ISSN 2616-5872 Vol.4, Issue 8: 35-40, DOI: 10.25236/AJEE.2022.040807

Analysing the characteristics of precipitation variation in Hefei in recent years based on multiple linear regression

Pan Chunhua

College of Geography and Environmental Sciences, Zhejiang Normal University, Jinhua, China

Abstract: *Based on the multiple linear regression model, the precipitation change characteristics of Hefei from 2009 to 2019 are analyzed, and the long-term precipitation data show that the annual precipitation in northwest China, Jianghuai and eastern coast of China is increasing, and the annual precipitation in North China and central China is decreasing, and the precipitation change between regions is obviously different. In recent decades, extreme precipitation events have shown a frequent trend in China, especially in the northwest and southeast coasts and the middle and lower reaches of the Yangtze River, which not only increases the frequency of extreme precipitation but also increases the intensity of precipitation. In this context, the characteristics of precipitation change in Hefei from 2009 to 2019 were studied, and the results showed that the change of precipitation in Hefei was related to the minimum temperature and year. The change of precipitation is obvious, there is an interannual and interdecadal variation law of fluctuation change, the fluctuation amplitude increases, the uncertainty of precipitation increases, and the lowest temperature has a negative correlation with the annual precipitation, that is, the lower the lowest temperature, the greater the precipitation, which is consistent with the change trend of precipitation in the entire Jianghuai River Basin. In order to provide reference for meteorological research and the prevention of meteorological disasters.*

Keywords: *Multiple linear regression model, SPSS, Precipitation, Hefei City*

1. Introduction

Water resources are basic natural resources and strategic economic resources[1], and atmospheric precipitation, as the most important part of the water cycle, plays an important role in surface runoff and water availability. Although the national average precipitation has not seen obvious trend changes in the past hundred years and nearly 60 years, there has been a clear decadal change after 1970[2]. Since the beginning of the 21st century, with the global warming, the precipitation characteristics of various regions have also undergone different responses. Under such a premise, in-depth study of the characteristics and mechanisms of regional climate change and then prediction of its future precipitation trend is an important research content of global and regional climate change. Ren Guoyu et al. [3] show that the average precipitation in the middle and lower reaches of the Yangtze River has shown a clear trend of increase throughout the year and in summer (since 1951). Wang Sheng et al. [4] concluded that extreme precipitation events in Anhui Province have an increasing trend, and the frequent phenomenon of drought and flood disasters has become more and more obvious since the 90s of the 20th century. Gu Junqiang et al. [5] pointed out that while the annual precipitation in Zhejiang Province increased, the number of rainfall days did not increase equally.

Jiang Junjie et al. [6] used the daily precipitation data of Anhui Province from 1959 to 2007, and analyzed the trend and change characteristics of nearly 49a precipitation by linear regression method, Mann-Kendall method, cumulative distance flat method and Morlet wavelet analysis method. The results show that there is a significant increase trend of summer and winter precipitation in Anhui (significance<0.05), and the change trend of spring, autumn and annual precipitation is not significant. There is a significant decrease in the total number of rainfall days in the whole region (significant < 0.01), but the number of heavy rainfall days has increased. The total number of rainy days decreased from south to north, but the climatic tendency rate of heavy rainfall days tended to increase from south to north. There are roughly two main cycles of annual precipitation and four-season precipitation in the province, and the shock trend of the 2-4A cycle is gradually obvious. In Anhui, the number of rainy days decreased, heavy precipitation increased, and precipitation tended to be concentrated. After the mid-90s of the 20th century, the trend of wide oscillation of precipitation became prominent, and the uncertainty increased, resulting in a significant increase in the possibility of drought, floods, landslides, debris flows and other disasters, which will adversely affect local human production and life, such as agricultural production.

Under the climate background of global warming, the frequency and intensity of extreme precipitation events have been greatly enhanced, so it is of great significance for the study of water vapor and precipitation. Anhui Province is located in the Yangtze River Delta region, bordering the river and the sea, with the Yangtze River waterway inside and the coastal zone outside. Hefei, the capital of Anhui Province, is located between 30°56'-32°33'N latitude and 116°40'-117°58'E longitude, located in the middle of Anhui, between Jianghuai and the west wing of the Yangtze River Delta. Hefei bears the important position of the central city of Hefei metropolitan area, the sub-central city of China's Yangtze River Delta urban agglomeration approved by the State Council, the core city of the Anjiang urban belt, the central city of the G60 science and technology innovation corridor, the "Belt and Road" and the strategic dual-node city of the Yangtze River Economic Belt, with the continuous expansion of the city scale and the rapid increase of the permanent population, as of the end of 2021, the permanent population of Hefei is 9.465 million, and the urbanization rate has reached 84.04%. It has had a significant impact on Hefei's urban ecological environment and regional climate, among which disasters such as dam failure and flooding caused by precipitation and urban waterlogging have not only caused greater economic losses, but also seriously affected people's normal life.

There are many domestic research literature on precipitation changes[7-11], and the average annual precipitation in Daxi'an showed a downward trend, and the spatial distribution showed a gradual decrease from southeast to northwest[12]. Li Na et al.[13] found that the average annual precipitation in Xi'an has shown a downward trend in the past 38 years.

In summary, most of the data series used in the studies are old, can not reflect the latest changes in precipitation, and the research results are uncertain, so the latest precipitation data are used to analyze the temporal and spatial changes of precipitation in Hefei, and the law of precipitation changes is found and summarized. At present, there are few regional precipitation studies in Hefei City, Anhui Province, and there is no systematic analysis and summary. Anhui Province is an important grain-producing area, but the span between north and south is large, and the temporal and spatial distribution of precipitation varies greatly, resulting in frequent droughts and floods, resulting in agricultural economic losses and even casualties. Therefore, understanding and mastering the temporal trend and spatial distribution characteristics of precipitation in this region is of great guiding significance for analyzing precipitation changes under the background of climate warming, rationally planning agricultural production in the region, and actively preventing and mitigating disasters.

Therefore, on the basis of referring to previous research results, this paper intends to select Hefei City in central Anhui Province as the research area to study the characteristics of annual precipitation changes in Hefei City, focusing on the potential correlation between precipitation in Hefei City, the capital of Anhui Province, and its average temperature, minimum temperature and year. In this paper, multivariate statistical analysis is taken as the basis and premise, and a series of methods are verified by combining relevant data, and the influence of the results on the overall multiple linear regression equation is verified. Through the combination of other test methods, the test results are more realistic, the accuracy and efficiency are improved, and the original regression results can be optimized to the greatest extent.

2. Multiple linear regression analysis of precipitation in Hefei

2.1 Raw data

The experimental data used in this paper are precipitation, annual average temperature and minimum temperature data of Hefei from 2009 to 2019 (see Table 1), with a total of 11 samples, and the relevant factors are assumed: among them, precipitation is used as the dependent variable, the annual average temperature, the lowest temperature and the year are used as the independent variables, the annual precipitation is Y, the lowest temperature is X1, the year is X2, and the average temperature is X3, and a multiple regression model is established to analyze the influence of various factors on the precipitation Y in Hefei.

It is assumed that the regression model form that meets the data of this experiment is:

$$Y=\beta_0+\beta 1X1+\beta_2X2+\mu$$

where is the constant term and is the error term.

ISSN 2616-5872 Vol.4, Issue 8: 35-40, DOI: 10.25236/AJEE.2022.040807

Table 1: Some meteorological data of Hefei from 2009 to 2019

year	Average temperature(℃)	Annual precipitation(mm)	Lowest temperature(℃)
2009	16.78	504	-8
2010	16.68	259.1	-4
2011	16.41	504.9	-7
2012	16.60	499.6	-6
2013	17.23	426.5	-6.6
2014	16.55	609.4	-6.4
2015	16.78	624.5	-4.6
2016	17.05	741.2	-8.9
2017	17.18	481.3	-4.5
2018	17.07	1052.5	-9.5
2019	16.48	579	-5.1

2.2 Regression analysis

Table 2: Goodness-of-fit test

	Statistics						
model	R	R 2	F	df1	df2	Sig. F	Durbin-Watson
1	0.918	0.843	12.508	3	7	0.003	2.647

a. Predictors: (constant), lowest temperature, year, average temperature

b. Dependent variable: Annual precipitation

The goodness of fit test is one of the important contents of statistical significance test with Chi-square statistics. Table 2 shows a good goodness of fit.

Table 3: ANOVA results

model		Sum of squares	df	mean square	F	Sig.
1	regression	341643.747	3	113881.249	12.508	0.003 b
	Residuals	63731.382	7	9104.483		
	total	405375.129	10			

a. Dependent variable: Annual precipitation

b. Predictors: (constant), lowest temperature, year, average temperature

Meanwhile, ANOVA is used to analyze the relationship between classified data and quantitative data. The significance in Table 3 is less than 0.05.

Table 4: Factors

model		The coefficients are not standardized		Normalization coefficients	t	Saliency	Colinearity statistics	
		B	Standard error	Beta			Tolerance	VIF
1	(constant)	-73544.012	18700.088		-3.933	0.006		
	year	37.237	9.508	0.613	3.916	0.006	0.916	1.092
	Average temperature (℃)	-81.705	112.060	-.118	-.729	0.490	0.857	1.167
	Lowest temperature	-76.767	17.133	-.696	-4.481	0.003	0.932	1.073

a. Dependent variable: Annual precipitation

In regression analysis, if there are two or more independent variables, it is called multiple regression. In fact, a phenomenon is often associated with multiple factors, and it is more efficient and realistic to predict or estimate the dependent variable by the optimal combination of multiple independent variables than to predict or estimate with only one independent variable. Therefore, multiple linear regression is

ISSN 2616-5872 Vol.4, Issue 8: 35-40, DOI: 10.25236/AJEE.2022.040807

of greater practical significance than univariate linear regression. This article has several independent variables, choosing multiple linear regression. Select the software to build the model as SPSS, import the data into SPSS to select analysis, regression, linearity, and check the corresponding indicators and detection methods.

It can be seen from the table:

(1) R2: Indicates the fit of the model, the closer to 1, the better. The R2=0.843 of the three independent variables and the dependent variable in the multiple linear regression model correspond to a significant value of 0.003, indicating that the established model fits well.

(2) Durbin-Watson test:DW inspection for short，It is by far the most commonly used method for testing autocorrelation.

In general, the DW test values are distributed between 0-4, when the DW test value is significantly close to 0 or 4, there is an autocorrelation, the closer to 2, the absence of (first-order) autocorrelation, the greater the possibility that the observations are independent of each other, the less obvious the autocorrelation of the independent variables, and the better the model design. In this study, the DW test value was 2.647, indicating that the observations of multiple linear regression in this study were highly independent of each other, and the factors did not affect each other.

(3) Significance: As can be seen from Table 4, the significance of the independent variable year and the lowest temperature are 0.006 and 0.003, respectively, while the significance of the average temperature is 0.490, obviously, the average temperature has little effect on the dependent variable.

(4) Correlation coefficient and tolerance/variance inflation factor (VIF) are used to test for multicollinearity. The variance inflation factor is the reciprocal of tolerance (1/tolerance). VIF is used for collinearity diagnosis (correlation between variables): when 0<VIF<10, there is no multicollinearity; When 10≤VIF < 100, there is a strong multicollinearity; When VIF≥ 100, there is severe multicollinearity. Therefore, according to the above table, it can be found that the VIF of the year, the lowest temperature, and the average temperature are all less than 10, so there is no multicollinearity.

2.3 Gradual regression

Table 5: Variables entered and removed

model	The input variable	The removed variable	method
1	Lowest temperature	.	Step (condition: probability of F to enter < = 0.050, probability of F to be removed >= 0.100).
2	year	.	Step (condition: probability of F to enter < = 0.050, probability of F to be removed >= 0.100).

a. Dependent variable: Annual precipitation

From the above analysis, it can be seen that the fitness of the multiple linear regression model is good, and the significant correlation is shown after the F test, and the SPSS stepwise regression procedure is used to test it to exclude the interaction between the dependent variables. From the input and output variables in Table 5, it can be seen that the probability of F of the lowest temperature input of the independent variable is less than or equal to 0.05, and for the independent variable year, the probability of F of the input is less than or equal to 0.05.

Table 6: Model summary

mode 1	R	R 2	Adjusted R2	Errors in standard estimates	Durbin Watson
1	0.703a	0.495	0.439	150.86467	
2	0.912b	0.831	0.789	92.58199	2.656

a. Predictors: (constant), lowest temperature
b. Predictors: (constant), lowest temperature, year
c. Predicks: (Constance), Lowst Tempe Latour, Hyères

In Table 6, the model summary is 0.495 for model 1, 0.439 for adjusted R, 0.831 for model 2, and 0.789 for adjusted R.

Table 7: Stepwise regression coefficients

model		The coefficients are not standardized		Normalization coefficients	t	Saliency	Colinearity statistics	
		B	Standard error	Beta			Tolerance	VIF
1	(constant)	72.852	173.908		0.419	0.685		
	Lowest temperature	-77.629	26.153	-0.703	-2.968	0.016	1.000	1.000
2	(constant)	-70927.979	17807.283		-3.983	0.004		
	Lowest temperature	-73.602	16.081	-0.667	-4.577	0.002	0.996	1.004
	year	35.266	8.845	0.581	3.987	0.004	0.996	1.004

a. Dependent variable: Annual precipitation

As can be seen from Table 7, the significance of model 2 is more obvious, respectively: 0.004, 0.002, 0.004, indicating that the coefficients of each variable are very significant and statistically significant, and the stepwise regression equation can be derived from the above analysis: Y=-73.602X1+35.266X2-70927.979

3. Conclusion

In summary, from the above calculation process data analysis to obtain the regression model of Hefei's annual precipitation, according to the regression model, it can be seen that the factors affecting precipitation and fixing other factors meet the significance of the independent variable on the dependent variable, and show the criterion of linear correlation. The multiple linear relationship between annual precipitation in Hefei and the year and minimum temperature can be expressed by the equation: Y=-70927.979-73.602X1+35.266X2 (Y is the annual precipitation, X1 is the lowest temperature, X2 is the year). Through the established multiple linear regression model of precipitation in Hefei from 2009 to 2019, the results show that the change of precipitation in Hefei is related to the minimum temperature and year. The change of precipitation is obvious, there is an interannual and interdecadal variation law of fluctuation change, the fluctuation amplitude increases, the uncertainty of precipitation increases, and the lowest temperature has a negative correlation with the annual precipitation, that is, the lower the lowest temperature, the greater the precipitation, which is consistent with the change trend of precipitation in the entire Jianghuai River Basin. And from the data, it can be found that as the minimum temperature drops by 1 degree Celsius, precipitation increases by about 73.602 mm. In fact, there are many factors that affect precipitation, and only the influence of time factors and temperature on it is analyzed here.

References

[1] Chen Lei. Strict water resources management to ensure sustainable development ——— written on the occasion of World Water Day and "China Water Week" in 2010[J]. China Water Resources, 2010(6): 3 -4.

[2] Ding Yihui, Wang Huijun. New understanding of scientific issues of climate change in China in the past 100 years [J]. Chinese Science Bulletin, 2016, 61(10): 1029-1041.

[3] REN Guoyu, WU Hong, CHEN Zhenghong. Spatial characteristics of precipitation trends in China [J]. Journal of Applied Meteorological Science, 2000, 11(3):322-330.

[4] Wang Sheng, Lu Jun, Wu Biwen, et al. Study on summer precipitation changes and its impact on drought and flood in Anhui Province [J]. Journal of Anhui Agricultural Sciences, 2008, 36(7): 2870-2873.

[5] Gu Junqiang, Shi Neng, Xue Genyuan. Climate change of precipitation and rainy days in Zhejiang Province in the past 40 years [J]. Journal of Applied Meteorological Science, 2002, 13(3):322-329.

[6] Jiang Junjie, Sun Weiguo. Analysis of temporal and spatial variation characteristics of precipitation in Anhui Province from 1959 to 2007[J].China Agrometeorology, 2012,33(01):27-33.

[7] HE Zhiming, DENG Shiru, WU Qiong, et al. Analysis of temperature and precipitation changes in

ISSN 2616-5872 Vol.4, Issue 8: 35-40, DOI: 10.25236/AJEE.2022.040807

Nanchang in 1963~2012[J].Journal of Meteorology and Environment, 2018,34(2):35-43.

[8] Ma Jianing, Gao Yanhong. Analysis of average annual precipitation and extreme precipitation changes in the upper reaches of the Yellow River in the past 50 years [J]. Plateau Meteorology, 2019, 38(1): 124-135.).

[9] WANG Tingting, FENG Qi, GUO Xiaoyan, et al. Precipitation variation characteristics and abrupt analysis in Gulang River Basin in 1959~2014[J].Plateau Meteorology, 2019,38(6):1251-1262.

[10] Wu Xiulan, Zhang Taixi, Wang Hui, et al. Analysis of regional climate change characteristics in Xinjiang in 1961~2017[J].Desert and Oasis Meteorology, 2020, 14(4):27-34.

[11] Xue Yuting, Li Xiehui, Wang Lei, et al. Spatial-temporal variation characteristics of different levels of precipitation in summer in southwest China in 1976~2017[J]. Journal of Southwest University: Natural Science Edition, 2022,44(2):137-145.).

[12] Li Xiaojuan. Journal of Xi'an University of Architecture and Technology: Natural Science Edition, 2012, 44(4):489-494.

[13] Li Na, Zhou Weibo. Spatial-temporal variation characteristics and abrupt analysis of precipitation in Xi'an[J]. Western Development: Land Development Engineering Research, 2017, 12(6): 58-63, 70.