Pisco Med Publishing

Original Research Article

Research on Rainfall Factors in Zhengzhou Based on Gray Correlation Analysis

Jianguang Xu, Canyu Guo

Hebei University of Engineering, Handan, Hebei, 056038

Abstract: Firstly,the Pearson correlation coefficients between rainfall and other meteorological indicators were calculated in this paper,and significance tests were performed to determine whether they were correlated and the strength of the correlation. At the same time,considering the small data samples and the existence of data errors,gray correlation analysis can be used,and then the main effects of dew point,air temperature and wind speed on the variation of precipitation can be derived.The years 2003,2004,2016,and 2018 with high precipitation were screened out through M-K mutation test,cluster analysis based on the characteristics of total annual precipitation and number of days of rain during the year.

Keywords: Rainfall factors; Gray correlation analysis;M-K mutation test

1. Introduction

Many parts of Henan were severely affected by the historically rare heavy rainfall in July 2021.Among them,Zhengzhou suffered an extraordinarily heavy rainfall,which is the scale of a once-in-a-millennium event.All of the single-day precipitation on July 20 broke the 60-year record since Zhengzhou established a weather station in 1951,reaching the extreme value of hourly rainfall on land in China[1].The extreme rainfall brought huge losses and disasters to Zhengzhou,which has a population of tens of millions-subways were flooded,stations leaked,streets turned into rivers,and flooding and secondary disasters from the rainstorm have posed serious threats to the lives,safety and property of local people[2].

2. Establishment of precipitation factor model

It is for us to analyze the data in the annexes and analyze the Pearson correlation coefficients between rainfall and other meteorological indicators in two counterparts and perform significance tests to determine whether they are correlated and the strength of the correlation[3].At the same time,considering the small data sample and the existence of data errors,grey correlation analysis can be used,which in turn leads to the main effects of dew point,air temperature and wind speed on the variation of precipitation.The years with higher precipitation are filtered out by M-K mutation test,cluster analysis based on the total annual precipitation,the number of days of rain in the year and other characteristics.

A correlation is a non-deterministic relationship,and the correlation coefficient is a quantity that studies the degree of linear correlation between variables.Due to the different objects of study,the correlation coefficient is defined in several ways as follows.

Pearson's correlation coefficient:also known as the correlation coefficient or linear correlation coefficient,generally denoted by the letter r,is used to measure the linear relationship between two variables with the following formula:

$$r=\frac{n\sum xy-\sum x\sum y}{\sqrt{n\sum x^2-(\sum x)^2}\sqrt{n\sum y^2-(\sum y)^2}} \tag{1}$$

where the value of the correlation coefficient r is between-1 and+1,i.e.-1≤r≤+1.It can generally be classified in three levels:|r|<0.4 for low linear correlation;0.4≤|r|<0.7 for significant correlation;and 0.7≤|r|<1 for high linear correlation.

To further investigate the correlation between precipitation and other meteorological indicators,we used grey correlation analysis,first we manually processed the data in the annexes,and the meteorological indicators in the annexes were DEWP,FRSHTT, GUST,MAX,MIN,MXSPDSLP,SNDP,ISTP,TEMP,VISIB,WDSP twelve influential factors,which are mean dew point,occurrence of unusual hazards,gust wind speed,maximum temperature,minimum temperature,maximum sustained wind speed,precipitation,mean sea level pressure,snow depth,mean station pressure,mean temperature,mean visibility,and mean wind speed.

The formula for calculating the grey correlation coefficient is as follows[4].

$$\xi_i(k)=\frac{\underset{s}{min}\underset{t}{min}|x_0(t)-x_s(t)|+\rho\underset{s}{max}\underset{t}{max}|x_0(t)-x_s(t)|}{|x_0(k)-x_i(k)|+\rho\underset{s}{max}\underset{t}{max}|x_0(t)-x_s(t)|} \tag{2}$$

The above equation is defined as the grey coefficient,i.e.the correlation coefficient of each influencing factor is:

doi: 10.18282/l-e.v10i6.2790

$$y(x_0(t), x_i(t)) = \frac{a + \rho b}{|x_0(t) - x_i(t)| + \rho b} \tag{3}$$

where a is the bipolar minimum difference,b is the bipolar maximum difference,and ρ is the resolution factor.

$$a = \max_i \max_t |x_0(t) - x_i(t)| \tag{4}$$

Define the gray correlation $y(x_0, x_i)$,i.e.,find the mean value of each column of the matrix for which the correlation coefficients are obtained.

$$y(x_0, x_i) = \frac{1}{n}\sum_{k=1}^{n} y(x_0(t), x_i(t)) \tag{5}$$

The influence of the correlation of different meteorological indicators on precipitation can be derived by comparing the grey correlation.

3. Model solving

The Pearson correlation coefficients were derived from the SPSS analysis of precipitation and other meteorological indicators in the annex,which yielded the following results.

Tab1 Correlation data of precipitation PRC and dew point DEWP.

		PRCP	DEWP
PRCP	Pearson Correlation	1	.069**
	Sig.(bobtail)		.000
	Number of cases	19447	19447
DEWP	Pearson Correlation	.069**	1
	Sig.(bobtail)	.000	
	Number of cases	19447	19447

**.Significant correlation at the 0.01 level(two-tailed).

By analyzing the data in Table 1,we can conclude that the Pearson correlation coefficient is close to 1 and the two-tailed is at the level of 0.01,we can conclude that the precipitation PRC is significantly correlated with the dew point DEWP.

Tab2 Data table of correlation between precipitation PRC and temperature TEMP.

		PRCP	TEMP
PRCP	Pearson Correlation	1	.029**
	Sig.(bobtail)		.000
	Number of cases	19447	19447
TEMP	Pearson Correlation	.029**	1
	Sig.(bobtail)	.000	
	Number of cases	19447	19447

**.Significant correlation at the 0.01 level(two-tailed).

Analysis of the data in Table 2 shows that the Pearson correlation coefficient is close to 1 and the two-tailed is at the level of 0.01,we can conclude that the precipitation PRC is significantly correlated with the temperature TEMP.

Tab3 Correlation data table between precipitation PRC and gust wind speed WDSP.

		PRCP	WDSP
PRCP	Pearson Correlation	1	.022**
	Sig.(bobtail)		.002
	Number of cases	19447	19447
WDSP	Pearson Correlation	.022**	1
	Sig.(bobtail)	.002	
	Number of cases	19447	19447

**.Significant correlation at the 0.01 level(two-tailed).

By analyzing the data in Table 3,we can conclude that the Pearson correlation coefficient is close to 1 and the two-tailed is at the level of 0.01,we can conclude that the precipitation PRC is significantly correlated with the wind speed WDSP.

The magnitude of Pearson's correlation coefficient led to the conclusion that dew point,air temperature and wind speed play a major role in precipitation variation and passed the significance test.

According to the grey correlation analysis can be obtained according to the grey correlation of each indicator in the sub-series ar e:09892,0.9894,0.9888,0.9889,0.9890,0.9892,0.9888,0.9888,0.9888,0.9889,0.9887,0.9894,comparing the magnitude of the scores of different indicators can be the following table.

Tab4 Table of grey correlation between precipitation and other meteorological indicators

targets	connotation	gray correlation
DEWP	Average dew point	0.9892

FRSHTT	Occurrence of unusual disasters	0.9894
GUST	gust wind speed	0.9888
MAX	highest temperature	0.9889
MIN	minimum temperature	0.9890
MXSPD	Maximum sustained wind speed	0.9892
SLP	precipitation(meteorology)	0.9888
SNDP	snow depth	0.9888
STP	Average station pressure	0.9888
TEMP	average temperature	0.9889
VISIB	Average visibility	0.9887
WDSP	Average wind speed	0.9894

The ranking of the above data table of grey correlation between precipitation and other meteorological indicators shows that FRSHTT occurrence of anomalous hazard,WDSP average wind speed,and DEWP average dew point have the greatest influence on precipitation with grey correlation of 0.9894,0.9894,and 0.9892 respectively.

To observe the trend of precipitation at different meteorological monitoring points in Zhengzhou,we performed a linear fit to the precipitation data collected at meteorological stations 1,2 and 3 from 1984 to 2021,respectively,as shown in the following figure.

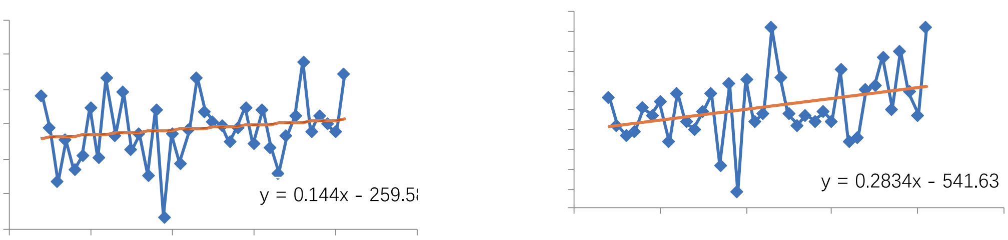

Fig1 Linear Regression Plot of Zhengzhou Weather Station 1(left)and Station 2(right),1984-2021.

Observing the trend of annual precipitation changes in three meteorological stations in Zhengzhou,we find that the precipitation from 1984 to 2021 all show an upward trend.The linear regression equation of station 1 is:y=0.144x-259.58;the linear regression equation of station 2 is:y=0.2834x-541.63;the linear regression equation of station 3 is:y=0.1203x-215.61.The regression coefficients of all three regression equations are greater than zero,so the annual precipitation of Zhengzhou shows an upward trend.

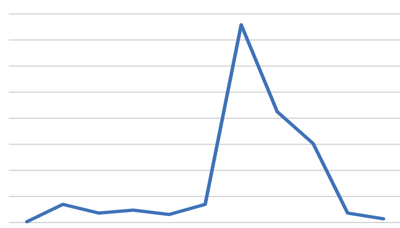

Fig2 The month-by-month precipitation folding statistics for 2021 in Zhengzhou.

By plotting the month-by-month precipitation folding statistics for 2021 in Zhengzhou,we find that the peak precipitation in 2021 was reached in July 2021,and the precipitation in August 2021 declined,but was still higher than the precipitation in other months.the single day precipitation on July 20 had both broken the 60-year historical record since the establishment of the weather station in Zhengzhou in 1951.At the same time,the heavy rainfall in Zhengzhou lasted for many days.In the previous rainfall history of Zhengzhou,the heaviest rainfall was at 3pm on July 3,1986,and the cumulative six-hour precipitation for that time was as high as 201mm.

References:

[1] Zhou S.F.,Liu T.X.,Duan L.M.,Zhang W.R.,Ji R.,Sun L..Analysis of simulated precipitation prediction based on time series in the Hailiu Rabbit River basin[J].Soil and water conservation research,2021,28(05):88-94.

[2] Wu Wanqin,Qian Hong.Research on the prediction analysis of monthly precipitation in Kunming city--based on SARIMA model and Holt-Winters summation model[J].Journal of Yunnan University for Nationalities(Natural Science Edition),2021,30(04):365-370.

[3] Zhou Xiaoxu.Application of hierarchical clustering-based LSTM neural network model for precipitation prediction in Jiangsu Province[D]. Shandong University,2020.

[4] Zhang Caihong.Analysis and prediction of precipitation trends in Weinan City based on ARIMA model[J].Value Engineering,2019,38(34):197-199.