conditions of WRF to rainfall simulation, indicating that the sensitivity of various components was influenced by each other. However, the biggest change comes from the physical parameter scheme. Pennelly et al. (2014) used the WRF model to simulate the precipitation of three flood events in Alberta, Canada. Through four evaluation indices, they evaluated the simulation effects of different physical schemes. It was found that the Kain-Fritsch (KF) scheme was the best for heavy rainfall events, and each evaluation index of the

combination of the KF scheme and explicit cumulus parameterization schemes performed well. Ding et al. (2019) used two radiation schemes, three cumulus parameterizations, four microphysics, and two PBLs of WRF model to form eight parameter scheme combinations in Qinba mountain and simulated the precipitation and temperature in January and July 2015 by using NCEP Final (FNL) data. It shows the precipitation simulation effect of the Dudhia, KF, Ferrier, Yonsei University (YSU) combination in winter is the best, and the combination of Dudhia, KF, Lin, YSU performs best in summer. The above researches show that the sensitivity of the parameter schemes in different regions has a certain difference. For a specific area, optimizing the combination of parameterization schemes suitable for the region can effectively improve the accuracy of precipitation simulation, thereby further improving the accuracy of flood forecasting.

Based on the above research, this study carried out flood forecasting by a land–atmosphere coupling model in the Taihang Piedmont basin. A phased multi-index evaluation system was established to evaluate the rainfall output by the climate model in both time and space. A multi-attribute decision-making model combining Technique for Order Preference by Similarity to an Ideal Solution (TOPSIS) optimization and grey correlation degree was constructed to optimize the parameter scheme. Then based on WRF model and HEC-HMS, typical flood events in the study area from 1975 to 2018 were simulated. After spatial and temporal scale conversion, the rainfall data output by WRF model was used to drive HEC-HMS for flow prediction. This paper explored the applicability of the coupling scheme between these two models in the study area, providing better support for real-time flood forecasting in Taihang Piedmont basin.

2 Data and methods

2.1 Study area and data

The research area is located in Xingtai City, Hebei Province, east of Taihang Mountains (Fig. 1). The catchment belongs to Li River Basin, with a total area of 1684 km^2, and the outlet station is Duanzhuang hydrological station. The study area is located in a typical temperate continental monsoon climate area with high terrain in the west and low terrain in the east. Affected by the topographic uplift of Taihang Mountains and low latitude weather system, the precipitation mainly occurs in summer with short duration and high intensity. Its precipitation in summer accounts for 70% of the whole year and is distributed unevenly throughout the year. Two large- and medium-sized reservoirs have been built in this area. Zhuzhuang reservoir is a large-sized water conservancy engineering focusing on flood control, irrigation, and also considering comprehensive utilization
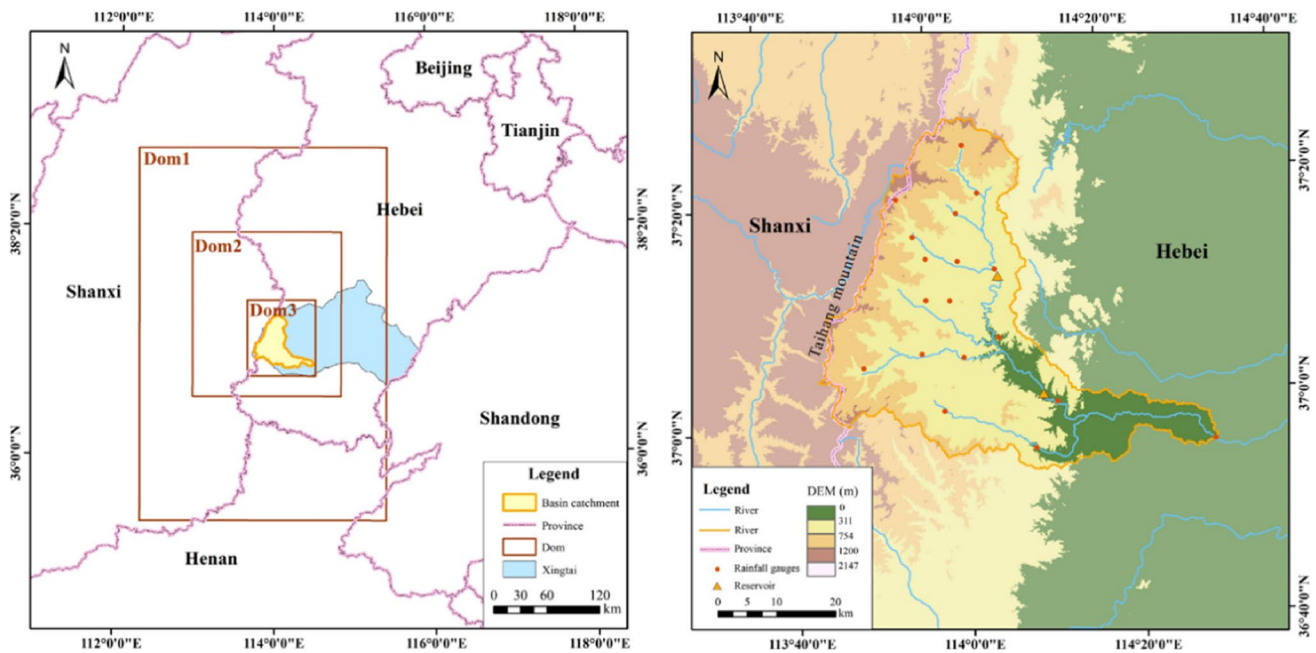
Fig. 1 Locations and DEM of study area with the gauges, reservoirs, river networks, and the settings of the WRF model nested domains

of power generation and urban water supply, with a total storage capacity of 416.2 million m^3. Another medium-sized reservoir is the Yegoumen reservoir, which is mainly used for flood control, with a total storage capacity of 50.4 million m^3. The historical maximum peak discharge of the study area reaches 6100 m^3/s. The flood rises and falls steeply, and its confluence time is short, which is particularly likely to cause personnel and property losses.

The rainfall data are collected from 18 rainfall gauges, including Dukou, Jiangshui, Luluo, Podi, and so on. Their locations are shown in Fig. 1. The runoff data of Duanzhuang hydrological station is from 1975 to 2018. Six typical rainfall events of two different magnitudes are selected for the scheme parameter optimization of the WRF model, and 12 typical flood events are used for the calibration and verification of the hydrological model. The six 48 h typical rainfall processes screened from historical precipitation data of Taihang Piedmont basin are shown in Table 1, and the Thiessen polygon method is used to calculate the areal mean rainfall (Fiedler 2003). Among them, the cumulative precipitations of event 1–event 3 are more than "100 mm," and event 4–event 6 are less than "100 mm."

2.2 Methods

2.2.1 WRF

WRF model is a unified mesoscale numerical weather prediction model, which can be used to simulate real weather cases. With the development of computer technology in

Table 1 Areal mean rainfall at different times of each storm events

Event ID	Time ID	Areal mean rainfall (mm)		
		0–24 h	24–48 h	0–48 h
1	20160721	320.86	51.48	372.34
2	20000706	211.28	44.94	256.21
3	20010726	98.00	5.53	103.52
4	20130709	49.02	30.83	79.85
5	20110701	49.40	18.03	67.43
6	20160731	42.75	3.91	46.65

recent years, it has been deeply studied and widely used. The model is mainly composed of preprocessing module WPS (WRF preprocessing system), main module ARW, and postprocessing module ARWpost. Model prediction mainly considers the resolution of 1–10 km. More details about WRF model can be found on the website (http://www2.mmm. ucar.edu/wrf/users/). The NCEP Final (FNL) is chosen as the driving data of WRF model to provide the initial and boundary conditions. The resolution of FNL is 1°×1° with a 6 h time interval. The grid center of the chosen domain is 113.86°E, 37.25°N, and three-layer grid nesting is adopted. The horizontal grid spacing of the innermost domain is set to be 1 km, and downscaling ratio is set to be 1:3 (Yang et al. 2012). To match the time step of the hydrological model, the output time interval of the climate model is 0.5 h. Taihang Piedmont basin is located in the middle latitude, so Lambert projection is selected. The physical parameterization schemes affecting the simulation of meteorological elements

in the WRF model mainly include microphysical process, cumulus parameterizations, land surface process, boundary layer scheme, longwave radiation, and shortwave radiation.

According to previous studies, land surface process, longwave radiation, and shortwave radiation mainly affect energy propagation and dissipation, while having relatively less impact on precipitation formation and water vapor transport (Evans et al. 2012). Therefore, this paper focuses on studying the impact of microphysical processes, planet boundary layer scheme, and cumulus parameterizations in the WRF model on rainfall simulation. Among them, YSU scheme and Mellor-Yamada-Janjic (MYJ) scheme are two typical planet boundary layer schemes (Janjic 1994). Microphysical schemes Purdue-Lin (Lin) and Single-Moment6 (WSM6) are two widely used schemes in the WRF model. Thompson scheme can provide the mixing ratio of water vapor, cloud water, rain, ice, snow, and graupel, which is suitable for high-resolution simulation research (Qi 2019). In the cumulus parameterizations scheme, the shallow convection KF scheme considers the influence of the basic airflow movement on the cloud microphysical process, while the Betts-Miller-Janjić (BMJ) scheme and Grell-Devinji (GD) scheme are suitable for the simulation of high-resolution and strong convective weather (Janjic 2000; Grell and Freitas 2014). These physical parameter schemes are considered to form different parameter combinations to simulate the typical rainfall events, as shown in Table 2.

2.2.2 HEC-HMS

Model introduction HEC-HMS model is a new generation of hydrological model system derived from HEC-1 combined with GIS by the hydrological engineering center of the U.S. Army Engineering Corps (Anderson et al. 2002; Gyawali and Watkins 2012; Abushandi and Merkel 2013). It is a distributed hydrological model with physical basis, which is mainly used to simulate the rainfall and runoff process of tree watershed system (Knebl et al. 2005; Feldman 2000; Fleming and Scharffenberg 2012). The model fully considers the temporal and spatial variability of the underlying surface and rainfall data. It has been proved that it can produce flood simulation results with high precision.

To build an HEC-HMS hydrological model, it is necessary to collect the watershed topography, land use data, and soil classification data. The spatial resolution of the digital elevation model adopted in this study is 30 m × 30 m. The land use data and soil data are from the Resource and Environment Science and Data Center of the Chinese Academy of Sciences (https://www.resdc.cn). According to the national soil type distribution map, the soil types in the catchment mainly include cinnamon soil and brown soil. Its main land use types are forest and grassland. According to the topographic characteristics and underlying surface

Table 2 Designed 18 combinations of physical parameterization schemes

Scheme ID	PBL	Cumulus parameterization	Microphysics
1	YSU	KF	Thompson
2	YSU	KF	Lin
3	YSU	KF	WSM6
4	YSU	GD	Thompson
5	YSU	GD	Lin
6	YSU	GD	WSM6
7	YSU	BMJ	Thompson
8	YSU	BMJ	Lin
9	YSU	BMJ	WSM6
10	MYJ	KF	Thompson
11	MYJ	KF	Lin
12	MYJ	KF	WSM6
13	MYJ	GD	Thompson
14	MYJ	GD	Lin
15	MYJ	GD	WSM6
16	MYJ	BMJ	Thompson
17	MYJ	BMJ	Lin
18	MYJ	BMJ	WSM6

conditions of the study area, the HEC-GeoHMS toolbox is used to divide the research area. After the steps of filling the depression, calculating the flow direction, generating the river network, and setting the catchment area threshold, 11 sub-basins are finally divided. The basin generalization map of the HEC-HMS model is shown in Fig. 2. Based on the hydrological characteristics of this catchment, the Soil Conservation Service (SCS)-Curve Number curve method, SCS unit hydrograph method, Muskingum method, and regression curve method are selected as the calculation methods of runoff generation, confluence, river evolution, and base flow of the model. The operation rules of the two reservoirs are considered in this model. Their basic information such as the maximum storage capacity, maximum water level, and storage capacity curve of the reservoir is set according to their conditions using the reservoir module. According to different flood events, the initial water level or initial storage capacity of the reservoir is set to represent the initial state of the reservoir when the flood comes, and the outlet curve evolution method is selected for water storage simulation. Nash efficiency coefficient (NSE) and relative error are used to evaluate the accuracy of hydrological model simulation. Generally, the initial values of the model parameters can be calculated through watershed characteristics, and the final value of the parameters can be obtained through the combination of manual and automatic calibration.

Method of hydrological model coupling There based on the optimization of the WRF parameter scheme and the

Fig. 2 Watershed generalization of HEC-HMS model
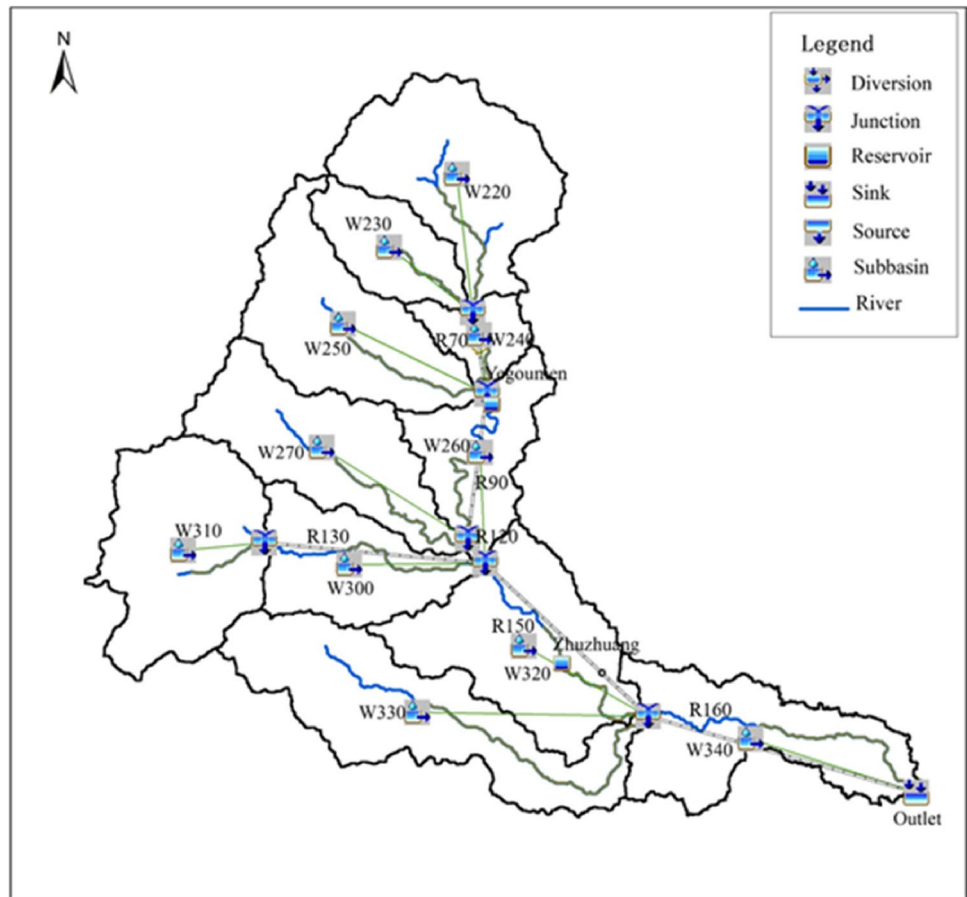
calibration of HEC-HMS parameters, the unidirectional coupling method is adopted to carry out the application of the land–atmosphere coupling model. Firstly, scale conversion should be conducted in time and space. Temporally, the output time interval of the WRF model is set to 0.5 h, which is the same as the time step size of HEC-HMS. Spatially, the forecast rainfall data at the gauge stations should be extracted from the grid rainfall data output by the climate model for matching. Based on the prediction results of the climate model, the prediction accuracy is tested by comparing the simulation performance of different rainfall data inputs with or without future predicted rainfall on typical flood events. It means that one is to use the WRF model to predict the rainfall in the next 48 h as the input of hydrological model and the other is after the release time of coupling model results, the rainfall input is zero, its output results are completely decided by the obtained rainfall before the release time.

After considering the running time of the climate model and the duration of each typical event, the coupling model carries out rolling staggered prediction, as shown in Fig. 3. For discussion purposes, the differences between the running time of different events and different physical schemes are excluded according to the actual modeling conditions, and it takes 7 h for the warmup of WRF model and prediction of the next 48 h rainfall.
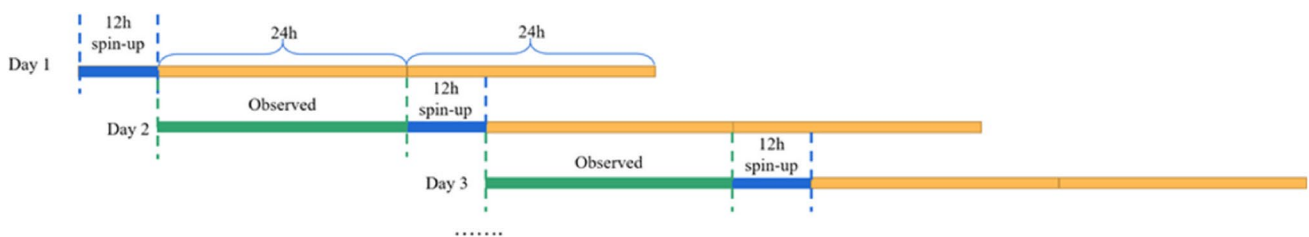

Fig. 3 Rolling forecast of flood events

2.3 Evaluation indices and methods

2.3.1 Evaluation indices of rainfall simulation

The rainfall amount is evaluated by the relative error, as shown in Eq. (1),

$$RE = \frac{|P - Q|}{Q} \times 100\% \tag{1}$$

where P is the areal mean rainfall predicted by the WRF model, and Q is the observed areal mean rainfall. For the heavy rainfall process (≥ 25 mm/day), the prediction can be rated as good with a relative error of less than 20%. Therefore, 20% is considered as the threshold of relative error to evaluate the cumulative areal rainfall in this study.

In addition, the temporal and spatial distribution of the rainfall simulation is evaluated by three-point rainfall evaluation indices including critical success index (CSI), probability of detection (POD), and false alarm ratio (FAR). These indices are constructed based on the second meteorological classification described in Table 3. To avoid taking the very small rainfall obtained in the process of solving the integral equations of WRF model into account, the conditional amount for determining the occurrence of rainfall is greater than 0.01 mm (Wang 2018). According to the physical meaning of each index calculation formula, POD reflects the detection ability of the climate model for a certain precipitation level and represents the proportion of the predicted precipitation actual area in all actual precipitation areas, with a value range of 0–1. FAR represents the proportion of the area without actual precipitation in the total predicted precipitation area, with a value range of 0–1. CSI reflects the proportion of correctly simulated rainfall frequency in all possible rainfall conditions, and it is a widely used index in meteorology.

$$POD = \frac{1}{N} \sum_{i=1}^{N} \frac{Na_i}{Na_i + Nc_i} \tag{2}$$

$$FAR = \frac{1}{N} \sum_{i=1}^{N} \frac{Nb_i}{Na_i + Nb_i} \tag{3}$$

Table 3 Rain/non-event contingency table for the simulation and observation

Simulated	Observed	
	Rain	Non-event
Rain	Na (Hit)	Nb (False)
Non-event	Nc (Miss)	Nd (None)

$$CSI = \frac{1}{N} \sum_{i=1}^{N} \frac{Na_i}{Na_i + Nb_i + Nc_i} \tag{4}$$

The calculation formulas of the three indices are shown in Eq. (2), Eq. (3), and Eq. (4). For the spatial accuracy evaluation, at time step i, the simulated rainfall is compared with the observed values at different gauges, and the quantities of Na, Nb, Nc, and Nd are counted respectively at each time step, and then the values of the indices at all time steps are averaged to obtain POD, FAR, and CSI values. When evaluating the temporal accuracy, at gauge i, the simulated rainfall is compared with observed values at different time steps, the number of Na, Nb, Nc, and Nd is counted respectively at each gauge, and then the indices of all the gauges are averaged to obtain the final values.

When optimizing various schemes, different evaluation indices can evaluate the simulation results output by WRF model from different angles and comprehensively evaluate the simulation of the six selected rainfall events. However, different indices may not come to a unified conclusion. According to the definition of the indices, the trend of CSI is consistent with POD, but the evaluation of CSI is more comprehensive than POD. In addition, FAR is the only index that indicates the false alarm of simulation, which can reflect the accuracy from the reverse side. According to the physical meaning of each index, CSI focuses on the accurate part of model prediction while FAR focuses on the false alarm.

To evaluate the performance of different physical parameterization schemes in simulating the typical rainfall events and select the best scheme accurately, a new index CSI/FAR is constructed as a comprehensive evaluation index of simulation accuracy in both time and space for each rainfall event. The higher the CSI, the higher the proportion of accurate model prediction, and the lower the FAR, the less false alarm area of model simulation. Therefore, the larger the CSI/FAR value, the better the simulation results.

Root mean square error (RMSE) quantitative evaluation index is also introduced to the point rainfall evaluation. P_j and Q_j are the simulated and observed values of rainfall, respectively. In the evaluation of spatial accuracy, they are the simulated observed cumulative areal rainfall in a specific position j with the whole observation period. When the temporal accuracy evaluation is carried out, P_j and Q_j are the simulated and observed areal mean rainfall in the study area at the observation time j, respectively.

2.3.2 A multi-attribute decision-making model based on TOPSIS optimization and grey correlation degree

TOPSIS is a multi-attribute decision-making method developed by Hwang and Yoon (1981; Zhou et al. 2014).

The main contents are as follows: firstly, the initial decision matrix ($R_{m \times n}$) is normalized to determine an ideal point and a negative ideal point (Xu et al. 2004), assuming that the value of each attribute is monotonically increasing, define the possible optimal value combination of all attributes in the scheme set as the ideal vector (R_i^+) and the possible worst value combination of all attributes as the negative ideal vector (R_i^-), then evaluate the Euclidean distance (Li et al. 2012a, b) (D_i^+, D_i^-) between each scheme and the ideal point or the negative ideal point. The scheme with the maximum relative closeness to the ideal point is selected as the optimal scheme. The grey correlation analysis method is a statistical analysis method considering multiple elements. By comparing the correlation degree of each element between two objects, we can measure the proximity of the two objects, and that is the correlation degree. This method was first applied to compare the shape similarity of two-time series. Now it is used in more and more fields (Wang 2021; Meili 2021; Liu et al. 2020). For the multi-attribute decision-making problem, the core principle of the grey correlation analysis method is to compare the grey correlation degree of each scheme relative to the ideal scheme. The greater its value, the closer the scheme to the ideal scheme. Different from the TOPSIS method, the grey correlation degree (Liu et al. 2013) (ρ_i^+, ρ_i^-) is used to construct the closeness degree.

Based on the Euclidean distance ideal solution method, when sorting and optimizing the evaluation of various parameter schemes, it only considers the distance between the data sequences and reflects the positional relationship between data curves, but cannot reflect the situation change of each data sequence. The grey correlation degree of the grey system theory can analyze the situation change well. It is a measure of the curve similarity between two groups of data series. The closer the curve shape is, the greater the correlation degree of the corresponding series is. Therefore, it can fully consider the proximity of each scheme to the ideal scheme in terms of position and shape similarity and improve the comprehensive closeness of the parameter optimization scheme by adding the grey correlation degree. The technical process of constructing the mixed closeness degree is shown in Fig. 4. A mixed closeness degree (C_i) is constructed by combining the Euclidean distance and the grey correlation degree, which fully considers the difference and correlation between the scheme and the ideal solution, making this mathematical model more scientific and reasonable.

3 Results and discussions

3.1 WRF parameter scheme optimization

The process of optimizing different physical parameterization schemes of the climate model is divided into two stages. In the first stage, six typical rainfall events are simulated according to the combination of 18 schemes including two planet boundary layer schemes, three microphysics schemes, and three cumulus parameterizations. Several combinations of physical parameterization schemes are quantitatively selected according to the relative error of areal mean rainfall. In the second stage, the accuracy of the rainfall simulation in time and space by several physical parameterization schemes optimized in the first stage will be evaluated, and the optimal

Fig. 4 Flow chart of constructing mixed closeness method
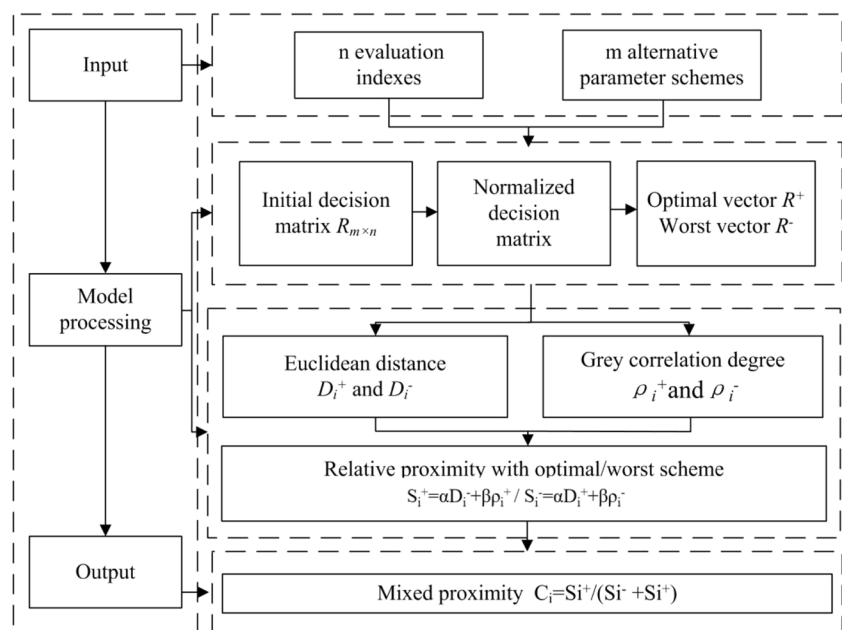
combination of the physical parameterization schemes will be selected in this stage.

3.1.1 First stage scheme optimization

The first stage of the scheme optimization is to screen the 48 h cumulative areal mean rainfall of each event. Figure 5 shows the relative error of the cumulative areal mean rainfall for each typical event under 18 parameter schemes. The principle of screening is to compare the relative error of the overall period 0–48 h in each scheme firstly, as

shown in Fig. 5a, and then analyze the relative error of two Sects. 0–24 h and 24–48 h respectively as shown in Fig. 5b,c. For the cumulus parameterization, in the overall simulation, when the KF scheme is selected, the simulation results for the cumulative 48 h of six typical precipitation processes are the most unsatisfactory, and the relative error is greater than 20% for 20 times in the 36 tests. And GD scheme is slightly better than the KF scheme with 17 times greater than 20% in the 36 tests. When the BMJ scheme is selected, the simulation effect for 48 h of the cumulative precipitation is the most ideal, and in the 36
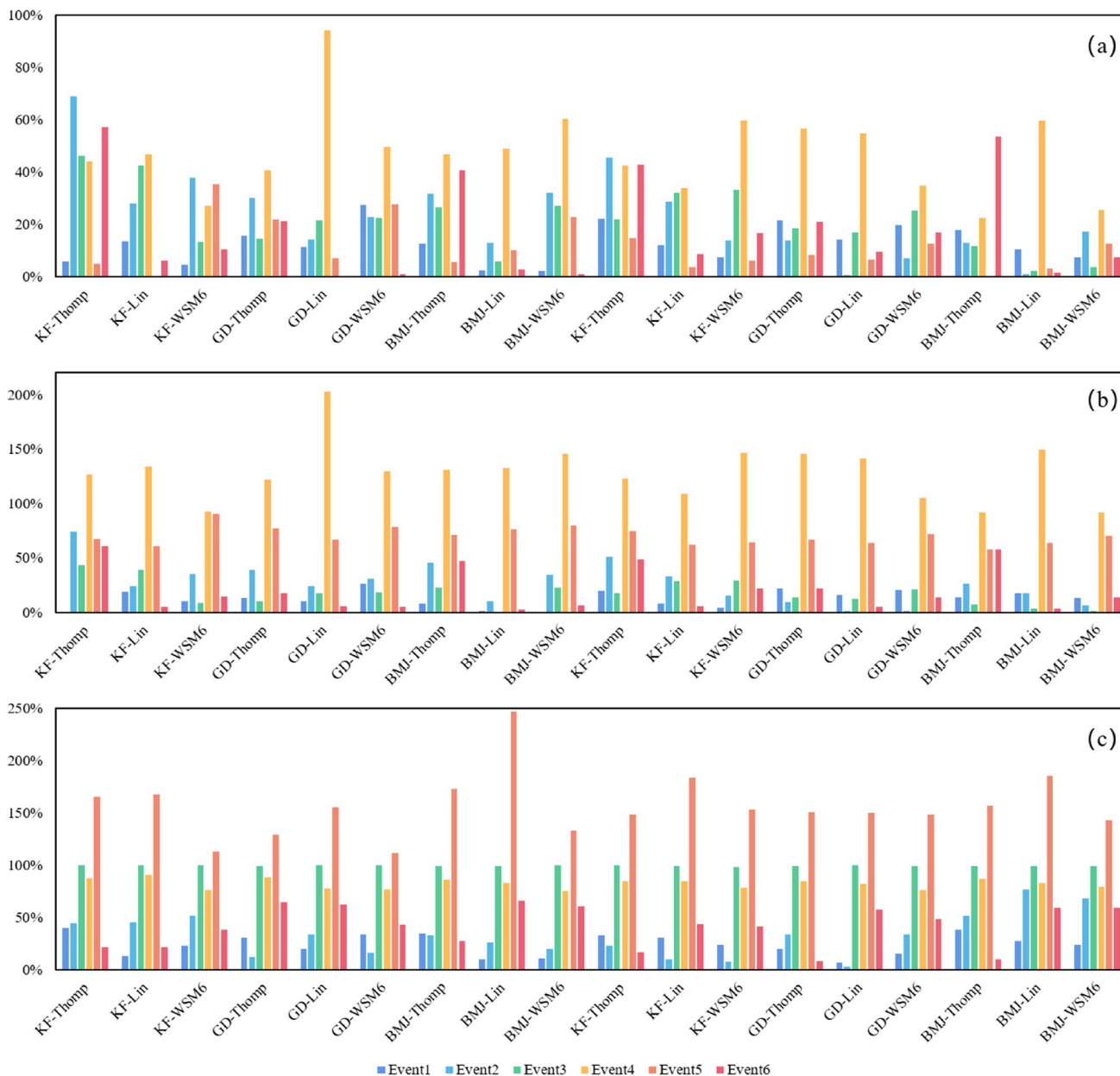

Fig. 5 Relative error of different parameterization schemes in different periods: **a** 0–48 h, **b** 0–24 h, **c** 24–48 h and planet boundary layer scheme YSU for schemes 1–9, MYJ for schemes 10–18

tests, there are 13 times that the error is greater than 20%. In the analysis of daily simulation error, the number of the relative errors greater than 20% is still the most in the KF scheme with 57 times in its 72 tests, while the performance of the GD scheme is slightly better than BMJ, and the number of relative error greater than 20% is three times less than that of BMJ scheme. For microphysical schemes, the Thompson scheme has the worst results in both the overall and two-section simulations, with 79 times greater than 20% in the 108 tests, while the Lin scheme has the best performance with 58 times greater than 20%. Among the 18 different parameter schemes, the first nine schemes apply the YSU scheme, and the other nine schemes apply the MYJ scheme. Through the statistical analysis of each scheme, it can be seen that in the overall and two-section simulations of 48 h, the number of times that the simulation results of the MYJ scheme exceed 20% is 96 in its 162 tests, while the YSU scheme is 110. Thus, the simulation effects of the MYJ scheme are better than the YSU scheme.

When considering the combination of planetary boundary layer scheme, microphysical scheme, and cumulus parameterization scheme, MYJ is selected as the boundary layer scheme. Then, when BMJ scheme is selected as the cumulus parameterization scheme, the performance of the three microphysical schemes is generally better than other combinations. When GD scheme is selected, the results of Lin scheme and WSM6 scheme are better than Thompson scheme. Therefore, according to the above results, five combination schemes of MYJ-BMJ-Lin, MYJ-BMJ-WSM6, MYJ-BMJ-Thomp, MYJ-GD-Lin, and MYJ-GD-WSM6 are selected as the optimal combinations of the first stage.

3.1.2 Second stage scheme optimization

Spatial accuracy evaluation of rainfall simulations In the first stage, five combination schemes (scheme 14–scheme 18)

have been selected through quantitative indicators. Considering the temporal and spatial accuracy of the typical rainfall simulations, point rainfall evaluation indicators are used for screening. The point rainfall evaluation indices of spatial accuracy are shown in Fig. 6. Among all indices, the POD value is highest for event 1, which reaches more than 0.95, and the POD of event 2, event 3, and event 4 also reaches more than 0.8, while the POD of some schemes for event 6 reaches 0.7, and the worst results are event 5. In this event, all the POD is less than 0.5 except the scheme 14 (GD-Lin), indicating that the overall effect of WRF model on spatial simulation of event 5 is unsatisfactory. Then the best evaluation results of FAR are event 3 and event 2. Their FAR values are less than 0.1, and other events are also relatively low except event 5. Event 5 has the highest false alarm rate, which is more than 0.5. Consistently, the CSI values are ideal for the other typical events, which are more than 0.5, except event 5. The CSI of event 5 is low, and its best result is only 0.3. In general, it can be seen from these indices that the spatial accuracy of rainfall results simulated by the climate model is good, but for some individual rainfall events, it is poor, which may be caused by different temporal and spatial distribution characteristics of rainfall.

According to the comprehensive evaluation index CSI/FAR constructed in the previous section, it can be seen from Table 4 that different typical rainfall events and different parameter combinations have different simulation performances. The simulation performance of scheme 18 (BMJ-WSM6) for event 3 ranks first, events 1, 4, and 5 rank second, and events 2 and 6 rank third, while it has the best overall performance for the simulation of the six typical rainfall events. The performance of scheme 14 (GD-Lin) for event 4 and event 6 ranks first, but for events 1 and 5 are general, and for event 3 ranks fifth. Its simulation performance in space is slightly worse than scheme 18. Schemes 17 (BMJ-Lin) and 16 (BMJ-Thomp) have good rainfall simulation performance on some individual events, but the overall simulation
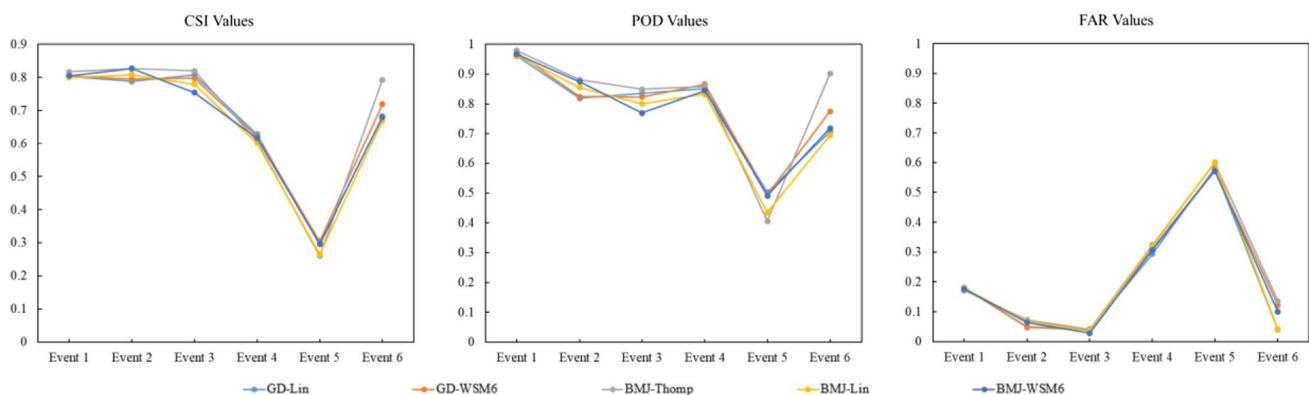

Fig. 6 Evaluation index of spatial accuracy in different events

Table 4 Ranking of CSI/FAR calculation results of spatial distribution

Ranking	Events					
	Event 1	Event 2	Event 3	Event 4	Event 5	Event 6
	20160721	20000706	20010726	20130709	20110701	20160731
1	Scheme 16	Scheme 15	Scheme 18	Scheme 14	Scheme 15	Scheme 14
	4.7608	16.6589	26.587	2.1367	0.5283	17.2327
2	Scheme 18	Scheme 14	Scheme 17	Scheme 18	Scheme 18	Scheme 17
	4.5475	16.0098	21.529	2.011	0.5179	15.9307
3	Scheme 14	Scheme 18	Scheme 15	Scheme 15	Scheme 14	Scheme 18
	4.5115	12.8005	20.4385	1.9976	0.5073	6.7917
4	Scheme 17	Scheme 17	Scheme 16	Scheme 17	Scheme 17	Scheme 15
	4.4611	11.8771	20.3776	1.8807	0.4412	5.8838
5	Scheme 15	Scheme 16	Scheme 14	Scheme 16	Scheme 16	Scheme 16
	4.4489	11.2419	19.4477	1.8741	0.4343	5.8162

performance ranking is low. In particular, scheme 16 only ranks first for event 1, but ranks fifth for events 2, 4, 5, and 6. Therefore, the performance of scheme 18 is the best in the spatial simulation, and scheme 14 ranks second.

Temporal accuracy evaluation of rainfall simulations The evaluation of the rainfall simulations in time is shown in Fig. 7. Among the evaluation indices, POD has the best result for event 1, reaching more than 0.97, indicating that the model is more accurate in the temporal simulation of event 1. At the same time, the simulation on other rainfall events in terms of POD is also relatively good, all above 0.7, which shows that the overall performance of the WRF model on the temporal simulation of rainfall is satisfactory. The lowest value of FAR is event 3, lower than 0.1. The FAR of events 1, 2, 4, and 6 are also relatively low, basically between 0.1 and 0.2. Event 5 has the highest FAR value, close to 0.4. For the CSI, event 1 has a high value. It is also good in events 2, 3, 4, and 6, which is all above 0.5. Event 5 has low CSI scores, which are all lower than 0.5.

In conclusion, according to the results of each evaluation index, it can be seen that the simulation results for event 5 are relatively poor, and other events are good in terms of temporal accuracy.

According to the results from Table 5, each scheme has different temporal accuracy for events 1–6. Scheme 14 (GD-Lin) ranks first in the simulation of event 2 and event 4, ranks third of events 1 and 3, but the overall simulation performance of the six events is the best. Scheme 18 (BMJ-WSM6) has the best simulation performance on event 1, the second for event 3 and event 5, but its temporal accuracy in simulation of the six typical events is slightly worse than scheme 14. The simulation performance of schemes 15 (GD-WSM6), 16 (BMJ-Thomp), and 17 (BMJ-Lin) on each rainfall event is not stable. Although the indices of some individual events are good, the simulation performance of other events is relatively bad. Especially scheme 17, except for the results of event 6, that of other events are all in the fourth place. In addition, for all the six typical rainfall events, scheme 15 and scheme 16 ranked last for two and three of
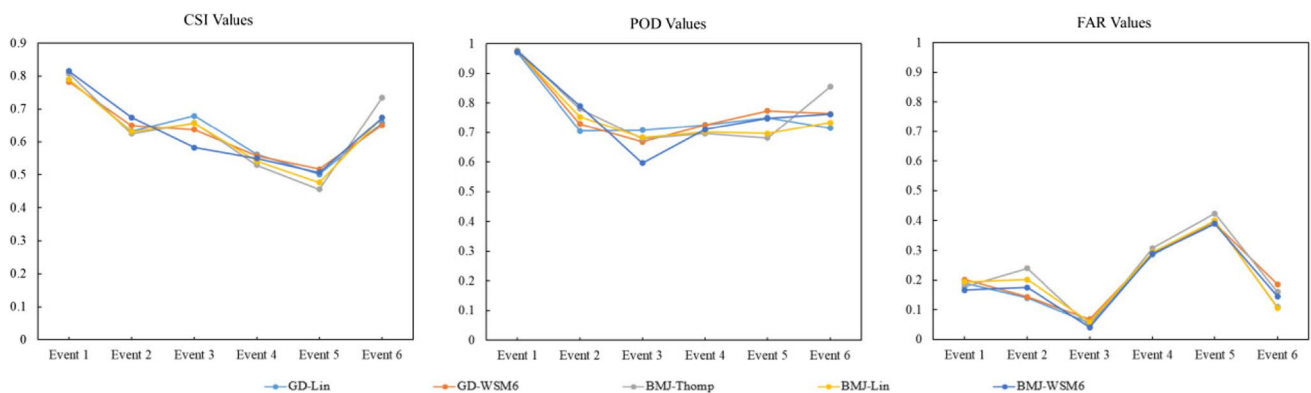

Fig. 7 Evaluation index of temporal accuracy in different events

Table 5 Ranking of CSI/FAR calculation results of temporal distribution

| Ranking | Different events | | | | | |
	Event 1 20160721	Event 2 20000706	Event 3 20010726	Event 4 20130709	Event 5 20110701	Event 6 20160731
1	Scheme 18 4.9156	Scheme 14 4.54	Scheme 16 14.4733	Scheme 14 1.9661	Scheme 15 1.3235	Scheme 17 6.3484
2	Scheme 16 4.5825	Scheme 15 4.5266	Scheme 18 14.2737	Scheme 15 1.9225	Scheme 18 1.2978	Scheme 14 6.0475
3	Scheme 14 4.1461	Scheme 18 3.8595	Scheme 14 12.1335	Scheme 18 1.8989	Scheme 14 1.2563	Scheme 18 4.643
4	Scheme 17 4.0854	Scheme 17 3.1297	Scheme 17 11.2241	Scheme 17 1.8389	Scheme 17 1.1902	Scheme 16 4.5886
5	Scheme 15 3.8892	Scheme 16 2.6077	Scheme 15 9.349	Scheme 16 1.7195	Scheme 16 1.0733	Scheme 15 3.508

them, respectively. Therefore, the simulation performance of these three schemes in the temporal accuracy is not satisfactory. To sum up, scheme 14, followed by scheme 18, has stable temporal simulation performance and better results.

3.1.3 Improvement of optimization method

In the above scheme optimization, phased optimization method is adopted. Firstly, five schemes with relatively accurate areal rainfall prediction are quantitatively selected, and then the spatio-temporal point rainfall evaluation indices are used to screen these five schemes. However, in the second stage, there is no specific quantitative analysis on the prediction error of point rainfall, and the proposed comprehensive index is only processed by simple mathematics. On this basis, it is improved by introducing the RMSE to quantitatively analyze the spatio-temporal prediction of point rainfall and constructing a multi-attribute decision-making model based on TOPSIS and grey correlation degree. The improved optimization method can evaluate each scheme more comprehensively. The temporal and spatial distribution of each typical rainfall under the optimal scheme selected by the improved method is given in Figs. 8 and 9.

By constructing the multi-attribute decision-making model, the overall result of scheme ranking is close to the original scheme ranking, which further consolidates the test results of the phased index system, as shown in Tables 6 and 7. According to the spatial and temporal accuracy evaluation, the optimal schemes selected by the two methods are scheme 18 (MYJ-BMJ-WSM6) and scheme 14 (MYJ-GD-Lin), respectively. However, with the scheme optimization method described in the previous section, both scheme 14 and scheme 18 perform better than other schemes significantly, but there is little difference between them. In the improved optimization method, it can be seen that the

advantages of the optimal scheme are more obvious. To some extent, the optimization method weakens the influence of the contingency of typical events on the scheme optimization. In the spatial accuracy evaluation, scheme 18 only ranks first in event 3. In the improved method ranking, it ranks first among three rainfall events. In the original evaluation of temporal accuracy, scheme 14 and scheme 18 are significantly better than others, while in the improved evaluation, the rankings of scheme 15 (MYJ-GD-WSM6) and scheme 18 are almost the same, and they are significantly worse than scheme 14. On this basis, according to the cumulative area rainfall of six typical events and the spatial–temporal distribution accuracy, the overall effect of the model on the simulation of large rainfall (cumulative rainfall above 100 mm) is better than that of small rainfall (cumulative rainfall below 100 mm). Among them, scheme 18 performs better in simulating rainfall below 100 mm, and scheme 14 is ideal in simulating rainfall above 100 mm.

3.2 Calibration and verification of HEC-HMS

According to the actual situation of rainstorms and floods above Duanzhuang hydrological station in Taihang Piedmont basin, 12 flood events from 1975 to 2018 are selected, including nine flood events for model calibration and three for model verification. The optimal values of the model parameters are shown in Table 8. Among the 12 storm flood events, the relative errors of nine floods are less than 20%. It can be seen from Table 9 that after each event passes the calibration, the variation range of *RE* for peak discharge is 6.49–19.57%, that of runoff depth is −19.26–14.61%, and its average NSE is 0.79. The relative error of each flood event in the verification period is less than 20%, and its average NSE is 0.80. Therefore, the calibrated parameter values are reasonable, and the model can be used to simulate the flood in the basin above Duanzhuang station.
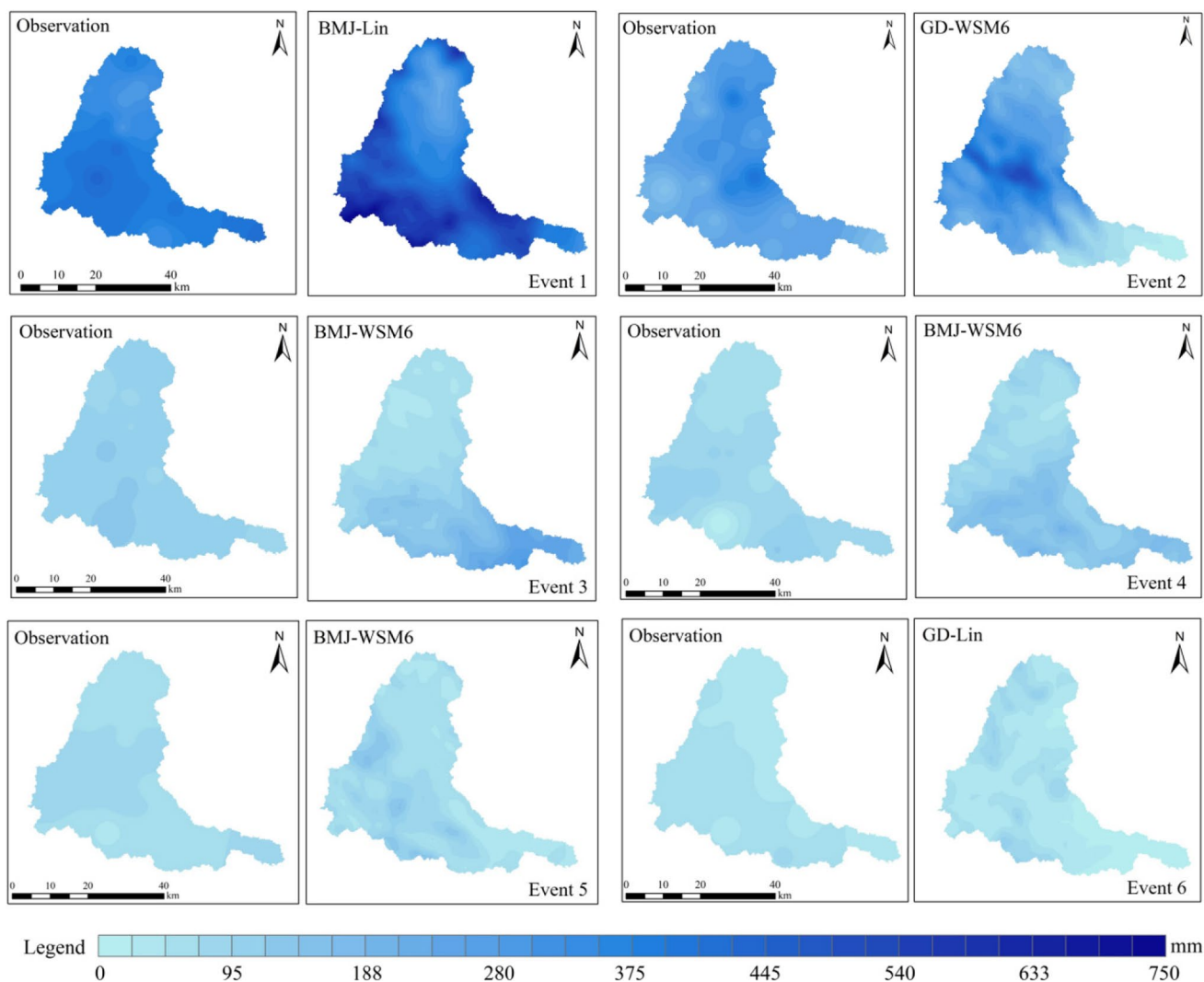
Fig. 8 Spatial distribution of observed and simulated accumulation rainfall from events 1 to 6

3.3 Simulation results of the coupling model

In this part, three typical flood events are selected for the coupling model prediction. The starting time and peak time of each flood event are given in Table 10, and the release time of coupling model results is preset according to the actual occurrence time of flood. Figure 10 shows the rolling flood forecast results of each flood event. The specific analysis is as follows:

(1) For event 20,040,813, according to the coupling model forecast results released at 8:00 on August 11, it can be seen that due to the regulation and storage of the two reservoirs, by 8:00 on August 11, a small amount of runoff formed by the observed rainfall and the future 24 h rainfall predicted by WRF model had been stored, resulting in no flood and runoff at the downstream out-

let section. After inputting the rainfall predicted by the climate model for the next 48 h, the hydrological model can capture the flood peak. The peak discharge is 14% smaller than the observed, and the peak time is 5.5 h earlier than that of the actual events, with an NSE of 0.678. The coupling model extends the flood forecast period to 43.5 h, and the overall simulation performance is good. In the results released on the 12th, the runoff simulated by the observed rainfall through the hydrological model was also stored by reservoirs. If the predicted rainfall for the next 24 and 48 h is input into the hydrological model, the flood can be simulated. Its peak flow is 30% smaller than the observed, NSE is 0.858, and the peak time is 1 h ahead of an observed flood event. It can be seen that with the progress of rolling forecast, the coupling model can simulate the flood form and peak time more accurately, but the simulation
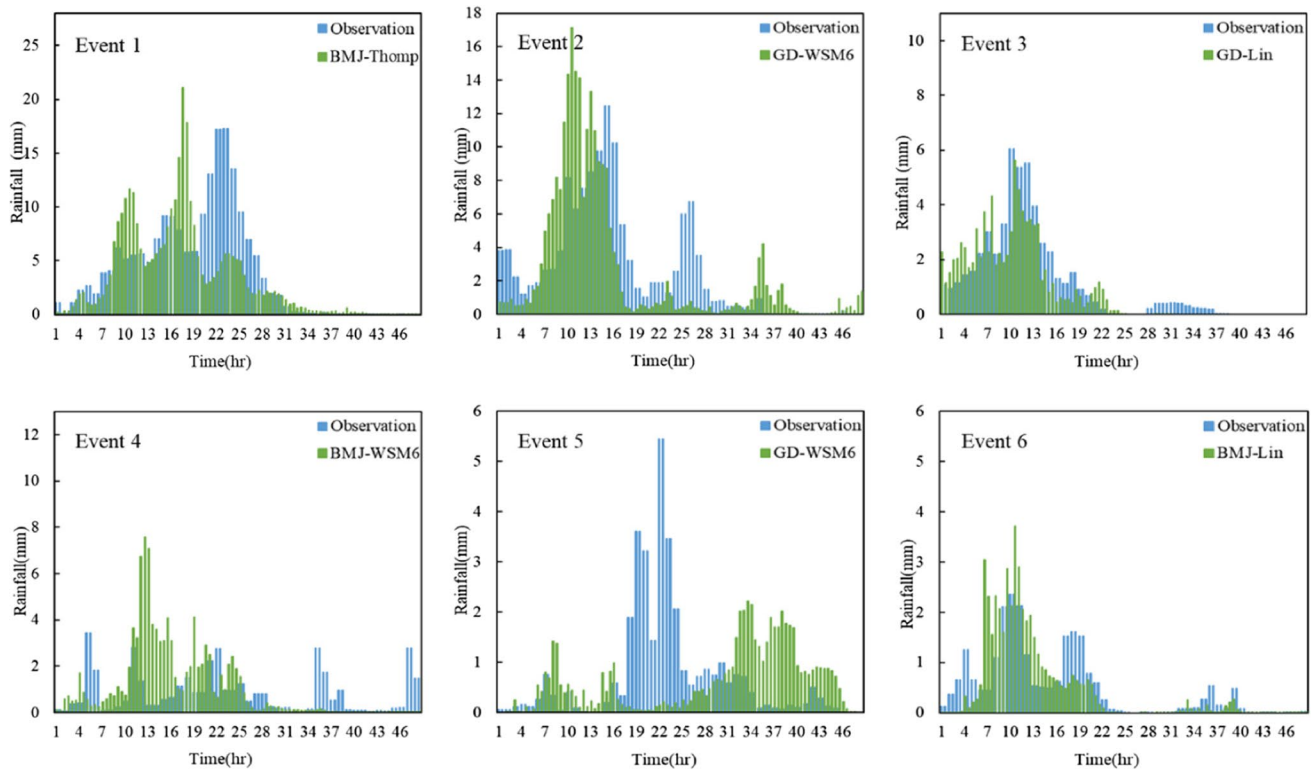
Fig. 9 48 h time distribution of observed and simulated values of the optimal scheme for each event

Table 6 Spatial values of C_i

Ranking	Event 1 20160721	Event 2 20000706	Event 3 20010726	Event 4 20130709	Event 5 20110701	Event 6 20160731
1	Scheme 17 0.6858	Scheme 15 0.7463	Scheme 18 0.709	Scheme 18 0.7115	Scheme 18 0.6693	Scheme 14 0.7025
2	Scheme 18 0.6761	Scheme 14 0.7183	Scheme 16 0.3704	Scheme 15 0.6915	Scheme 14 0.5979	Scheme 17 0.6455
3	Scheme 16 0.3793	Scheme 17 0.4622	Scheme 17 0.3563	Scheme 16 0.5699	Scheme 15 0.496	Scheme 15 0.414
4	Scheme 14 0.3434	Scheme 18 0.3848	Scheme 15 0.346	Scheme 14 0.4967	Scheme 16 0.4872	Scheme 18 0.3984
5	Scheme 15 0.2781	Scheme 16 0.2591	Scheme 14 0.2779	Scheme 17 0.259	Scheme 17 0.3646	Scheme 16 0.3037

performance of flood peak discharge is not ideal due to the error of the climate model prediction. In the simulation results on the 13th, the main rainfall forming the flood has occurred. According to the observed rainfall, a good flood form can be simulated. The peak flow is 14.4% smaller than the observed, NSE is 0.726, and the peak time is 3 h ahead of the actual event, while the 24 h rainfall predicted by the WRF model only affects the water runoff and reduces the runoff relative error.

(2) According to the prediction results released on July 19 for event 20,160,721, it can be seen that the observed

rainfall failed to simulate the formation of a complete flood process by the hydrological model while putting the 24 h WRF model prediction results into the model can capture the flood peak, and the peak flow is 37% smaller than the measured, the peak time lags 1.5 h, and NSE is 0.599. This flood event is affected by the artificial regulation of reservoir, so its shape is complex. The hydrological model has a large error, and the error of peak flow is 16.21% smaller than that of actual flood events. The error is further increased due to the temporal and spatial difference of rainfall prediction caused

Table 7 Temporal values of C_i

Ranking	Event 1 20160721	Event 2 20000706	Event 3 20010726	Event 4 20130709	Event 5 20110701	Event 6 20160731
1	Scheme 16 0.6678	Scheme 15 0.7387	Scheme 14 0.6431	Scheme 18 0.7064	Scheme 15 0.7402	Scheme 17 0.6817
2	Scheme 14 0.6203	Scheme 14 0.723	Scheme 16 0.4604	Scheme 15 0.4939	Scheme 18 0.7253	Scheme 14 0.5955
3	Scheme 18 0.5585	Scheme 18 0.3967	Scheme 18 0.4237	Scheme 16 0.43	Scheme 14 0.6556	Scheme 15 0.4938
4	Scheme 15 0.5241	Scheme 16 0.3036	Scheme 17 0.4027	Scheme 14 0.3742	Scheme 17 0.2764	Scheme 18 0.3928
5	Scheme 17 0.2891	Scheme 17 0.28	Scheme 15 0.3038	Scheme 17 0.2779	Scheme 16 0.232	Scheme 16 0.3179

Table 8 Optimal parameters of HEC-HMS model

Sub-basin	Area (km²)	Curve number	Recession constant	Lag time (min)	Reach	Muskingum K	Muskingum x
W220	193.93	46.06	0.34	92.00	R70	3.00	0.20
W230	78.59	42.34	0.30	278.00	R90	4.80	0.20
W250	172.91	37.00	0.35	150.00	R120	5.00	0.20
W240	43.00	60.06	0.36	84.14	R130	1.61	0.20
W270	168.52	45.06	0.26	184.16	R150	2.23	0.20
W310	200.92	37.00	0.57	220.00	R160	2.00	0.20
W300	123.37	35.00	0.30	192.00			
W260	104.15	42.07	0.25	105.92			
W330	230.99	59.41	0.53	605.81			
W320	200.90	68.17	0.32	250.00			
W340	168.58	50.00	0.20	258.26			

Table 9 Calibration and verification of each flood event

Period	Events	Peak discharge			Runoff			NSE
		Observed (m³/s)	Simulated (m³/s)	Relative error (%)	Observed (mm)	Simulated (mm)	Relative error (%)	
Calibration period	19960721	6100	5123	16.01%	256.28	304.24	−18.71%	0.87
	19750808	402	376	6.49%	68.71	59.92	12.79%	0.88
	19770730	209	191	8.66%	39.08	33.37	14.61%	0.59
	20000722	173	139	19.57%	3.8	4.43	−16.58%	0.76
	19760819	172	144	16.51%	2.8	3.29	−17.50%	0.79
	19820803	170	140	17.65%	2.44	2.91	−19.26%	0.86
	19820812	129	105	18.68%	1.71	2.02	−18.13%	0.81
Validation period	20000706	365	426	−16.77%	11.83	13.05	−10.31%	0.85
	20040813	117	101	13.85%	6.69	5.75	14.05%	0.74

by the climate model. Compared with 24 h prediction results, the results of 48 h have no difference in the simulation of peak discharge and peak time, and the NSE is 0.608. Because the rainfall forming the main peak has been predicted in the first 24 h, and a small amount of rainfall in the next 24 h has little impact on the overall flood simulation, it only influences the runoff relative error, increasing it to 2.5%. It can be seen from the

Table 10 Description of flood events

Flood events	Starting time of HEC-HMS	Peak time	Starting time of WRF model	Release time of coupled model forecast
20040813	2004/08/10/12:00	2004/08/13/09:00	2004/08/10/12:00	2004/08/11/08:00
				2004/08/12/08:00
				2004/08/13/08:00
20160721	2016/07/19/12:00	2016/07/20/20:30	2016/07/19/12:00	2016/07/19/18:00
				2016/07/20/18:00
				2016/07/21/18:00
20000706	2000/07/05/00:00	2000/07/06/14:00	2000/07/05/00:00	2000/07/05/08:00
				2000/07/06/08:00
				2000/07/07/08:00
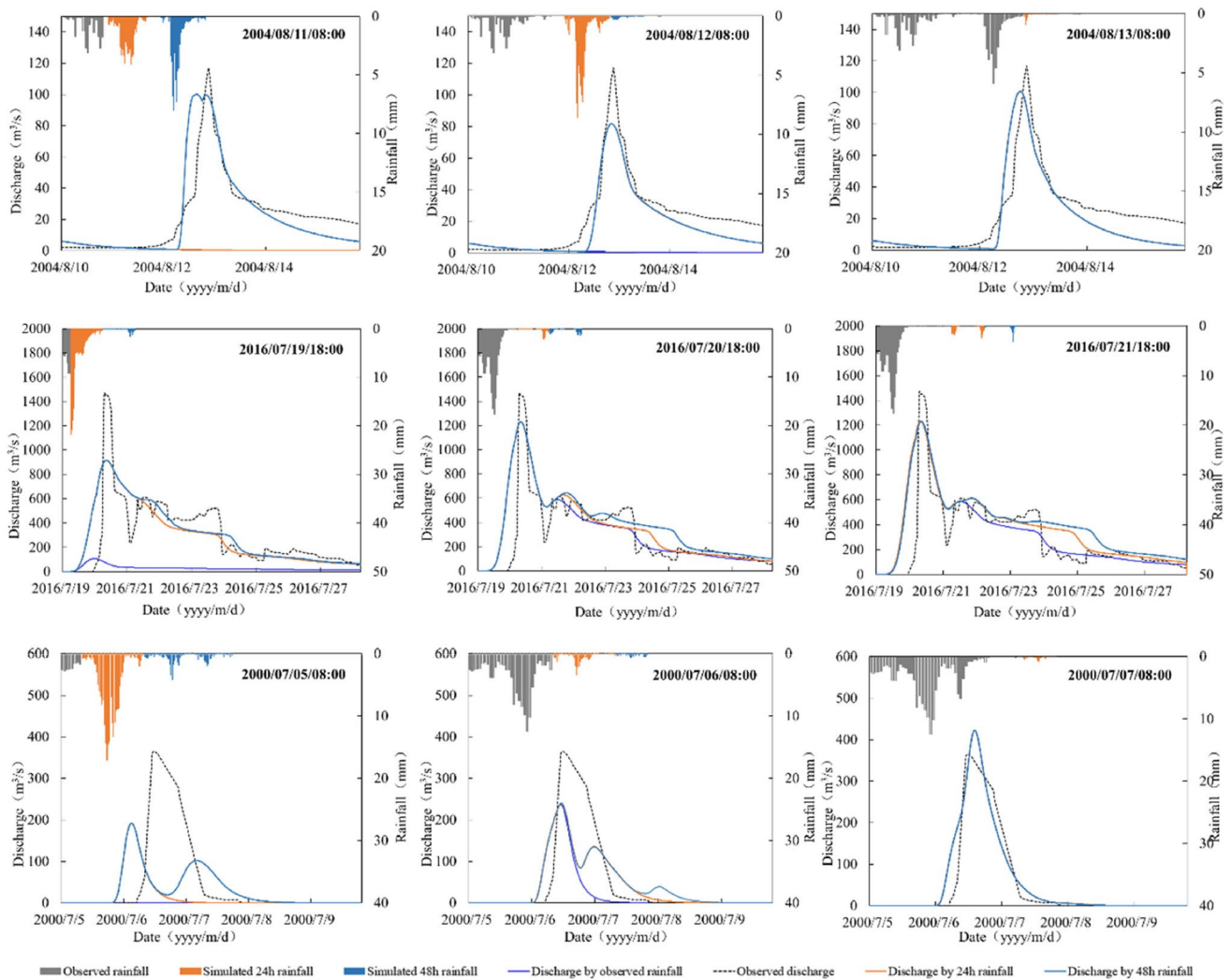
Fig. 10 Coupled model cycling flood forecasting process and the results of coupling model

forecast results released on the 20th the main rainfall forming flood runoff has occurred, so the forecast of the climate model only has an impact on the runoff and regression process in the later stage. When inputting the observed rainfall sequence, the simulated peak discharge is 16.21% smaller than the actual, its peak time is 1.5 h later, and the NSE is 0.543. When using the 24 h and 48 h rainfall prediction results, the simulated peak flow and peak time are consistent with the observed rainfall series, the NSE are 0.527 and 0.467, respectively. Due to the large error of the hydrological model in the runoff simulation of this flood event, with the input of 24 h and 48 h rainfall prediction, the runoff simulation error increases and the NSE decreases.

(3) For event 20,000,706, the prediction performance of the coupling model is unsatisfactory, it does not capture the flood peak, mainly due to the error of the climate model. On the other hand, the short-term flood event requires higher simulation accuracy of the climate model in time and space. Especially in the time dimension, the duration of the flood is short. Based on the existing observed rainfall, if there is a deviation in time prediction, it will lead to the failure to form a centralized rainfall sequence, resulting in inaccurate simulation of the flood peak by the coupling model.

From the error of rolling forecast simulation results, it can be seen that the flood peak form of event 20,040,813 is great, its incoming water is concentrated. The accuracy of hydrological model simulation is high, which provides a good land surface basis for coupled model flood forecasting. The numerical weather model's error has become the key factor affecting flood forecasting. Moreover, the errors caused by climate model in the prediction of rainfall falling area and main rainfall peak will directly lead to the errors of flood discharge and peak time. Due to the great influence of artificial reservoir regulation, the flood form of event 20,160,721 is complex, and there are large errors in hydrological model simulation. Event 20,000,706 belongs to the flood caused by short-term heavy rainfall. This flood event is formed in a relatively short time, and the simulation effect is the least ideal in three events. The low simulation results of each typical flood event are caused by the parameters of hydrological model, which also makes the prediction results dangerous. Through the simulation results of typical events in this basin, it can also be seen that when establishing the hydrological model in the basin with large-scale water conservancy projects, the interference factors of manual regulation cannot be ignored. The parameter setting of reservoir calculation module should be more flexible and fit with reality, it should also be taken into account in the calibration and verification of model parameters. Based on hydrological model, for the error caused by numerical weather forecast, it

is also necessary to combine the synoptic theory and practical experience of forecast, effectively correct the precipitation forecast results, comprehensively analyze and construct ensemble forecast, draw more accurate forecast conclusions, and further improve the accuracy of flood forecast.

4 Discussions

According to the six typical rainfall events, the rainfall parameter combination MYJ-GD-Lin, MYJ-BMJ-WSM6 is selected through two stages and multiple indicators. However, in addition to the scheme of parameter combinations, we can also see the impact of single parameter in the WRF model. For the two planet boundary schemes, the MYJ scheme is better than YSU in this study area. The MYJ scheme uses the turbulence closure method to calculate the turbulent motion of the boundary layer, which is suitable for the stable boundary layer. Wang et al. (2014) used four different boundary layer schemes in the WRF model to simulate the southwest wind flow at the middle or low level in eastern China and the Western Pacific subtropical high, and for the precipitation caused by them, the simulation performance of MYJ is significantly better than that of YSU. This is also consistent with the performance of boundary layer parameter scheme evaluation in this study. For cumulus parameterizations, previous studies have shown that BMJ and GD schemes are suitable for the simulation of high-resolution and strong convective weather, which is in line with the climate characteristics of the study area, but KF scheme does not perform well in this area. The KF simply considers the effects of water vapor uplift and subsidence, including entrainment and roll-out as well as the rise and sink of airflow (Kain 2004). Some studies have shown that KF scheme has a good performance in other watersheds especially for the simulation of rainfall intensity and rainfall falling area (Tian et al. 2017; Wang et al. 2021). In this study, the parameter combination of KF scheme can also better simulate the precipitation of watersheds in individual scheme combinations, but the overall performance is not stable. It is proved that KF scheme can simulate a certain rainfall or a certain type of rainfall, but its applicability in the study area is not high. Therefore, this parameter scheme is not recommended in the further construction of rainfall ensemble forecasts in this study area. For microphysical schemes, the Lin and WSM6 schemes are widely used and their simulation performances in China are generally good (Niu 2007; Zhang et al. 2013; Yan 2007). As a two-dimensional cloud rain model (Lin et al. 1983), Lin scheme not only considers water vapor, rain, and cloud water but also considers water in phase states such as snow, cloud, and ice. WSM6 is similar to Lin, which still includes the conversion between complex phase states and adds graupel treatment. It is the most complex scheme of

WSM series. Thompson scheme is the most complex cloud microphysics scheme at present. Its process is complex and suitable for high-resolution research. In the rainfall simulation of the study area, Thomp's simulation performance is acceptable under some scheme combinations for the rainfall events above 100 mm, but its evaluation performances are generally not ideal below 100 mm. At the same time, considering its low calculation efficiency and higher hardware conditions, it is not recommended to be used in the scheme screening and not recommended to further construct the rainfall ensemble forecast.

Based on this study, to further improve the model prediction accuracy of this area, more studies can be carried out from the following aspects. Firstly, when selecting the initial and boundary conditions of the climate model, according to previous studies, the FNL of NCEP data (Cheng et al. 2019; Li et al. 2012a, b) is applied to the mesoscale climate model for rainfall simulation, and their results are satisfactory. Therefore, this dataset is selected in this study and optimized for the parameters of the WRF climate model. It is worth noting that this dataset began in 1999, but under the influence of global warming, the precipitation in Xingtai area showed an obvious decreasing trend from 1954 to 2010 (Cheng 2012), and at the same time, the precipitation in recent years is not abundant. However, due to the influence of the Western Pacific subtropical high and the annual Summer Typhoon on the air movement, there is still the possibility of occurring serious flood disasters in this area. Due to these reasons, the selection of typical rainfall in the basin is limited to a certain extent. Therefore, to improve the rainfall prediction ability of WRF model in operational prediction, we can try to replace a series of longer datasets and increase the sample base of rainfall event selection as much as possible in further research, to reduce the error caused by the uncertainty of typical event selection. In addition, when evaluating the rainfall predicted by the climate model, the observation value of the rainfall station is selected, which will be limited by the number and spatial distribution of rainfall gauges in the basin and cannot reflect the spatial distribution characteristics of rainfall well. In further research, the observed value of the evaluation standard can be more reliable by integrating the remote sensing data with the observed station data.

5 Conclusions

In this paper, the area above Duanzhuang hydrological station in Taihang Piedmont basin is selected as the research area. HEC-HMS is driven by the rainfall data output from the WRF model. In this way, unidirectional coupling of atmospheric and hydrological models is established to realize the real-time flood forecasting of the study area. NCEP Final (FNL) is used as the initial and boundary conditions of WRF model to simulate six typical rainfall events. By combining different microphysics, planet boundary layers, and cumulus parameterizations, 18 scheme combinations are designed to optimize the parameter schemes and the most suitable scheme combination is selected for this basin. The screening method by two stages is adopted in this study. In the first stage, taking the relative error as the quantitative evaluation index, five parameter schemes of MYJ-BMJ-Lin, MYJ-BMJ-WSM6, MYJ-BMJ-Thomp, MYJ-GD-Lin, and MYJ-GD-WSM6, which are relatively accurate for rainfall prediction, are selected. In the second stage, based on those five schemes, CSI/FAR comprehensive indices are constructed from temporal and spatial accuracy to screen the optimal scheme. In addition, by adding the quantitative index RMSE and building a multi-attribute decision-making model based on TOPSIS and grey correlation degree, the scheme optimization method is further improved. Finally, the parameter scheme MYJ-BMJ-WSM6 is selected by calculating the mixed closeness degree, which performs well in the simulation of spatial accuracy. The parameter scheme MYJ-GD-Lin is more accurate in the simulation of temporal for each typical event. Scheme 18 performs well in the simulation of rainfall below 100 mm, while scheme 14 has better performance in the simulation of rainfall above 100 mm. Based on this conclusion, in the rolling forecast, scheme 18 is adopted for the rainfall simulation of flood event 1, and scheme 14 is adopted for the rainfall simulation of flood events 2 and 3.

By collecting the rainstorm and flood data, hydrological simulation is carried out in the study area using HEC-HMS model and obtained a satisfactory result, which shows that the model has good applicability in this area. Then the rainfall output by WRF model after scale conversion is used to drive the hydrological model to carry out the real-time rolling prediction of the discharge at the outlet station. Taking three typical flood events as examples, by comparing the coupling prediction results of obtained rainfall and the future rainfall simulated by climate model respectively input into the hydrological model, it can be found that the coupling prediction method shows certain advantages. It can prolong the flood prediction period and its accuracy can be accepted. Analyzing the coupling model simulation prediction error of different typical fields and the verification of rainfall prediction, it can be concluded that the coupling the coupling model has its advantages. The error of the combined model mainly comes from two aspects. Firstly, due to the influence of water conservancy projects in the basin, the reservoir has a significant regulating effect, resulting in a significant impact of downstream rainfall station data on flood runoff at the outlet station. Secondly, the simulation of climate models on rainfall forecasting has shown good overall simulation results in large-scale

rainfall events, but in short duration heavy rainfall events. The storage of reservoirs within a watershed will further expand the impact of spatio-temporal distribution errors in climate models on coupled models. In the later stage, real-time updates of the initial field of the climate model can be continued through assimilation of radar or remote sensing data to improve the prediction accuracy of the climate model, enabling the coupled model to achieve better prediction results. At the same time, in this study, the grid rainfall output from WRF is converted into station values, and then the Tyson polygon method is used to calculate the average value of the watershed, which may have some errors. In the future, other rainfall calculation methods such as mod clark simulation of grid rainfall can be tried, which may further improve the accuracy of forecasting. In addition, due to the limitations of the original dataset NCEP FNL and the actual climate change in this watershed, the selection of typical rainfall events in this study has been affected. The next step can also be to increase the sample base of rainfall events as much as possible through a series of longer datasets, in order to reduce errors caused by the uncertainty of selecting typical events.

Data availability The data that support the fundings of this study are available from the corresponding author upon request.

Code availability Not applicable.

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

Abushandi E, Merkel B (2013) Modelling rainfall runoff relations using HEC-HMS and IHACRES for a single rain event in an arid region of Jordan. Water Resour Manage 27(7):2391–2409

Anderson M, Chen Z, Kavvas M et al (2002) Coupling HEC-HMS with atmospheric models for prediction of watershed runoff. J Hydrol Eng 7(4):312–318

Cheng X, Zhang Q (2012) Analysis of climate change trend in Xingtai City in recent 57 years. J Anhui Agr Sci 40(02):978–980

Cheng H, Wang Y, Wen B et al (2019) Reliability analysis of rain and snow weather diagnosis in Kelan area, Shanxi Province Based on NCEP FNL data. J Meteorol Environ 35(02):23–31

Ding Y, Li H, Shao A (2019) Simulation of 2m temperature and precipitation in Qinba Mountain area based on different parametric scheme combinations of WRF model. Gansu Sci Technol 35(09):39–43+46

Evans JP, Ekström M, Ji F (2012) Evaluating the performance of a WRF physics ensemble over South-East Australia. Clim Dyn 39(6):1241–1258

Feldman AD (2000) Hydrologic modeling system HEC-HMS: technical reference manual. US Army Corps of Engineers, Hydrologic Engineering Center

Fiedler F (2003) Simple, practical method for determining station weights using Thiessen polygons and isohyetal maps. J Hydrol Eng 4(8):219–221

Fleming M, Scharffenberg W (2012) Hydrologic Modeling System (HEC-HMS): new features for urban hydrology, USACE Hydrologic Engineering Center, Davis

Flesch TK, Reuter GW (2012) WRF model simulation of two Alberta flooding events and the impact of topography. J Hydrometeorol 13(2):695–708

Grell GA, Freitas SR (2014) A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. Atmos Chem Phys 14(10):5233–5250

Gyawali R, Watkins DW (2012) Continuous hydrologic modeling of snow-affected watersheds in the Great Lakes basin using HEC-HMS. J Hydrol Eng 18(1):29–39

Hwang C, Yoon K (1981) Multiple attribute decision making. Springer, Berlin, Heidelberg, p 269

Janjic Z (1994) The step-mountain eta coordinate model: further developments of the convection, viscous sublayer, and turbulence closure schemes. Mon Wea Rev 122(5):927–945

Janjic Z (2000) Comments on development and evaluation of a convection scheme for use in climate models. J Atmos Sci 57(21):3686–3686

Jasper K, Kaufmann P (2003) Coupled runoff simulations as validation tools for atmospheric models at the regional scale. Q J R Meteorol Soc 129(588):673–692

Kain J (2004) The Kain-Fritsch convective parameterization: an update. J Appl Meteorol 43(1):170–181

Knebl MR, Yang ZL, Hutchison K et al (2005) Regional scale flood modeling using NEXRAD rainfall, GIS, and HEC-HMS/RAS: a case study for the San Antonio River Basin Summer 2002 storm event. J Environ Manage 75(4):325–336

Li L, Jin J, Zhu Y (2012) Discussion on some problems in the application of the TOPSIS method. Water Resour Power 30(3):51–54

Li A, He H, Zhang Y (2012) Numerical simulation of the influence of land surface parameter disturbance of WRF model on a rainstorm in Northwest China. Plateau Meteorol 31(1):65–75

Li J, Chen Y, Wang H et al (2017) Extending flood forecasting lead time in a large watershed by coupling WRF QPF with a distributed hydrological model. Hydrol Earth Syst Sci 21(2):1279–1294

Lin Y, Farley RD, Orville HD (1983) Bulk parameterization of the snow field in a cloud model. J Climate Appl Meteor 22(6):1065–1092

Lin C, Wen L, Lu G et al (2006) Atmospheric-hydrological modeling of severe precipitation and floods in the Huaihe River Basin. Chin J Hydrol 330(1–2):249–259

Liu S, Cai H, Yang Y et al (2013) Advance in grey incidence analysis modeling. Syst Eng-Theor Prac 33(08):2041–2046

Liu H, Wu J, Chen X et al (2020) Analysis of grey correlation degree between water resources system and socio-economic system. Trop Geomorphol 41(01):31–36

Meili J (2021) Sustainable agriculture and food production in Qinghai: analysis based on grey correlation model. IOP Conf Ser Earth Environ Sci 831(1):12040

Niu J (2007) Influence of different microphysical parameterization schemes on the forecast of heavy rainfall, M.S.thesis, Academy of Meteorological Science, China, p 87 (in Chinese)

Pennelly C, Reuter G, Flesch T (2014) Verification of the WRF model for simulating heavy precipitation in Alberta. Atmos Res 135–136:172–192

Qi P (2019) Study on microphysical characteristics and precipitation formation mechanism of cumulus mixed cloud in the eastern foot of Taihang Mountain, M.S. thesis, Chinese Academy of Meteorological Science, China, p 69 (in Chinese)

Tang C, Dennis RL (2014) How reliable is the offline linkage of Weather Research & Forecasting model (WRF) and variable infiltration capacity (VIC) model? Global Planet Change 116:1–9

Tian J, Liu J, Wang J et al (2017) A spatio-temporal evaluation of the WRF physical parameterisations for numerical rainfall simulation in semi-humid and semi-arid catchments of Northern China. Atmos Res 191:141–155

Vincendon B, Ducrocq V, Dierer S et al (2009) Flash flood forecasting within the PREVIEW project: value of high-resolution hydrometeorological coupled forecast. Meteorol Atmos Phys 103(1–4):115–125

Wang C, Wang B (2021) Analysis of the influencing factors of the zero drift of capacitive acceleration sensor based on grey correlation degree. Int Core J Eng 8(7):191–202

Wang Z, Duan A, Wu G (2014) Effects of boundary layer parameterization scheme and air-sea coupling on WRF simulation of East Asian summer monsoon. Sci China Earth Sci 44(03):548–562

Wang R, Qiao F, Ding Y et al (2021) Impact of cumulus parameterization schemes with multigrid nesting on the high-resolution simulation of an extreme heavy rainfall event in Chongming. Shanghai Clim Environ Res 26(01):58–74

Wang Y (2018) Research on coupled atmosphere-hydrologic modeling for hydrologic simulation in watersheds based on different grids sizes, Ph.D. thesis, China Institute of Water Resources and Hydropower Research, p 158 (in Chinese)

Xu Y, Zhang W, Liu Y (2004) Research on standardization of decision matrix in multi-attribute decision making. In: China Youth Conference of operations research and management scholars, Qinhuangdao, China (in Chinese)

Yan Z, Deng LT (2007) Description of microphysical processes in WRF model and its prediction experiment. Desert Oasis Meteorol 06:1–6

Yang B, Zhang Y, Qian Y (2012) Simulation of urban climate with high-resolution WRF model: a case study in Nanjing. Chin Asia-Pac J Atmos Sci 48(3):227–241

Yu Z, Lakhtakia MN, Yarnal B et al (1999) Simulating the river-basin response to atmospheric forcing by linking a mesoscale meteorological model and hydrologic model system. J Hydrol 218(1–2):72–91

Zhang S, Sun K, Zhang H et al (2013) Impact from the microphysical scheme of WRF model on the simulation of a heavy rainfall process in Shanxi Province. Water Resour Hydropower Eng 44(9):8–11

Zhou J, Ouyang S, Wang X et al (2014) Multi-objective parameter calibration and multi-attribute decision-making: an application to conceptual hydrological model calibration. Water Resour Manage 28(3):767–783