Research Article

Uncertainty of Flood Forecasting Based on Radar Rainfall Data Assimilation

Xinchi Chen,[1,2] **Liping Zhang,**[1,3] **Christopher James Gippel,**[4]
Lijie Shan,[1,2] **Shaodan Chen,**[1,2] **and Wei Yang**[1,2]

[1]*State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China*
[2]*Hubei Collaborative Innovation Center for Water Resources Security, Wuhan University, Wuhan 430072, China*
[3]*College of Tourism Culture and Geographical Science, Huanggang Normal University, Huanggang 438000, China*
[4]*Australian Rivers Institute, Griffith University, Nathan, QLD 4111, Australia*

Correspondence should be addressed to Liping Zhang; zhanglp@whu.edu.cn

Precipitation is the core data input to hydrological forecasting. The uncertainty in precipitation forecast data can lead to poor performance of predictive hydrological models. Radar-based precipitation measurement offers advantages over ground-based measurement in the quantitative estimation of temporal and spatial aspects of precipitation, but errors inherent in this method will still act to reduce the performance. Using data from White Lotus River of Hubei Province, China, five methods were used to assimilate radar rainfall data transformed from the classified Z-R relationship, and the postassimilation data were compared with precipitation measured by rain gauges. The five sets of assimilated rainfall data were then used as input to the Xinanjiang model. The effect of precipitation data input error on runoff simulation was analyzed quantitatively by disturbing the input data using the Breeding of Growing Modes method. The results of practical application demonstrated that the statistical weight integration and variational assimilation methods were superior. The corresponding performance in flood hydrograph prediction was also better using the statistical weight integration and variational methods compared to the others. It was found that the errors of radar rainfall data disturbed by the Breeding of Growing Modes had a tendency to accumulate through the hydrological model.

1. Introduction

Flooding is a hazard with potentially serious consequences of loss of life, displacement of communities, and economic costs. The risks can be reduced by issuing warnings and implementing planned responses to avoid the threat. This relies on access to accurate real-time forecasts of flood peak magnitude and timing and duration at the time-scales of nowcasting (~2 hours) and short-range forecasting (~1-2 days). The time required for rainfall to be converted into runoff and be conveyed downstream through channels is sufficient that forecasting of flash floods can use recently observed rainfall data rather than less accurate forecast rainfall. Thus, accurate rainfall data is the key input to flood forecasting [1]. Such data is usually obtained from an automatic network of rain gauges, set up according to established standards to continuously measure rainfall volume at short

intervals with an acceptable accuracy. For practical reasons, the density of gauges in the network will often be too low to properly characterize the spatial and temporal distribution of rainfall throughout a catchment. Failure to capture data concerning the intensity of the main storm cell can lead to underestimation of flood peak magnitude and timing. Weather radar offers an alternative method of capturing rainfall data that has the potential to overcome this limitation of a rain gauge network or to enhance the information from the network. Radar echoes provide real-time, high spatial resolution information on cloud extent and thickness, evolution and direction of travel of storm cells, and precipitation distribution, but with a degree of uncertainty that could be considered less than desirable for reliable flood forecasting. The accuracy of radar-derived rainfall data can be improved by calibration with observed data using data assimilation techniques [2–4]. Data assimilation refers to

the process of incorporating observed data into the model state of a numerical model of a system and is employed in situations when the number of available observations of the system is much smaller than the number required to specify the model state. Postassimilation rainfall data can be input to distributed hydrological models to more accurately predict flood hydrograph characteristics [5–9].

There has been considerable progress in practical application of weather radar for estimation of surface rainfall. For example, the UK Met Office uses a network of 13 weather radars that provides 2 km × 2 km gridded radar echo data at 5-minute intervals, as well as an extrapolation based on radar data every 30 minutes, produced and accumulated in real time. Precipitation forecasts are provided every 6 hours [10]. The US Weather Bureau has operated a national river forecasting system since the early 1990s that combines weather forecasts with flood prediction models to disseminate flood warnings [11]. The precipitation input data are derived from satellite, ground radar, and rainfall gauges. China has established a network of 172 new generation Doppler weather radars, achieved a real-time data transmission every 6 minutes and provides composite weather radar data. This has helped to enhance the monitoring of sudden, potentially calamitous weather and improved the warning services for rainstorm disasters [12].

The performance of conceptual hydrological models depends to a large extent on the quality of the rainfall data but also on the selection of parameter values and model structure. Ideally, parameters are optimized using historical flood events, but model performance also relies to some extent on expert judgment and could be limited by the structure of the model, which is a simplification of a complex hydrological process [13–15]. Uncertainty associated with model structure and parameters has received attention [16–20], as has uncertainty associated with the quality of the rainfall data input [21, 22]. Error in remotely sensed precipitation data will be transmitted, and possibly amplified, through the hydrological simulation process. This paper examines the nature of uncertainty in flood forecasting due to transmission of errors associated with assimilated radar-based rainfall data.

In this paper the data assimilation methods of average calibration (AVG), optimum interpolation (OPT), Kalman filter (KLM), statistical weight integration (SWI), and variational calibration (VAR) were applied to radar rainfall data transformed from the classified Z-R (radar reflectivity-rainfall rates) relationship, and then the result was compared with precipitation data observed from a rain gauge network. The rainfall data were then used as input to the Xinanjiang hydrological model to predict flood hydrographs using the five sets of assimilated radar-based rainfall data and uncalibrated (UC) classified Z-R relationship rainfall data. This paper provides critical descriptive analysis of the main assimilation methods and the effect of the errors on uncertainty of flood forecasting using new data sets from the White Lotus River catchment, China. A novel aspect of this paper is quantification of the influence of the propagation of rainfall input error on simulated flood event peak flow magnitude, flood event volume, and flood peak timing. This

TABLE 1: Timing of the eight rainstorms chosen for the study.

Event	Starting time	Ending time	Duration (hours)
050408	2005-04-08, 22:00	2005-04-09, 14:00	16
050429	2005-04-29, 07:00	2005-05-01, 21:00	62
050609	2005-06-09, 22:00	2005-06-10, 14:00	16
060411	2006-04-11, 20:00	2006-04-13, 00:00	28
060508	2006-05-08, 08:00	2006-05-09, 23:00	39
060627	2006-06-27, 02:00	2006-06-27, 19:00	17
060705	2006-07-05, 10:00	2006-07-06, 08:00	22
060726	2006-07-26, 15:00	2006-07-27, 13:00	22

was done by perturbing radar rainfall data using the Breeding of Growing Modes (BGM) method.

2. Study Area

The White Lotus River, located in Hubei Province, China, has a catchment area of 1797 km². Its headwaters are in the Dabie Mountains and it forms a midcatchment tributary of the Xishui River, which flows into the lower Yangtze River (Figure 1). The catchment lies within the subtropical monsoon climate zone, with cold dry winters, hot wet summers, and distinct climatic seasonality. Mean annual temperature is 16.7°C, and annual average precipitation is 1366 mm, with the majority falling from June to August. In the wet season, rainstorms occur with a high frequency and intensity, and the short convergence time typically results in flood hydrographs with rapid rates of rise and fall.

3. Data and Method

3.1. Data Sets. The White Lotus catchment has 19 rain gauges and one river hydrological station at the outlet (Figure 1). This study used hourly recorded rainfall data from the years 2005 and 2006. The radar data were sourced from the CINRAD/SA Doppler weather radar located in Wuhan for the flood seasons of 2005 and 2006. The study area is located within the radar radius of 150 to 250 km. Generally, the 0.5° elevation [23] reflectivity factor plan position indicator (PPI) was used to develop the transformation of Z-R relationship. Radar estimated rainfall was quantified on a 1 km × 1 km grid at a 1-hour time step. Eight rainstorms were chosen for analysis (Table 1). The total sampled period for these events was 222 hours.

3.2. Radar-Based Rainfall Data Assimilation

3.2.1. Average Calibration (AVG). AVG is a simple and precise method for calculating regional rainfall volume [24, 25] whereby N rain gauges are used to calculate the mean calibration factor:

$$F = \frac{1}{N} \sum_{i=1}^{N} \frac{G_i}{I_i}, \tag{1}$$

where G_i is the observed rainfall of station i and I_i is the radar reflection rate of the station. AVG is applied by
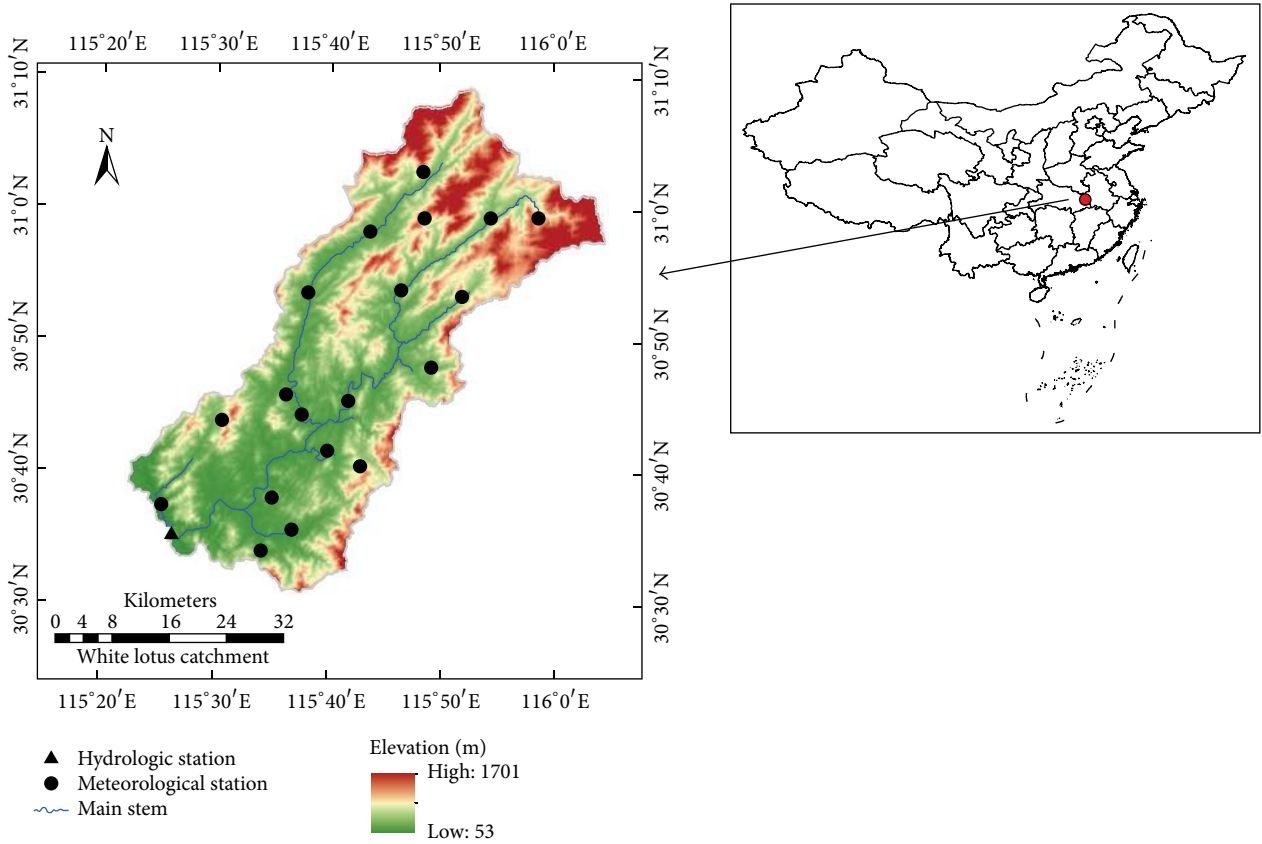
FIGURE 1: Location of White Lotus River catchment and the meteorological and river gauging stations within the catchment.

multiplying I by F to give regional rainfall distribution. In the situation of an inconsistent spatial pattern of rainfall over an area of interest, a number of regional AVG factors could be calculated.

3.2.2. Kalman Filter (KLM). In radar-based precipitation measurement, there exists not only system errors produced by unstable radar capability but also the random errors due to the unstable relation between Z and R and the influence of wind field. These random errors are called noise. KLM aims to find the optimum estimation of calibration factor f from noisy data.

The basic theory of KLM [26, 27] is that if f_1 and f_2, two independent estimations of the variable f, can be obtained by different equations, the relation between f, f_1, and f_2 is

$$f = (1 - \omega) f_1 + \omega f_2, \tag{2}$$

where ω is a weighting factor. The criterion for determining ω is minimizing the variance of the calibrated f. KLM is used to determine the weight, with the input being observed data and the output being the optimum estimation of f. The KLM algorithm is a real-time recursive filter, using present input data and the previously calculated state. It is a real-time recursive filter, with the accuracy of estimation improving continuously as the number of observations increases.

3.2.3. Optimum Interpolation (OPT). The OPT method uses radar rainfall data derived from echoes as the first-guess

background meteorological field and then corrects this using a linear combination of the differences between this first guess and observed data from the rain gauges. The weighting function is determined by minimizing the expected square error of the data [28, 29]. An OPT is generally expressed as follows:

$$A_g = B_g + \sum_{i=1}^{N} (O_i - B_i) P_i, \tag{3}$$

where A_g (B_g) is an analysis (first guess) value at a grid cell g to be interpolated, O_i (B_i) is an observed (first guess) value given at observation point i, and P_i denotes a weight function at observation point i. The optimum weight is computed by supposing that the errors are unbiased and uncorrelated and can be expressed as

$$\sum_{j=1}^{N} P_i \mu_{ij} = \mu_{gi} \quad (i = 1, 2, \ldots, N), \tag{4}$$

where μ represents the first-guess error correlation coefficient between two arbitrary grid points.

3.2.4. Variational Calibration (VAR). The VAR method [30, 31] uses radar-based and gauge-measured precipitation for each observation point to achieve a calibration factor $C\widetilde{R}(k)$ $(k = 1, 2, \ldots, n)$ and then obtain $C\widetilde{R}(i, j)$ in every grid cell (i, j) by interpolation. Next, an optimum calibration factor

field $CR(i, j)$ which minimizes the sum of squared errors between $CR(i, j)$ and $C\widetilde{R}(i, j)$ is sought. Finally, the analyzed precipitation in the grid (i, j) is

$$P(i, j) = P_R(i, j) + CR(i, j), \tag{5}$$

where P (P_R) is analysis (radar-based) value at grid (i, j).

3.2.5. Statistical Weight Integration (SWI).
The SWI method averages M results of precipitation intensity distribution, derived by M methods according to weighting that depends on the reciprocal of error statistics.

Error $e_{ik} = x_{ik} - y_k$, where x_{ik} is the output value of ith method in grid k and y_k is the measured value of automatic rain gauge in the same grid. Error variance $\sigma_i^2 = (1/N) \sum_{k=1}^{N} (e_{ik} - \bar{e}_i)^2$, where $\bar{e}_i$ is the mean error of each method. The reciprocal of the error variance of various methods is usually used as the weight of the integration analysis. The result after integrating is

$$\widehat{y}_k = \frac{\sum_{i=1}^{M} x_{ik}/\sigma_i^2}{\sum_{i=1}^{M} 1/\sigma_i^2}, \tag{6}$$

where $\widehat{y}_k$ is estimated results of the kth grid after integration.

3.2.6. Evaluation Criteria.
The absolute error (AE) and relative error (RE) were used as criteria to evaluate the five assimilation methods. The Nash-Sutcliffe (NS), peak relative error (PRE), and peak time difference (PTD) were used to evaluate the results of hydrological modelling.

$$NS = 1 - \frac{\sum_{i=1}^{N} \left(\widehat{Q}_i - Q_i\right)^2}{\sum_{i=1}^{N} \left(Q_i - \overline{Q}\right)^2},$$

$$PRE = \left(\frac{\sum_{i=1}^{N} \widehat{q}_i - \sum_{i=1}^{N} q_i}{\sum_{i=1}^{N} q_i}\right) \times 100\%, \tag{7}$$

$$PTD = \sum \frac{\left(|T_s - T_o| \leq 3\right)}{N},$$

where $\widehat{Q}_i$ is the predicted streamflow at time step i, Q_i is the observation at time step i, while $\overline{Q}$ is the mean of observations, N represents the total number of observations, $\widehat{q}_i$ is the predicted peak flow at time step i, q_i is the measured peak flow at the same time, T_s is the predicted peak time, and T_o is the observed peak time.

3.3. Precipitation-Runoff Model.
The Xinanjiang model developed by Zhao [32] has been applied extensively and successfully to predict runoff from precipitation in humid and semihumid regions [33]. This model conceptualizes that runoff is not produced until the soil moisture content of the aeration zone reaches field capacity, and thereafter runoff equals the rainfall excess without further loss. It consists of three submodels, a three-layer evapotranspiration submodel, a runoff generation submodel, and a runoff routing submodel. Many studies have described the structure of the

model in detail [32–35]. The inputs to the model are daily areal precipitation and pan evaporation, with observed daily discharge used for calibration. The outputs are the modelled daily discharge at the catchment outlet and the estimated actual evapotranspiration of the catchment [36]. The water balance equation of the model can be expressed as follows:

If $P_E + A < W'_{mm}$

$$R = P_E + W - W_M + W_M \left(1 - \frac{P_E + A}{W'_{mm}}\right)^{1+B}. \tag{8}$$

If $P_E + A \geq W'_{mm}$

$$R = P_E + W - W_M, \tag{9}$$

where R is the runoff, P_E is the effective precipitation which equals the precipitation minus evaporation, W is the initial water storage within the catchment and A is the ordinate value corresponding to W, W'_{mm} is the maximum storage capacity of the catchment, W_M is the areal mean tension water capacity, which is composed of the capacity of each soil layer and represents drought conditions, and B is a parameter that represents the heterogeneity of the water storage capacity of the catchment.

To reasonably describe the evapotranspiration process, the soil profile is divided into three layers: the upper soil layer, the lower soil layer, and the deepest layer. If the water content in the upper layer is sufficient, evaporation equals the evaporation capacity of the upper layer. If not, all upper water is evaporated and the residual evaporation is sourced from the lower layer. If the water content in the lower layer is insufficient to meet the residual evaporation capacity, the water in the deepest layer will be evaporated.

The generated runoff is divided into the surface flow, interflow, and groundwater using steady infiltration. The total runoff can be routed by a linear system before arriving at the outlet of the catchment [37]. Flow routing uses the Muskingum or piecewise continuous algorithm. Taking into consideration the uneven distribution of rainfall and the nature of the underlying surface, the catchment is divided into a set of subcatchments using an appropriate method, such as Thiessen polygons. Finally, the total catchment runoff is obtained after the hydrographs at the outlets of each subcatchment are simulated and flood routing is applied [35].

3.4. Uncertainty Methods.
Most research into uncertainty of hydrograph forecasting has focused on model parameters and structure rather than quality of input data. Toth and Kalnay [38] recommended perturbation of the initial condition with perturbation fields that are representative of errors present in the analysis and then using these perturbations as the input to the model in the analysis of input uncertainty. The main methods used to generate perturbations that reflect the initial uncertainty are Monte Carlo, Singular Vector, and Breeding of Growing Modes (BGM). Monte Carlo is a statistical method, whereby the smaller the sample number the less

TABLE 2: Comparison of calibration results among different methods. UC is uncalibrated, KLM is Kalman filter, AVG is average calibration, OPT is optimum interpolation, VAR is variational calibration, and SWI is statistical weight integration.

| Method | Statistic | Storm event start date (yymmdd) | | | | | | | | Average |
		050408	050429	050609	060411	060508	060627	060705	060726	
UC	AE (mm)	27.35	21.38	5.79	3.97	5.73	7.52	3.58	6.72	10.26
	RE (%)	70.02	44.89	23.06	13.49	14.13	26.93	11.78	37.02	30.17
	Results (mm)	11.71	26.25	19.32	33.39	34.8	20.42	26.79	24.88	24.70
KLM	AE (mm)	4.81	4.61	1.45	1.38	1.58	1.9	1.29	1.06	2.26
	RE (%)	12.32	9.68	5.76	4.69	3.89	6.79	4.26	5.84	6.65
	Results (mm)	34.25	52.24	23.67	28.04	38.95	26.05	29.08	17.1	31.17
AVG	AE (mm)	4.22	5.22	1.2	2	3.14	5.2	1.17	2.03	3.02
	RE (%)	10.81	10.96	4.8	6.8	7.75	18.59	3.86	11.17	9.34
	Results (mm)	34.83	52.85	23.91	31.42	37.38	22.75	29.2	16.13	31.06
OPT	AE (mm)	4.11	4.27	1.32	0.36	0.79	1.47	1.16	0.1	1.70
	RE (%)	10.52	8.96	5.25	1.24	1.94	5.27	3.81	0.57	4.70
	Results (mm)	34.95	51.9	23.79	29.06	39.74	26.47	29.21	18.05	31.65
VAR	AE (mm)	4.11	4.26	1.32	0.05	0.32	1.47	1.05	0.16	1.59
	RE (%)	10.53	8.94	5.27	0.18	0.78	5.27	3.47	0.88	4.42
	Results (mm)	34.95	51.89	23.79	29.37	40.21	26.47	29.31	18	31.75
SWI	AE (mm)	4.11	2.95	1.32	0.05	0.26	1.47	1.05	0.05	1.41
	RE (%)	10.53	6.18	5.27	0.18	0.65	5.26	3.47	0.29	3.98
	Results (mm)	34.95	50.58	23.79	29.37	40.26	26.47	29.32	18.1	31.61
Measured (mm)		39.06	47.63	25.11	29.42	40.53	27.94	30.37	18.16	32.28

powerful the statistics. However, the sample number is often very limited, and randomly chosen initial values do not perfectly match with the dynamical mode. Although close to actual weather at initial time, the predicted results often deviate rapidly from the actual atmospheric condition due to the adjustment of the mode itself [39]. The Singular Vector method is better able to deal with many nonquantitative hypotheses in the data assimilation, increase the number of ensemble members, and more easily capture the errors generated by analysis. It can also determine the direction of the fastest disturbance and have a better dispersion. Disadvantages of this method are that the components of errors that do not grow are ignored and the calculations are tedious [40, 41].

The BGM proposed by Toth and Kalnay [38, 42] was designed to model how growing errors are bred and maintained in a conventional analysis cycle through successive use of short-range forecasts, with the bred modes offering an estimate of possible growing error fields in the analysis. The growth rate obtained by this method is better than that of the Monte Carlo method and even higher than the growth rate of the lagged average forecasting. A formula for the specific perturbation is generated:

$$P = S \times \text{Rmse} \times \text{Rand}, \tag{10}$$

where P is random perturbation field and S is perturbation amplitude coefficient. In order to make the disturbance amplitude not too large, S is set equal to 1. Rmse is the mean square error between the measured rainfall and rainfall estimation from radar data assimilation. Rand is a random distribution function which is evenly distributed between

−1 and 1, where the equation [43] for producing a random number that is uniformly distributed in $[a, b]$ is

$$x_i = a + (b - a) r_i, \tag{11}$$

where r_i is a uniformly distributed random number that belongs to $[0, 1]$.

Using the measured data of eight storm events from 19 rain gauge stations and radar estimations calibrated by the KLM method, 100 sets of random perturbation field were produced for each event by the BGM of initial disturbance. Then 200 sets of disturbance for each event, formed after the superposition and deduction of the perturbation field, were used as the input to the Xinanjiang hydrological model. For the eight storm events, this resulted in 1600 predicted hydrographs. The influence of the propagation of rainfall input error on runoff simulation was quantified using the three previously described goodness-of-fit statistics.

4. Results and Discussion

4.1. Evaluation of Methods of Assimilation of Radar Rainfall Data.
A comparison of uncalibrated and calibrated results of five kinds of assimilation methods for eight storm events (Table 2) revealed that, among the methods, the SWI method produced results closest to the observed data, followed by the VAR method, with a mean absolute error smaller than that of the SWI method. The next best method was the OPT method, followed by the KLM method and, finally, the AVG method. The average of the results across the eight storm events (Table 2) allows a rapid comparison of the relative performances of the five assimilation methods.
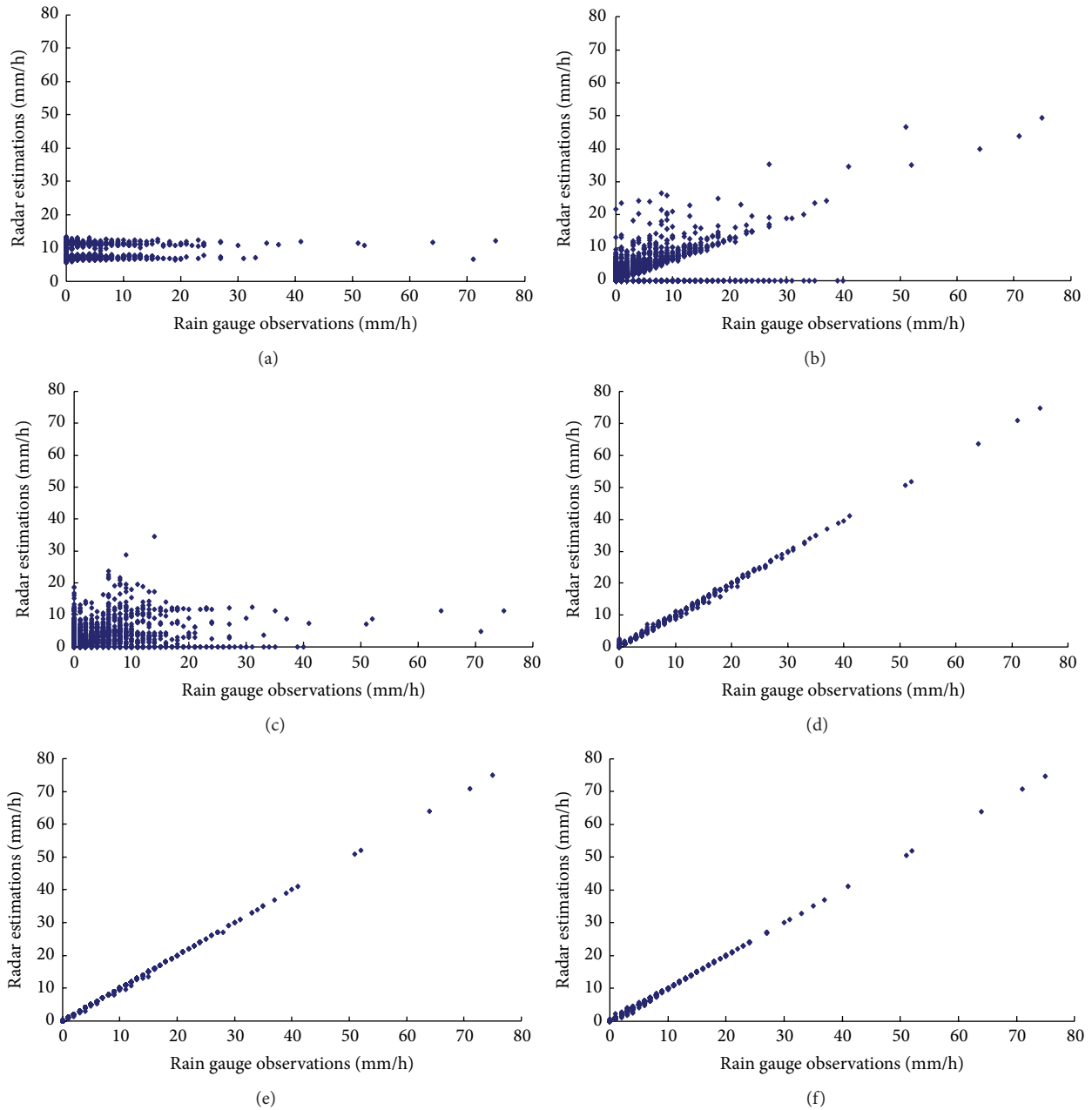
FIGURE 2: Hourly gauged rainfall observations compared to radar-based estimations: (a) uncalibrated classified *Z-R* data, (b) Kalman filter, (c) average calibration, (d) optimum interpolation, (e) variational calibration, and (f) statistical weight integration.

For AVG, the accuracy of regional total precipitation estimation improved after calibration, and the accuracy would improve further with more rain gauges. One problem with the AVG method is that it does not account for the shifting location of the focus of intense rainfall during storm events. AVG cannot assign greater weight to points independently identified to be of particular interest, and it does not consider the representativeness of the data derived by the automatic rain gauges and radar echoes. These problems explain the relatively low performance of the AVG method.

Variable terrain, synoptic conditions, and precipitation intensity impart a high degree of spatial and temporal variability to radar rainfall data that presents a challenge for data assimilation. The results indicated that the KLM method of data assimilation, with its strong theoretical basis and consideration of the distribution of error in determining the calibration factor, produced a better result than the AVG method. The KLM method can improve accuracy by reducing the relative deviation and degree of dispersion between the hyetometer and radar rainfall data and eliminating the intrusion of nonrainfall echoes from rainfall estimation. This method calibrates the radar estimated rainfall distribution field in the time domain rather than the spatial structure error of precipitation, which means it is more suitable for

the calibration of stable stratiform cloud precipitation and in places with low density of rain gauges. These may be the key factors that limit the calibration accuracy of the KLM method. The spatial structure and distribution of rain clouds are complicated and changeable, with mixed convective precipitation being a common phenomenon. Under these conditions, the OPT and VAR methods, which have similar results (Table 2), outperform the KLM method. This relative performance difference would also apply in places with a high density of rain gauges. While each precipitation calibration model has its advantages depending on the conditions, SWI combines the advantages of the other four methods and gives an appropriate weighting after the calculation. On this basis the SWI method would be expected to be superior to the other data assimilation methods tested here [44], and this was supported by the results (Table 2).

Scatterplots of hourly gauged rainfall and uncalibrated and calibrated radar-based rainfall (Figure 2) illustrated the good agreement between the rainfall distribution obtained by the SWI method and gauged rainfall. The VAR and OPT assimilation methods also performed well. The KLM and AVG method produced noticeably poorer results (Figure 2).

4.2. Hydrological Model for Flood Hydrograph Prediction. The Xinanjiang hydrological model was applied in the White Lotus River catchment taking into consideration temporal and spatial variation of rainfall and the nonlinear effects of watershed topography and river channel characteristics on the runoff concentration and dispersed inflow and nonlinear convergence. Model parameters were calibrated by the method of Rosenbrock [45]. The three indexes certainty factor, peak relative error, and peak time difference were used to evaluate the flood forecasting results. Goodness-of-fit statistics (Table 3) and simulated hydrographs (Figure 3) indicated that the methods of SWI, OPT, and VAR were superior in producing rainfall input data for storm hydrograph modelling. In particular, the SWI simulated hydrographs were in almost perfect agreement with the observed data. The rainfall data generated by the methods of KLM, AVG, and UC led to unsatisfactory hydrograph prediction.

4.3. Uncertainty Analysis of Hydrological Forecasting. The relative error of predicted results increased after assimilated rainfall data were randomly disturbed compared with the undisturbed peak flow, flood volume, and peak time difference. The peak flow and flood volume were increased by 5% and 8%, respectively (Figures 4 and 5), which means that the model input errors increased when they were processed through the hydrological model. Also, most of the predicted peak times lagged the observed times (Figure 6).

The error of precipitation input data in hydrological modelling is a significant source of error in predicted peak flow, flood volume, and peak time. Therefore, the spread of precipitation uncertainty has a great influence on runoff prediction. In China, practical applications usually adopt qualitative rainfall prediction and assumed rainfall input, and this would have implications for the forecast period and accuracy of flood forecasting. At present, both the accuracy and resolution of rainfall forecasts provided by numerical weather prediction (NWP) cannot meet the requirements of quantitative flood forecasting, and this situation may continue to exist into the foreseeable future. Weather radar and satellite remote sensing (SRS) can perform an important role in quantitative precipitation forecasting. Use of precipitation forecasts obtained by SRS as the input of hydrological models to simulate and predict runoff warrants further research attention. This could be a key step towards improving the accuracy of flood forecasting. It will be necessary to develop methods for combining and assimilating the SRS data and observed data from rain gauges for producing better precipitation forecasts. SRS has sufficiently high spatial and temporal resolution to effectively characterize the instantaneous distribution of precipitation throughout a catchment, while rain gauges can provide high-precision single-point observations.

The generation of floods is a complex and dynamic process. Modelling this process can be improved through the use of rainfall input data with high spatial and temporal resolution. How to best make use of rainfall information available from multiple sources, including gauge networks, radar, and satellite monitoring, is one of the key problems to be solved in hydrological modelling and forecasting.

5. Conclusions

Because of the high level of risk posed by floods, especially flash floods, it is critical that warning procedures make use of accurate nowcasting and short-range runoff forecasting capability. Precipitation data is the most important input to hydrological models used for forecasting hydrographs. While rainfall data is usually sourced from a network of gauges, an alternative is weather radar detector, which offers superior spatial and temporal resolution but at the cost of low precision and accuracy. This paper investigated methods of precipitation data assimilation with the objective of improving the quality of radar-based rainfall estimates as a means of achieving more reliable flood forecasts. Using data from a network of rain gauges in the White Lotus River in Hubei Province, China, and data from the weather radar detector based in nearby Wuhan, five established methods of assimilating radar data estimated by the classified *Z-R* relationship were applied. The assimilated data was compared with the observed rainfall data and then used in a hydrological model to predict historical flood hydrographs. Finally, BGM was used for the perturbation of radar data to evaluate the effect of the rainfall data input error on runoff simulation. The main conclusions arising from this work were as follows:

(1) Statistical weight integration, variational calibration, and optimum interpolation were superior methods of data assimilation, while Kalman filter and average calibration methods gave a less satisfactory result.

(2) For modelling storm event hydrographs, rainfall data generated by the statistical weight integration, variational calibration, and optimum interpolation methods resulted in hydrographs that were the closest fit to the gauged data. Rainfall data generated by the average calibration and Kalman filter methods of data assimilation resulted in unsatisfactory hydrograph prediction.

TABLE 3: Comparison of flood hydrograph prediction using different rainfall data assimilation methods. MPF is measured peak flow and FPF is forecast peak flow.

| Method	Flood event	MPF (m^3/s)	FPF (m^3/s)	NS	PRE (%)	PTD (h)
UC	050408	313.3	58.4	−0.73	81.36	1
	050429	618.2	167.6	−1.05	72.89	32
	050609	187.9	96.8	−0.47	48.48	2
	060411	232.7	157	−0.87	32.53	15
	060508	240.9	108.1	−0.59	55.13	10
	060627	189.6	76.6	−2.95	59.6	4
	060705	310.9	134.8	−3	56.64	1
	060726	147.3	113.4	0.2	23.01	8
KLM	050408	313.3	275.6	0.83	12.03	2
	050429	618.2	562.2	0.87	9.06	1
	050609	187.9	166.7	0.9	11.28	1
	060411	232.7	214.7	0.9	7.74	1
	060508	240.9	224.1	0.9	6.97	1
	060627	189.6	168.5	0.84	11.13	2
	060705	310.9	262.7	0.82	15.5	2
	060726	147.3	136.1	0.86	7.6	1
AVG	050408	313.3	283.3	0.79	9.58	1
	050429	618.2	651.5	0.86	5.39	5
	050609	187.9	155.1	0.88	17.46	2
	060411	232.7	196.2	0.83	15.69	0
	060508	240.9	212.2	0.85	11.91	1
	060627	189.6	155.2	0.44	18.14	4
	060705	310.9	250.5	0.65	19.43	1
	060726	147.3	108.2	0.22	26.54	0
OPT	050408	313.3	311.7	0.89	0.51	2
	050429	618.2	615.4	0.9	0.45	1
	050609	187.9	187.6	0.9	0.16	3
	060411	232.7	231.7	0.89	0.43	2
	060508	240.9	240.3	0.9	0.25	2
	060627	189.6	188.6	0.89	0.53	1
	060705	310.9	307.1	0.89	1.22	2
	060726	147.3	144.6	0.9	1.83	3
VAR	050408	313.3	313.3	0.9	0.54	0
	050429	618.2	618.2	0.91	0.36	0
	050609	187.9	187.9	0.89	0.85	1
	060411	232.7	232.7	0.9	1.16	2
	060508	240.9	240.9	0.9	0.23	1
	060627	189.6	189.6	0.89	0.63	1
	060705	310.9	310.9	0.89	1.21	2
	060726	147.3	147.3	0.74	0.47	0
SWI	050408	313.3	312.5	0.89	0.26	1
	050429	618.2	616.6	0.89	0.26	0
	050609	187.9	187.9	0.9	0.34	2
	060411	232.7	207.2	0.87	10.96	1
	060508	240.9	240.6	0.89	0.12	0
	060627	189.6	188.6	0.89	0.53	1
	060705	310.9	307.2	0.89	1.19	1
	060726	147.3	146.8	0.9	0.34	2
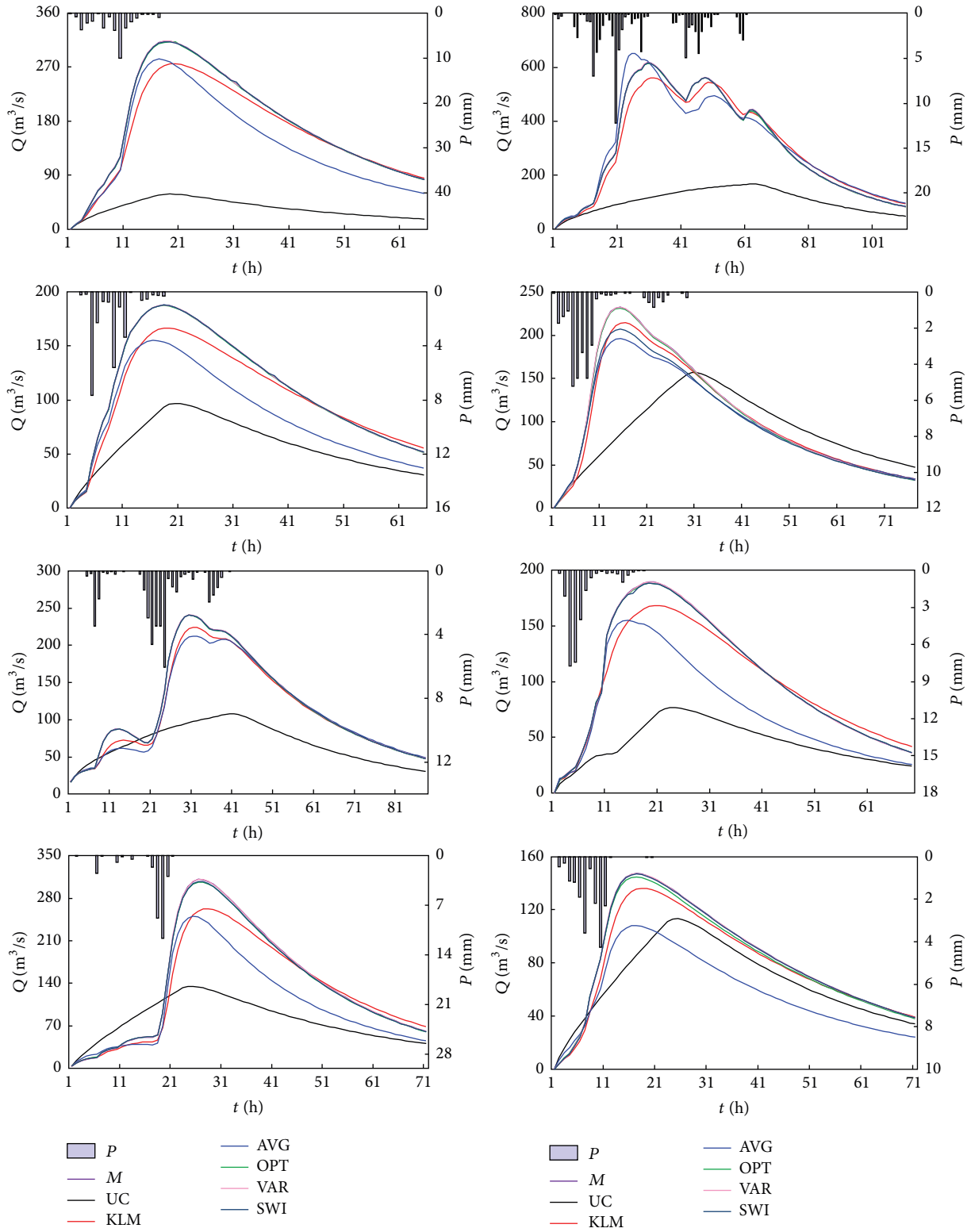
FIGURE 3: Comparison of simulated hydrographs for eight storm events using rainfall input data derived from five assimilation methods and uncalibrated radar data. M is measured hydrograph.
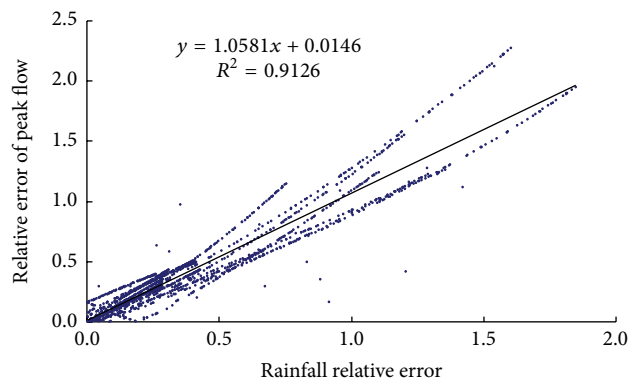
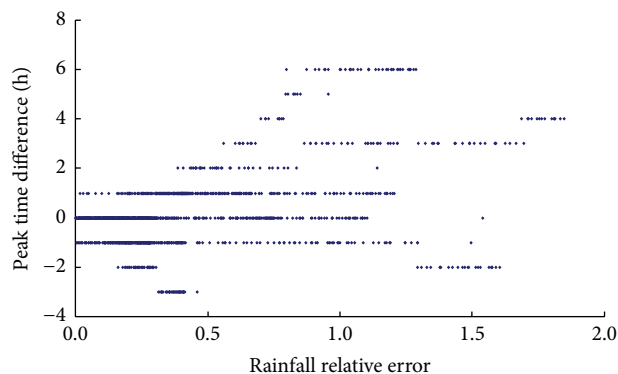FIGURE 4: Relationship between relative error in predicted flood peak magnitude and rainfall relative error.

FIGURE 5: Relationship between relative error in predicted flood volume and rainfall relative error.
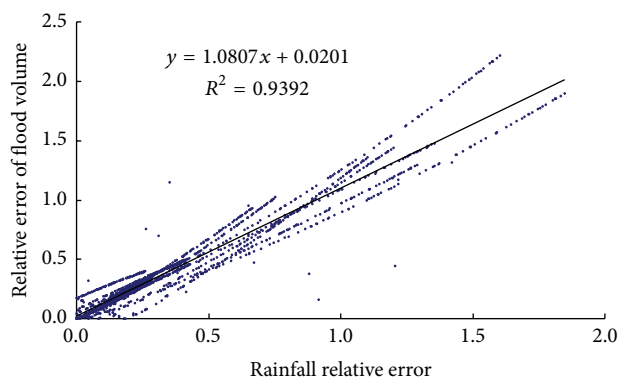

FIGURE 6: Relationship between error in predicted flood peak time and rainfall relative error.

(3) The random errors generated by model input had a tendency to increase during the process of running the hydrological model. Also, most of predicted flood peak times were lagged.

(4) As hydrometeorology ensemble forecasts become more widely available and the data acquisition methods develop, there will be an ongoing need for evaluation of uncertainty in input data.

(5) This paper provides critical descriptive analysis of the main assimilation methods and the effect of the errors on uncertainty of flood forecasting. It demonstrates that weather radar has potential to improve the precision of hydrological forecasting, which has implications for improved flood control and disaster mitigation. The influence of the propagation of rainfall input error on three simulated factors of hydrological forecast is quantified.

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

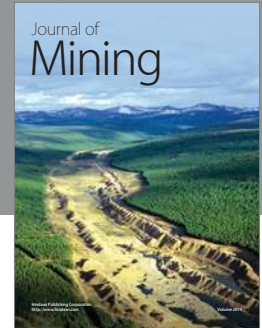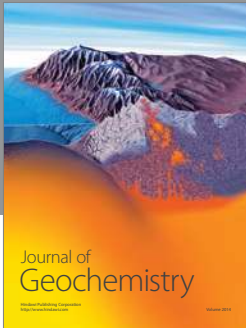[1] I. Yucel, A. Onen, K. K. Yilmaz, and D. J. Gochis, "Calibration and evaluation of a flood forecasting system: utility of numerical weather prediction model, data assimilation and satellite-based rainfall," *Journal of Hydrology*, vol. 523, pp. 49–66, 2015.

[2] L. P. Zhang, L. Li, A. Z. Ye et al., "Contrast research on precision of estimating regional precipitation with weather radar and rain gauge," *Engineering Journal of Wuhan University. Engineering Edition*, vol. 40, no. 1, pp. 1–5, 2007 (Chinese).

[3] J. Liu, M. Bray, and D. Han, "A study on WRF radar data assimilation for hydrological rainfall prediction," *Hydrology and Earth System Sciences*, vol. 17, no. 8, pp. 3095–3110, 2013.

[4] I. Maiello, R. Ferretti, S. Gentile et al., "Impact of radar data assimilation for the simulation of a heavy rainfall case in central Italy using WRF-3DVAR," *Atmospheric Measurement Techniques*, vol. 7, no. 9, pp. 2919–2935, 2014.

[5] M. Borga, "Accuracy of radar rainfall estimates for streamflow simulation," *Journal of Hydrology*, vol. 267, no. 1-2, pp. 26–39, 2002.

[6] Q. Dai and D. Han, "Exploration of discrepancy between radar and gauge rainfall estimates driven by wind fields," *Water Resources Research*, vol. 50, no. 11, pp. 8571–8588, 2014.

[7] F. Giannoni, J. A. Smith, Y. Zhang, and G. Roth, "Hydrologic modeling of extreme floods using radar rainfall estimates," *Advances in Water Resources*, vol. 26, no. 2, pp. 195–203, 2003.

[8] Z. Sokol, "Utilization of regression models for rainfall estimates using radar-derived rainfall data and rain gauge data," *Journal of Hydrology*, vol. 278, no. 1–4, pp. 144–152, 2003.

[9] C. A. Velasco-Forero, D. Sempere-Torres, E. F. Cassiraga, and J. Jaime Gómez-Hernández, "A non-parametric automatic blending methodology to estimate rainfall fields from rain gauge and radar data," *Advances in Water Resources*, vol. 32, no. 7, pp. 986–1002, 2009.

[10] X. Sun, R. G. Mein, T. D. Keenan, and J. F. Elliott, "Flood estimation using radar and raingauge data," *Journal of Hydrology*, vol. 239, no. 1–4, pp. 4–18, 2000.

[11] J. Z. Sun, "Convective-scale assimilation of radar data: progress and challenges," *Quarterly Journal of the Royal Meteorological Society*, vol. 131, no. 613, pp. 3439–3463, 2006.

[12] X. D. Yu, X. P. Yao, T. N. Xiong et al., *The Principle and Application of Doppler Weather Radar*, China Meteorological Press, Beijing, China, 2007.

[13] M. Borga, E. N. Anagnostou, and E. Frank, "On the use of real-time radar rainfall estimates for flood prediction in mountainous basins," *Journal of Geophysical Research Atmospheres*, vol. 105, no. 2, pp. 2269–2280, 2000.

[14] H. G. Zhang, S. L. Guo, and X. L. He, "Recent advancement and prospect of hydrological forecasting uncertainty study," *Journal of Shihezi University (Natural Science)*, vol. 24, no. 1, pp. 15–21, 2006 (Chinese).

[15] G. Villarini and W. F. Krajewski, "Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall," *Surveys in Geophysics*, vol. 31, no. 1, pp. 107–129, 2010.

[16] K. Beven and A. Binley, "The future of distributed models: model calibration and uncertainty prediction," *Hydrological Processes*, vol. 6, no. 3, pp. 279–298, 1992.

[17] F. Quintero, D. Sempere-Torres, M. Berenguer, and E. Baltas, "A scenario-incorporating analysis of the propagation of uncertainty to flash flood simulations," *Journal of Hydrology*, vol. 460-461, pp. 90–102, 2012.

[18] L. Xiong and K. M. O'Connor, "An empirical method to improve the prediction limits of the GLUE methodology in rainfall-runoff modeling," *Journal of Hydrology*, vol. 349, no. 1-2, pp. 115–124, 2008.

[19] X. R. Yin, J. Xia, and X. Zhang, "Recent progress and prospect of the study on uncertainties in hydrological modeling and forecasting," *Water Power*, vol. 32, no. 10, pp. 27–31, 2006 (Chinese).

[20] D. Zhu, Y. Xuan, and I. Cluckie, "Hydrological appraisal of operational weather radar rainfall estimates in the context of different modelling structures," *Hydrology and Earth System Sciences*, vol. 18, no. 1, pp. 257–272, 2014.

[21] A. S. Falck, V. Maggioni, J. Tomasella, D. A. Vila, and F. L. R. Diniz, "Propagation of satellite precipitation uncertainties through a distributed hydrologic model: a case study in the Tocantins-Araguaia basin in Brazil," *Journal of Hydrology*, vol. 527, pp. 943–957, 2015.

[22] P. A. Mendoza, J. McPhee, and X. Vargas, "Uncertainty in flood forecasting: a distributed modeling approach in a sparse data catchment," *Water Resources Research*, vol. 48, no. 9, Article ID W09532, 2012.

[23] V. T. Wood, R. A. Brown, and S. V. Vasiloff, "Improved detection using negative elevation angles for mountaintop WSR-88Ds. Part II: simulations of the three radars covering Utah," *Weather and Forecasting*, vol. 18, no. 3, pp. 393–403, 2003.

[24] T. P. Dai and D. S. Fu, "Precision of precipitation amount in the joint detection areas by weather radar and the auto-rain Gauge network," *Journal of Nanjing Institute of Meteorology*, vol. 13, no. 4, pp. 592–597, 1990 (Chinese).

[25] J. W. Wilson and E. A. Brandes, "Radar measurement of rainfall—a summary," *Bulletin of the American Meteorological Society*, vol. 60, no. 9, pp. 1048–1058, 1979.

[26] S. Chumchean, A. Seed, and A. Sharma, "Correcting of real-time radar rainfall bias using a Kalman filtering approach," *Journal of Hydrology*, vol. 317, no. 1-2, pp. 123–137, 2006.

[27] M. Costa and M. Monteiro, "Bias-correction of Kalman filter estimators associated to a linear state space model with estimated parameters," *Journal of Statistical Planning and Inference*, vol. 176, pp. 22–32, 2016.

[28] S. I. Kako, A. Isobe, and M. Kubota, "High-resolution ASCAT wind vector data set gridded by applying an optimum interpolation method to the global ocean," *Journal of Geophysical Research: Atmospheres*, vol. 116, no. 23, Article ID D23107, 2011.

[29] S. Kako and M. Kubota, "Relationship between an El Niño event and the interannual variability of significant wave heights in the North Pacific," *Atmosphere-Ocean*, vol. 44, no. 4, pp. 377–395, 2006.

[30] W. Castaings, D. Dartus, F.-X. Le Dimet, and G.-M. Saulnier, "Sensitivity analysis and parameter estimation for distributed hydrological modeling: potential of variational methods," *Hydrology and Earth System Sciences*, vol. 13, no. 4, pp. 503–517, 2009.

[31] P. C. Zhang, T. P. Dai, D. S. Fu et al., "Principle and accuracy of adjusting the area precipitation from digital weather radar through variational method," *Atmospheric Science*, vol. 16, no. 2, pp. 248–256, 1992 (Chinese).

[32] R.-J. Zhao, "The Xinanjiang model applied in China," *Journal of Hydrology*, vol. 135, no. 1–4, pp. 371–381, 1992.

[33] A. A. Alazzy, H. S. Lü, and Y. H. Zhu, "Assessing the uncertainty of the Xinanjiang rainfall-runoff model: effect of the likelihood function choice on the GLUE method," *Journal of Hydrologic Engineering*, vol. 20, no. 10, Article ID 04015016, 2015.

[34] J. Liu, X. Chen, J. Zhang, and M. Flury, "Coupling the Xinanjiang model to a kinematic flow model based on digital drainage networks for flood forecasting," *Hydrological Processes*, vol. 23, no. 9, pp. 1337–1348, 2009.

[35] C. Zhang, R.-B. Wang, and Q.-X. Meng, "Calibration of conceptual rainfall-runoff models using global optimization," *Advances in Meteorology*, vol. 2015, Article ID 545376, 12 pages, 2015.

[36] H. Xu, C.-Y. Xu, N. R. Sælthun, B. Zhou, and Y. Xu, "Evaluation of reanalysis and satellite-based precipitation datasets in driving hydrological models in a humid region of Southern China," *Stochastic Environmental Research and Risk Assessment*, vol. 29, no. 8, pp. 2003–2020, 2015.

[37] S. S. Meng, X. H. Xie, and X. Yu, "Tracing temporal changes of model parameters in rainfall-runoffmodeling via a real-time data assimilation," *Water*, vol. 8, no. 1, article 19, 2016.

[38] Z. Toth and E. Kalnay, "Ensemble Forecasting at NMC: the generation of perturbations," *Bulletin of the American Meteorological Society*, vol. 74, no. 12, pp. 2317–2330, 1993.

[39] S. L. Mullen and D. P. Baumhefner, "Monte Carlo simulations of explosive cyclogenesis," *Monthly Weather Review*, vol. 122, no. 7, pp. 1548–1567, 1994.

[40] R. Buizza and T. N. Palmer, "Impact of ensemble size on ensemble prediction," *Monthly Weather Review*, vol. 126, no. 9, pp. 2503–2518, 1998.

[41] F. Molteni, R. Buizza, T. N. Palmer, and T. Petroliagis, "The ECMWF ensemble prediction system: methodology and validation," *Quarterly Journal of the Royal Meteorological Society*, vol. 122, no. 529, pp. 73–119, 1996.

[42] Z. Toth and E. Kalnay, "Ensemble forecasting at NCEP and the breeding method," *Monthly Weather Review*, vol. 125, no. 12, pp. 3297–3319, 1997.

[43] F. Wu, "Several methods of creating stochastic numbers and applications," *Journal on Numerical Methods and Computer Applications*, vol. 27, no. 1, pp. 48–51, 2006 (Chinese).

[44] Y.-H. Shao and W.-C. Zhang, "Application of statistical weight matrix method in estimation of regional rainfall," *Advances in Water Science*, vol. 20, no. 6, pp. 755–762, 2009 (Chinese).

[45] X. Y. Song, Y. J. Li, H. Y. Yu et al., "Application of quantitative precipitation estimation by radar to flood forecast," *Engineering Journal of Wuhan University. Engineering Edition*, vol. 40, no. 2, pp. 51–54, 2007 (Chinese).