

Rainfall interpolation based on flood simulation in northern China

ZHAO Ran-hang¹, ZHOU Lu¹, Li Hua-xing¹, LI Hong-tao², QI Zhen¹, QIU Ting¹, GAN Tian¹

¹School of Civil Engineering and Hydraulic Engineering, Shandong University; Jinan 250061, China

²Jinan Hydrographic office, Jinan 250014, China

Abstract. Due to the short duration, high intensity and sudden intensity of torrential rain, mountain torrents are easily formed in the northern hilly area. The distributed hydrological model is used as the main means of rain-flood forecasting. Rainfall as an important input, its spatial interpolation accuracy and time scale directly affect the forecast results. Therefore, in this paper, the spatial interpolation calculation and analysis of rainstorm process with hourly scale is carried out in the northern hilly area. Taking Licheng district in Jinan as the research area, the spatial interpolation methods, such as IDW, OK and OCK, are used to calculate the spatial interpolation of 16 time-by-time rainstorm processes for 11 typical rainfall with different duration, and the interpolation results are cross-validated and error analysis is carried out. The results show that the accuracy of the collaborative Kriging interpolation considering the elevation is higher in the short duration rainstorm process of the hilly area.

1 Introduction

As an important input factor of distributed hydrological model, rainfall spatial interpolation accuracy and time scale directly affect the prediction results of rain flood. How to interpolate discrete rainfall data to obtain more accurate rainfall spatial distribution information has become a hot topic in hydrology. Li Jinjie¹ and others compared the interpolation method of average monthly rainfall in Southwest China, and proposed that the optimal method should be selected based on different regions and time scales; analysis of rainfall interpolation in different terrain areas by Fanglin Dong² shows that considering the influence of topography fluctuation on rainfall, the interpolation accuracy can be improved; Cara Tobin³ compare the spatial interpolation of multi-factor composite precipitation fields and input the hydrological model to explore the influence of interpolation results on runoff. The results show that the spatial interpolation method combined with auxiliary variables can better estimate the instantaneous rainfall and capture the spatial distribution.

The rainstorm in the hilly area of northern China has the characteristics of short duration, high intensity and sudden intensity, and its rain flood forecast is still the focus of mountain flood disaster research, and the input of hydrological model is mostly hourly scale. And domestic and foreign research is more inclined to day or above scale. Therefore, it is necessary to carry out the time-by-time interpolation analysis of rainstorm process. IDW, OK, OCK are selected to analyze and compare, to explore the applicability of different methods and the degree of improvement of elevation factor. The temporal spatial interpolation method of rainstorm process to meet

the actual demand is selected to provide a reference for the simulation and research of rainstorm process in northern hilly area.

2 Research areas and data sources

2.1 Regional overview of the study

The elevation of Licheng District of Jinan City decreases from south to north, as shown in Fig. 1, the south and middle are low hills, steep terrain, accounting for 83.4% of the total area of the whole area. Average annual rainfall in Licheng District is 665.7 mm, decreasing from south to north with the change of topography, and the annual distribution is concentrated in summer, and most of them are short duration heavy precipitation. There are 13 rainfall stations in Licheng District. The distribution map and list are shown in Fig. 1 and Table 1.

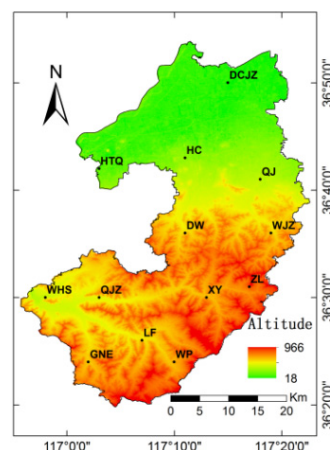


Fig. 1. DEM of LiCheng and Distribution of rainfall station.

Under the influence of topography, geomorphology and climatic conditions, the types of mountain torrents in Licheng District are mostly river flood disasters. The middle and upper reaches of the river are fast, Floods have risen sharply, the peak value is high and the peak time is short, The downstream terrain is gentle, often due to slow flood discharge, south to north flooding caused serious flood disaster.

Table 1. List of rainfall stations in Licheng, Jinan.

Station	East longitude	North latitude	absolute altitude(m)
HTQ	117° 03'	36° 42'	24
DW	117° 11'	36° 36'	201
HC	117° 11'	36° 43'	53
WJZ	117° 19'	36° 36'	251
QJ	117° 18'	36° 41'	96
SCJZ	117° 15'	36° 50'	23
ZL	117° 17'	36° 31'	377
XY	117° 13'	36° 30'	282
QJZ	117° 03'	36° 30'	162
WP	117° 10'	36° 24'	292
LF	117° 07'	36° 26'	221
GNE	117° 02'	36° 24'	252
WHS	116° 58'	36° 30'	163

2.2 data sources Formatting the text

According to the analysis of the hourly rainfall data of Licheng District from 2010 to 2014, the rainfall duration of Licheng District is relatively short, which is basically within 6 h. Taking into account the rainfall coverage and rainfall intensity standards, according to the principle that all rainfall stations have rainfall and at least one rainfall station meets the rainstorm standard (>16 mm/h), a total of 11 typical rainfall are selected. The duration and hourly rainstorm process are shown in **Table 2**, and 16 hourly rainstorm processes are obtained.

Table 2. Typical rainfall and the hourly rainstorm process.

NO.1	Typical rainfall	Duration	NO.2	Rainstorm process
A	20120708	5h	1	2012070801
			2	2012070802
			3	2012070803
			4	2012070804
			5	2012070805
B	20130608	2h	6	2013060805
			7	2013060806
C	20100617	1h	8	2010061720
D	20100820	1h	9	2010082003
E	20100907	1h	10	2010090712
F	20110702	1h	11	2011070213
G	20120710	1h	12	2012071004
H	20120818	1h	13	2012081818
I	20130713	1h	14	2013071309
J	20130716	1h	15	2013071603
K	20130812	1h	16	2013081200

3 Research method

3.1. Inverse Distance Weight

IDW, based on the principle of similar similarity, it is considered that the weight coefficient of the rainfall station near the center point of the grid is a power function of the distance between the rainfall station and the interpolation point. The rainfall at the center point of the grid is obtained by calculating the weighted average value and as the average surface rainfall⁴.

$$P_0 = \sum_{i=1}^n \frac{1}{(D_i)^n} P_i \left(\sum_{i=1}^n \frac{1}{(D_i)^n} \right)^{-1} \quad (1)$$

$$D_i = \sqrt{(x_0 - x_i)^2 + (y_0 - y_i)^2} \quad (2)$$

The P_0 is the estimated rainfall value; the P_i is the measured rainfall of the i (n) rainfall station; the D_i is the distance between the first rainfall station and the point to be evaluated; the n is the power of the inverse distance D_i , and the selection standard is to reach the minimum average absolute error.

3.2. Ordinary Kriging

OK is a method for unbiased optimal estimation of regionalization variables in a limited region based on the theory of differential function and structural analysis⁵. The applicable premise is that the regionalization variable has spatial correlation and satisfies the second order stationary hypothesis.

$$P^*(X_0) = \sum_{i=1}^n \lambda_i P(x_i) \quad (3)$$

The $P^*(X_0)$ is the estimated rainfall value, the $P(X_i)$ is the measured rainfall of the i ($i=1,2,\dots, n$) rainfall stations, and λ_i is its linear combination coefficient.

The λ_i criteria selected are:

unbiased: $E[P(X_0) - P^*(X_0)] = 0$;

Optimal: $Var[P(X_0) - P^*(X_0)] = \min$.

Combined with the second order stationary hypothesis:

$E[P(X)] = E[P(X + D)] = \bar{P}$ Calculated:

$$\begin{cases} \sum_{i=1}^n \lambda_i = 1; \\ \sum_{i=1}^n \lambda_i \gamma(x_i, x_j) + P = \gamma(x_i, x_0); \end{cases} \quad (4)$$

The γ represents a semi-variant function, $\gamma(h) = \frac{1}{2} E[P(X) - P(X + D)]^2$.

3.3. Co-Kriging

OCK is to optimize the estimation of the main variable on the basis of the common Kriging interpolation value, so as to improve the of interpolation accuracy⁶. Auxiliary variables need to be related to the main variable space and easy to obtain data. The set of cooperative regionalization variables can be represented by k sets of interrelated random functions $\{p_k(x), k = 1, 2, \dots, K\}$. Under the second order stationary hypothesis:

$$P_{V_{k_0}}^K = \sum_{k=1}^K \sum_{\alpha_k=1}^{n_k} \lambda_{\alpha_k} p_{\alpha_k} \quad (5)$$

The k_0 is $k=1,2,\dots$. The main variable to be estimated in the K regional variables, rainfall, the $P_{V_{k_0}}^K$ is the synergistic Kriging estimate of rainfall, and the $p_{\alpha_k} (\alpha_k = 1, 2, \dots, n_k)$ is the measured rainfall in the region. The

λ_{α_k} is the synergistic Kriging weighting coefficient to be determined, can be obtained by Lagrange multiplier method under unbiased and optimal conditions⁷:

$$\begin{cases} \sum_{k'=1}^K \sum_{\alpha_{k'}=1}^{n_{k'}} \lambda_{\beta_{k'}} \bar{C}_{k'k} (v_{\beta_{k'}}, v_{\alpha_k}) - \mu_k \\ = \bar{C}_{k_0,k} (V_{k_0}, v_{\alpha_k}) \quad \forall \alpha_k = 1, 2, \dots, n_k \\ \sum_{\alpha_{k_0}=1}^{n_{k_0}} \lambda_{\alpha_{k_0}} = 1 \\ \sum_{\alpha_k=1}^{n_k} \lambda_{\alpha_k} = 0 \quad \forall k \neq k_0 \end{cases} \quad (6)$$

The μ_k is Lagrange multiplier, the $\bar{C}_{a,b}$ is the mean value of mutual covariance, and the relationship between mutual variation function and mutual covariance is as follows:

$$\gamma_{k'k}(h) = C_{k'k}(0) - \frac{1}{2} [C_{k'k}(h) + C_{kk'}(h)] \quad (7)$$

3.4. Study on Spatial Interpolation of Time-by-Time Rainstorm in Regions

Taking the ArcGIS as the processing platform, the IDW, the OK method and the OCK method provided by the Geostatistical Analyst are used to analyze the data and adjust the parameters of 16 hourly rainstorm processes in Licheng District of Jinan City.

IDW mainly depends on the power value and the range of the search neighborhood. The most reasonable result can be obtained by taking the power value 0.5-3, further optimizing and determining the power value by minimizing the root mean square prediction error, and adjusting the smoothing coefficient and the short and short half axes according to the measured data with smoothing as the neighborhood type. Part of the results are shown in Fig. 2.

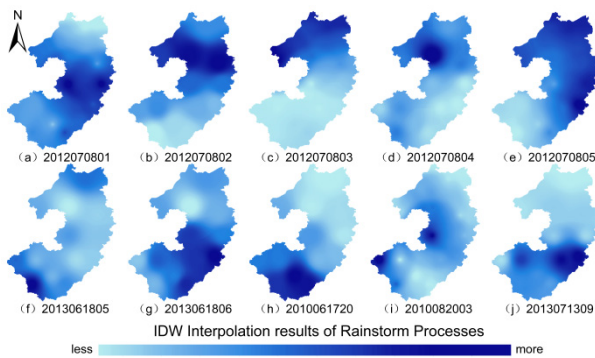


Fig. 2. Partial IDW results of Rainstorm Processes.

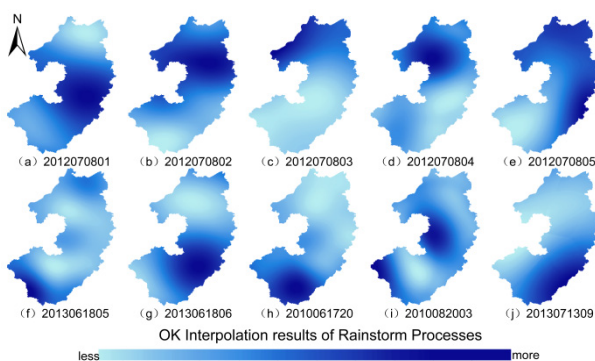


Fig. 3. Partial OK results of Rainstorm Processes.

The semi-variable function can check and quantify the spatial autocorrelation. In the case of real but difficult to calculate the semi-variable function directly, the semi-variable function is usually modeled, that is, the parameters of the mathematical curve are best fitted with the measured data.

OK is based on the semi-variable function cloud diagram, the fitting model is combined with the empirical semi-variable function, and the best fitting of the semi-variable function model is used to optimize the parameters such as step size, block gold, partial base table and variable path. Taking smoothing as the neighborhood type, the smoothing coefficient and the long and short half axis are adjusted. Some of the results are shown in Fig. 3.

Combined with the characteristics of annual average rainfall decreasing with elevation from south to north, elevation is selected as the auxiliary variable of rainfall. Cross covariance cloud map is used to analyze the cross correlation between rainfall and elevation at each rainfall station⁸. Different semi-variable function models of rainfall and elevation are fitted respectively, and the most suitable data structure is selected to couple the model and adjust the parameters to achieve the best fitting. Taking smoothing as the neighborhood type, the smoothing coefficient and the long and short half axis are adjusted. Some of the results are shown in Fig. 4.

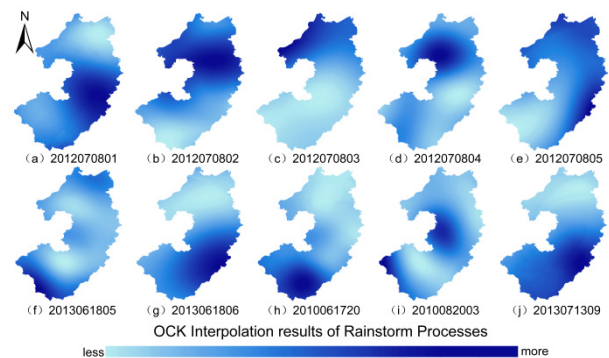


Fig. 4. Partial OCK results of Rainstorm Processes.

4 Results Analysis and Discussion

4.1. Analysis of interpolation results

The interpolation results of the 2012070801 and 2013061806 rainstorm processes are shown in Fig. 5. As can be seen, IDW, OK, OCK considering elevation can roughly reflect the spatial distribution of rainfall in the studied region. Among them, IDW the "bull's eye" at the rain station, There are more rainfall mutations in space, The change process is rough; OK、OCK reduced the "bull's eye" phenomenon in IDW, Equilibrate the maximum minima in the region, have more smooth performance. The OCK is slightly inclined to the highest elevation (southeast) relative to the OK rainstorm center, Rainfall is also lower at the lowest elevation (northern). It is consistent with the characteristics of annual average rainfall decreasing with elevation from south to north.

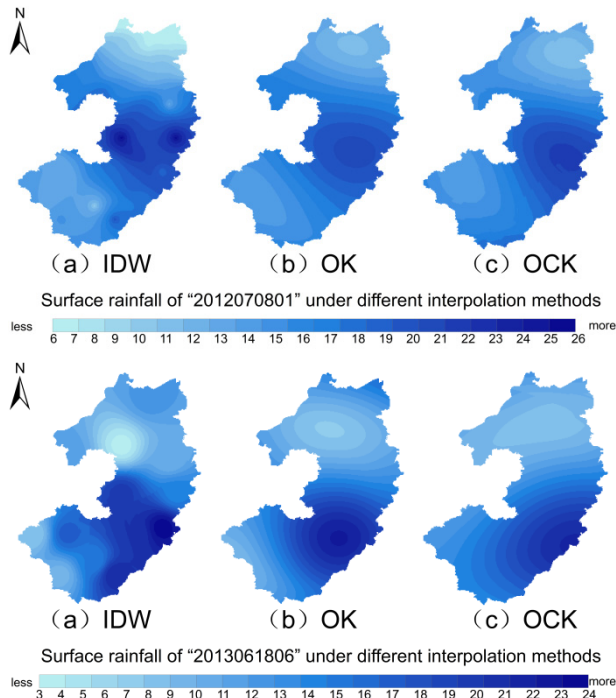


Fig. 5. Partial results comparison chart of Rainstorm Processes under different interpolation methods.

4.2. Cross validation analysis

The interpolation results are analyzed by cross verification: one known point is omitted as the verification data in order, and the other known points are used to predict the point, so as to obtain the error between all known points and their predicted values. The average absolute error (MAE) and root mean square error (RMSE) are usually used to reflect the actual situation of the predicted error and the degree of dispersion respectively.⁹

$$MAE = \frac{1}{n} \sum_i^n |X_i - x_i| \quad (8)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_i^n (X_i - x_i)^2} \quad (9)$$

The MAE is the average of the absolute error, RMSE the deviation between the predicted value and the measured value.

The error results of interpolation of 16 rainstorms are shown in Table 3 and Fig. 6.

Table 3. Error analysis table of cross validation results.

NO.2	Rainstorm process	MAE(mm)		
		IDW	OK	OCK
1	2010061720	5.204	4.789	4.254
2	2010082003	6.419	5.634	5.035
3	2010090712	5.022	5.003	4.734
4	2011070213	8.816	8.321	7.743
5	2012070801	4.408	4.129	4.056
6	2012070802	4.060	3.220	2.605
7	2012070803	4.439	3.668	3.607
8	2012070804	8.110	7.044	6.334
9	2012070805	2.812	2.128	1.532
10	2012071004	5.149	4.909	4.041
11	2012081818	7.805	5.751	4.741
12	2013061805	4.832	4.306	3.847
13	2013061806	4.600	4.145	3.980

14	2013071309	4.369	3.784	3.246
15	2013071603	8.158	7.348	6.472
16	2013081200	5.686	5.704	5.454
	mean value	5.618	4.993	4.480
NO.2	Rainstorm process	RMSE(mm)		
		IDW	OK	OCK
1	2010061720	5.781	5.577	4.988
2	2010082003	7.821	7.197	6.689
3	2010090712	6.792	6.246	6.080
4	2011070213	11.493	10.772	10.605
5	2012070801	5.579	5.249	5.024
6	2012070802	5.317	4.111	3.839
7	2012070803	6.544	6.011	5.918
8	2012070804	10.903	9.547	8.911
9	2012070805	3.531	2.409	1.724
10	2012071004	6.801	5.981	5.610
11	2012081818	9.989	7.771	6.774
12	2013061805	6.135	5.630	5.344
13	2013061806	5.240	4.659	4.341
14	2013071309	5.492	4.735	4.413
15	2013071603	10.246	8.538	7.648
16	2013081200	7.752	7.718	6.710
	mean value	7.214	6.384	5.914

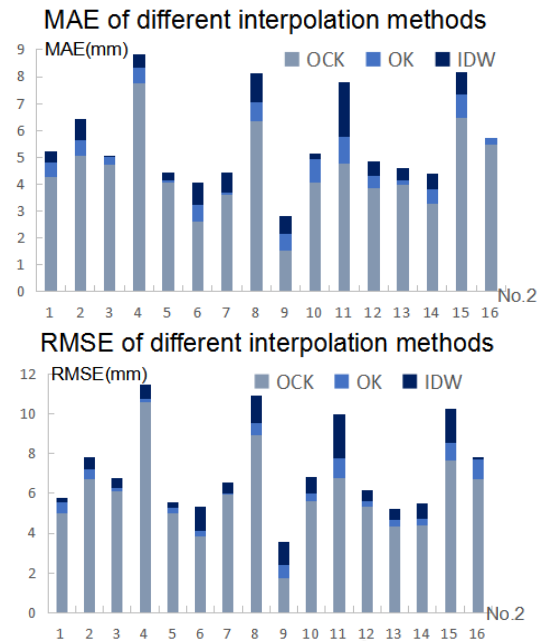


Fig. 6. Error analysis chart of cross validation results.

According to the chart above:

(1)Average absolute error: except that the error IDW during the 16th rainstorm is slightly less than OK, IDW>OK>OCK; for the remaining 15 Compared to IDW, The error between OK and OCK was reduced by 11.12% and 20.26% ,respectively. Therefore,IDW prediction error range is maximum, OCK minimum.

(2)Root mean square error:16 rainstorm processes are all IDW>OK>OCK,OCK,OK 18.02% and 11.51% lower than IDW, the OCK prediction deviation is smaller, the dispersion degree is smaller and the overall stability is better.

4.3. Discussion

According to the order of applicability of three spatial

interpolation methods for hourly rainstorm in Licheng District of Jinan City, OCK>IDW, are discussed as follows:

(1) The IDW error is obviously large because it only considers distance in the interpolation process, and has a higher dependence on the power value of distance and the search neighborhood, in which the power reciprocal of distance can not fully reflect the correlation of rainfall space. Therefore, the interpolation surface mutation is more, the error is larger.

(2) The spatial distribution of rainfall can be described more accurately by using the best fitting model of semi-variable function to quantify the autocorrelation of the data OK compared with the IDW. Therefore OK it has better overall correlation, avoids the phenomenon that the sampling point must be the extreme point, the surface of the interpolation result is smoother, and the prediction of the extreme point position (rainstorm center) is more accurate and the error is relatively small.

(3) OCK using elevation as auxiliary information and based on the cross-correlation analysis of rainfall and elevation, the semi-variant function model of rainfall and elevation is coupled to make the predicted rainstorm center more reasonable and more consistent with the actual situation in the research area.

5 Conclusion

Taking Licheng District of Jinan City as the research area, three different spatial interpolation methods are used to calculate the time-by-time spatial interpolation of 16 rainstorm processes, and the interpolation results are analyzed and the analyzed. The main conclusions are as follows:

(1) The Inverse Distance Weight has many mutations in space, and the phenomenon of "bull's eye" at rainfall station is obvious; the Ordinary Kriging and the Co-Kriging considering altitude equalizes the spatial mutation, which is relatively smooth on the whole and more reasonable in predicting the center position of rainstorm. The latter correlates rainfall with elevation, and interpolated surfaces are more realistic.

(2) The mean values of the mean absolute errors (MAE) of Inverse Distance Weight, Ordinary Kriging, Co-Kriging considering elevation are 5.618, 4.993, 4.480, respectively. The mean values of root mean square error (RMSE) are 7.214, 6.384, 5.914, respectively. The prediction error range of Co-Kriging interpolation considering elevation is the smallest, the dispersion is the smallest, the overall stability is the best, and the accuracy is the highest. Therefore, for the short duration rainstorm process in the northern hilly area with topographic undulation, the Co-Kriging considering elevation describes the spatial distribution of rainstorm better, the accuracy is higher, the Ordinary Kriging is the second, and the Inverse Distance Weighted interpolation method is poor.

Acknowledgments

Special project of central government guiding local

science and technology development (Z13505009017); Provincial water conservancy research and extension project in Shandong Province (SDSLKY201815, SDSLKY201222, SDSLKY201807).

References

1. LI Jinjie, WANG Aihui. 2019. Comparison of Spatial Interpolation Methods Based on Monthly Precipitation Observation Data of Station in Southwest China[J]. Climatic and Environmental Research (in Chinese), 24(1):50-60.
2. FANG Lindong, LIAO Weihong, WANG Mingyuan, SONG Wanzhen. Research of the Inverse Distance Weighted Interpolation Method Considering the Elevation. Yellow River, 2015, 37(9):38-41.
3. Tobin C, Nicotina L, Parlange M B, et al. Improved interpolation of meteorological forcings for hydrologic applications in a Swiss Alpine region[J]. Journal of Hydrology, 2011, 401(1-2):77-89.
4. HE Hongyan, GUO Zhihua, XIAO Wenfa. Review on spatial interpolation techniques of rainfall[J]. cje, 2005, (10): 1187-1191.
5. Feng Jingming, Zhao Tianbao, Zhang Yingjuan. 2004. Intercomparison of Spatial Interpolation Based on Observed Precipitation Data[J]. Climatic and Environmental Research (in Chinese), 9(2):261-277.
6. DENG Yu, Liu Shenghe, YAO Fengfeng, WANG Haijun. Standard Land Price Appraisal and Space Structure Analysis Based on Co-Kriging[J]. PROGRESS IN GEOGRAPHY, 2009, 28(3): 403-408.
7. Hou Jingru. Practical geostatistics [M]. Geological Press, 1998
8. Niu Hailing. Study on evolution characteristics of spatial pattern of China's oil resource flow [D]. Nanjing Normal University, 2011
9. Gu Chunlei, Yang Yang, Zhu Zhichun. Cross validation of interpolation accuracy of several DEM models [J]. Surveying and spatial geographic information, 2011 (05): 111-114