

THE EFFECTS OF WATERSHED FLOOD PEAKS ON THE HYDROLOGY AND WATER RESOURCES BASED ON THE LOGNORMAL DISTRIBUTION MODEL

WANG, Y.^{1,2} – FENG, M. Q.^{1*} – ZHANG, J.¹

¹*State Key Laboratory Base of Eco-Hydraulic Engineering in Arid Area, Xi'an University of Technology, Xi'an 710048, China*

²*Hydrology and Water Resources Survey Bureau of Shanxi Province, Taiyuan 030001, China*

**Corresponding author
e-mail: mqfeng@xaut.edu.cn*

(Received 3rd Apr 2019; accepted 17th May 2019)

Abstract. Flood is one of the most frequent natural disasters in the world; therefore, it is the focus of hydrologic research to reveal the effects of such factors as precipitation and peak discharge on the hydrological and water environments. Taking the Jialingjiang Watershed as an example, this paper collects the peak discharge data and historical data in the watershed and then studies the effects of watershed flood peaks on the hydrology and water resources and water environment in the Jialingjiang Watershed using the method combining data statistics and theoretical analysis based on the lognormal distribution model, and predicts the peak discharge in this area using the three-parameter lognormal distribution model. According to the research results, the parameter values are obtained for the Jialingjiang Watershed in the three-parameter lognormal distribution model; the three-parameter lognormal distribution model can reasonably reveal the scale effect law of the confluence area in the watershed peak discharge distribution; and the flood discharge distribution in a small watershed is significantly steeper than that in a large watershed. This study lays a theoretical foundation for the evaluation and prediction of the basic situation of hydrology and water resources in different watersheds.

Keywords: *watershed flood peak, hydrology and water resources, lognormal distribution model, runoff, confluence area*

Introduction

With the rapid economic and technological development, human activities are bringing significant impacts on the global natural environment (Del Giudice et al., 2014; Kuntiyawichai et al., 2014), resulting in changes in the runoff and hydrological cycle paths in the basin and ultimately leading to changes in the characteristics of hydrology and water resources (Al-Rawas and Valeo, 2010; Burton et al., 1997; Cannistraro et al., 2017; Sanjeev, 2017; Zhang et al., 2014). Studies have found that changes in the hydrology and water resources in the watershed can also cause errors in the hydrological analysis data, which will result in differences between the research results and the actual regional hydrological conditions and laws and consequently adversely affect the development and effective utilization of water resources (Ademila and Saloko, 2018; Olang and Fürst, 2011; Guo et al., 2016).

Runoff is formed throughout the process from precipitation to water flowing out of the area interface. Poorly developed runoff is one of the causes to flood (Chinnasamy et al., 2017; Hoseini et al., 2016). As one of the major natural disasters in the world, flood is always a threat to people's lives and properties and land resources (Te et al., 2010). In recent years, due to the over-exploitation of vegetation resources and reckless development of land resources, the ecological environment has been gradually

deteriorating, and floods have occurred frequently, causing serious losses to residents near the waters (Enzel et al., 1993; Drake, 2014; Koutroulis and Tsanis, 2010). Therefore, studying the occurrence rule and formation mechanism of flood is of significant social significance to preventing and predicting flood.

The lognormal distribution model is a data analysis method used in cases where there are random variables (Martinezgoytre et al., 1994; Xin and Venkataramana, 2012; Zhang et al., 2001). In the analysis of the formation and development of watershed flood peaks and their effects on the regional hydrological conditions, there are a series of random variables such as peak discharge and confluence area. For some of them, the distribution can be reasonably analyzed and the relationships between them determined with the lognormal analysis model.

Based on the lognormal distribution model, this paper analyzes the variation law of flood by analyzing the changes in rainfall discharge and flood characteristics in the watershed, and then analyzes the regional variation characteristics of watershed flood peaks and explores the effects of flood peak flow on the hydrology and water resources in a certain area, with a view to providing certain support for the development of water resources in watersheds and the summarization of the development patterns of hydrology and water resources.

Materials and methods

In recent years, in order to characterize the relationship between the confluence area and the peak discharge and their influences in the distribution of the annual maximum peak flow, some scholars have proposed different distribution models for different variable distribution forms (Gupta et al., 1994; Lanfredi et al., 1998). Studies have shown that the lognormal distribution model can organically relate the flood peaks of the watershed to its confluence area. Therefore, this paper uses the lognormal distribution model to analyze the influences of the flood peaks in the Jialingjiang Watershed on the hydrology and water resources.

Lognormal distribution function

The lognormal distribution is a continuous probability distribution function that can be used to characterize the right-skewed distribution. After transformation of the lognormal distribution function, if y obeys the lognormal distribution, then $\ln y$ obeys the normal distribution, and the lognormal distribution probability density function is shown in *Equation 1*, where μ_x is the mean of the parameter x ; and σ_x is the dispersion of the parameter x , and it is always greater than zero. From this, it can be seen that the curve of the logarithmic distribution function is determined by the corresponding mean and the dispersion parameter. According to the graph of the lognormal distribution probability density function as shown in *Figure 1*, when the logarithmic mean $\mu_{\ln y}$ is constant, the peak position of the probability density function graph is determined by the logarithmic dispersion $\sigma_{\ln y}$. The greater the logarithmic dispersion, the wider the convex part of the probability density function graph, and the more blunt the peak; and conversely, the sharper the peak. When the logarithmic dispersion is constant, the peak position of the probability density function graph is determined by the logarithmic mean $\mu_{\ln y}$. The smaller the logarithmic mean is, the more left the peak of the graph will go to,

i.e. the smaller the x-coordinate of the peak will be; and conversely, the more right it will go to and the greater its x-coordinate will be.

$$p(y) = \frac{1}{\sigma_{\ln y} y \sqrt{2\pi}} e^{-\frac{(\ln y - \mu_{\ln y})^2}{2\sigma_{\ln y}^2}}, \quad (0 < y < +\infty) \quad (\text{Eq.1})$$

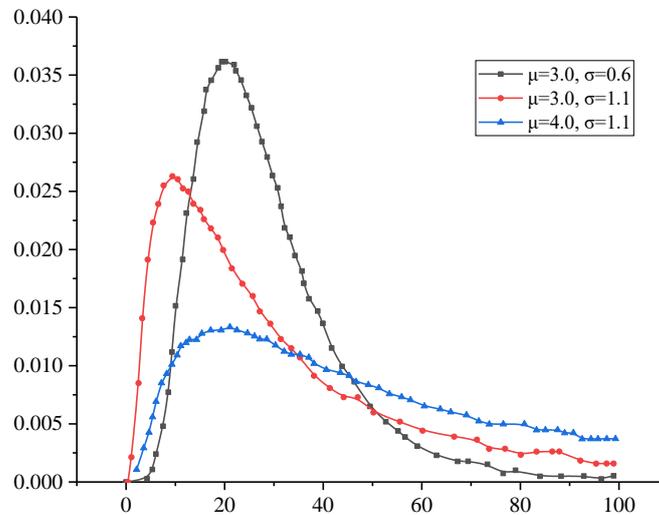


Figure 1. Lognormal distribution

The distribution characteristics of the above lognormal distribution function can be determined by two parameters, logarithmic mean and logarithmic dispersion. However, studies have shown that the shape of the lognormal distribution is generally asymmetrical, so the degree of skew the curve distribution is also essential. In the probabilistic statistical analysis, the coefficient of skewness C_s of the data distribution is defined in Equation 2, where μ^3 is the third-order center distance of y , and σ is the mean square error of y . When the coefficient of skewness is 0, the lognormal distribution is symmetric; when it is less than 0, the lognormal distribution is left-skewed; and when it is greater than 0, the lognormal distribution is right-skewed.

$$C_s = \frac{\mu^3}{\sigma^3} \quad (\text{Eq.2})$$

Two-parameter lognormal distribution model

The two-parameter lognormal distribution model is a distribution model proposed by Prof. Smith, in which a random variable y is set to satisfy the two-parameter lognormal distribution $LN(y: \mu, \sigma^2)$, where μ is the mean of the random variable $\ln y$; and σ^2 is the variance of $\ln y$. If x is a random variable obeying the standard normal distribution, the random variable y can be expressed by Equation 3. Both μ and σ^2 have a linear relationship with the confluence area S in the watershed analysis, then they can be expressed by Equations 4 and 5, respectively.

$$y = \exp(\mu + \sigma x) \quad (\text{Eq.3})$$

$$\mu = n_1 + n_2 \ln S \quad (\text{Eq.4})$$

$$\sigma^2 = m_1 - m_2 \ln S \quad (\text{Eq.5})$$

Combine the above equations, and we obtain *Equation 6* can be obtained. If the coefficient m_1 is much greater than m_2 , then *Equation 6* can be approximated to *Equation 7*, $f(z) = \exp(n_1 + m_1 x)$ and $\theta(z) = n_2 - 0.5m_2 x$. Therefore, when μ and σ^2 are in linear relationships with the confluence area of the watershed, the relationship between the flow Q of the watershed flood peak and the confluence area S is a power function.

$$y = \exp(n_1 + n_2 \ln S + \sqrt{(m_1 - m_2 \ln S)} x) \quad (\text{Eq.6})$$

$$Q = \exp[n_1 + n_2 \ln S + (m_1 - 0.5m_2 \ln S)x] = f(z)S^{\theta z} \quad (\text{Eq.7})$$

Three-parameter lognormal distribution model

Experts and scholars have improved the two-parameter lognormal distribution model and proposed a three-parameter lognormal distribution model to more accurately analyze the inherent law of hydrology and water resources. In this model, it is considered that the random variable y obeys the three-parameter lognormal distribution, that is, $\text{LN}(y: \mu, \sigma^2, a)$. The random variable x obeys the standard normal distribution, so the random variable y is expressed in *Equation 8*, and the mean of y is shown in *Equation 9*. The mean μ and variance σ^2 of the random variable $\ln y$ are still expressed by *Equations 4* and *5*, while the parameter a is shown in *Equation 10*. The meaning of each parameter is the same as above. Based on the above descriptions, the three-parameter lognormal distribution model can be determined by six parameters - $s_1, s_2, n_1, n_2, m_1, m_2$, and then the parameters related to the flood peak discharge distribution in each confluence area S in the watershed can be obtained to determine the distribution pattern of watershed flood peaks. At last, the regional analysis of the flood and the estimation of the peak discharge in the watershed are conducted.

$$y = \exp(\mu + \sigma x) + a \quad (\text{Eq.8})$$

$$E(y) = \exp(\mu + 0.5\sigma^2) + b \quad (\text{Eq.9})$$

$$b = E(y) - \exp(\mu + 0.5\sigma^2) = e^{s_1} A^{s_2} - e^{n_1 + m_1/2} A^{n_2 - m_2/2} \quad (\text{Eq.10})$$

Flood peak data of Jialingjiang Watershed

This paper selects the annual flood peak data acquired at 31 hydrological stations in Jialingjiang Watershed (as shown in *Fig. 2*) for study. The confluence area of each station is controlled within 50–150,000 km². The maximum peak discharge data of each station in different years is shown in *Table 1*.

Results

Analysis of annual watershed flood peaks

According to the data in the *Table 1*, a double logarithmic relationship graph between the controlled confluence area S at each station and the mean of the maximum

peak values Q is drawn. $f=0.01, 0.1$ and 0.4 are shown in *Figures 3, 4* and *5*, respectively. It can be seen that the relationship between $\log Q(S)$ and $\log S$ is approximately linear, indicating that the peak discharge Q is scale invariant with the confluence area. At different frequencies f , the slope of the line is different, which represents the scale index. In other words, as the confluence area S changes, the mean Q of the maximum peak values satisfies the multi-scale form.

Calculate the 1st-7th moment $M = E[Q^h(S)]$ of each station, and plot the double logarithmic relationships between the watershed area and the 2nd order matrix and the 4th order matrix, as shown in *Figures 6* and *7*. Similarly, the relationship between $\log M(S)$ and $\log S$ (see *Fig. 8*) is approximately linear, indicating that the peak discharge Q is scale invariant with the watershed area changing. Calculate the scale value θ at different frequencies f through linear regression analysis, as shown in *Table 2*. It can be seen that the scale index increases with the increase of frequency, indicating that the peak discharge of a large flood in the low return period grows more slowly than that of a small flood in the high return period. In other words, the peak discharge values of a small watershed are more steeply distributed, which is in line with the actual conditions.

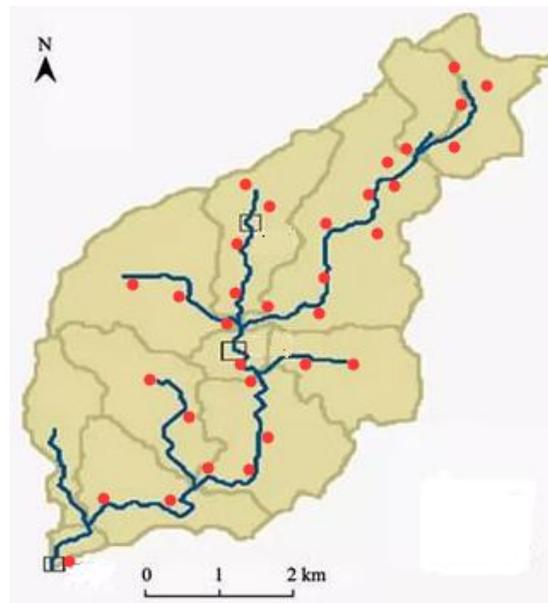


Figure 2. Map of the Jialingjiang Watershed

Table 1. Flood data at different stations in the watershed

No.	Confluence area (km ²)	Mean value of flood peak discharge (m ³ /s)	Years (up to 2018)
1	254	235	21
2	5187	5068	30
3	91	86	25
4	90546	90428	24
5	8874	9857	18
6	632	649	38
7	14812	15065	29
...
31	6857	6843	42

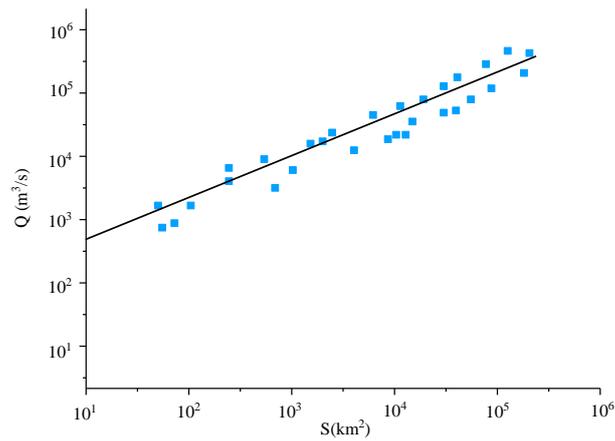


Figure 3. Double logarithmic relationship between peak discharge and confluence area ($f = 0.02$)

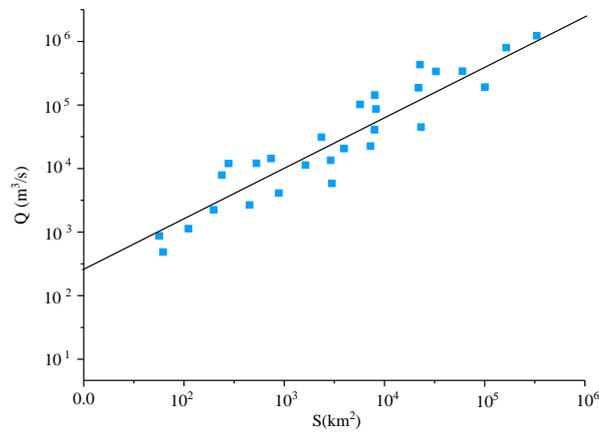


Figure 4. Double logarithmic relationship between peak discharge and confluence area ($f = 0.1$)

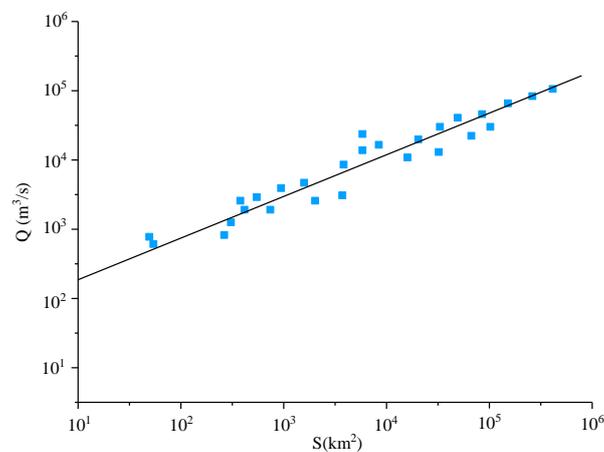


Figure 5. Double logarithmic relationship between peak discharge and confluence area ($f = 0.31$)

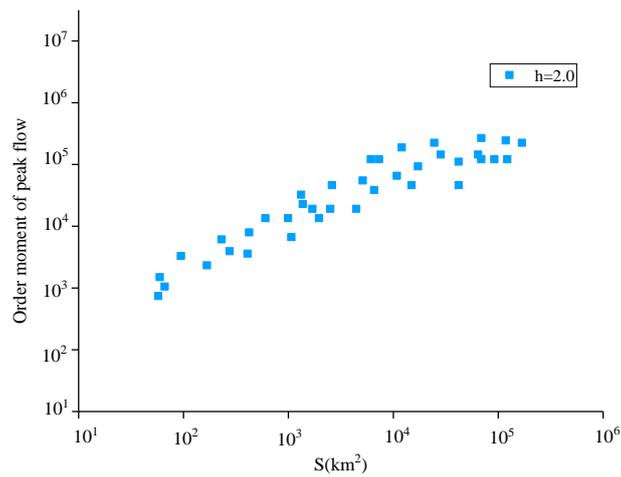


Figure 6. Relationship between second-order moment and watershed area

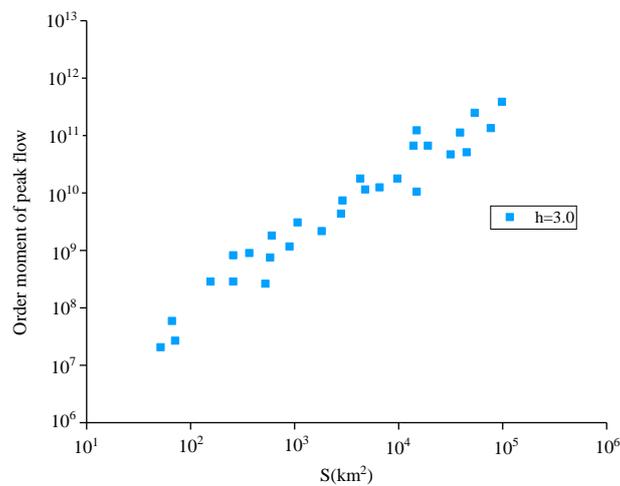


Figure 7. Relationship between third-order moment and watershed area

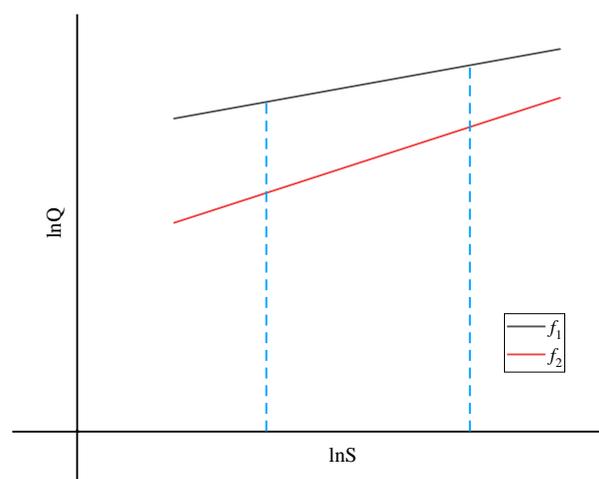


Figure 8. Relationship between peak discharge and confluence area

Table 2. Scaling index of peak discharge at different frequencies

P	0.1%	0.2%	0.5%	1%	2%	10%	15%	20%	31%
$\theta(f)$	0.538	0.542	0.547	0.552	0.578	0.609	0.657	0.673	0.701

Analysis on the application of the lognormal distribution model

According to the scale invariance feature, the peak discharge of the area where the no existing data is available can be calculated based on the annual peak discharge of the area where it is available. In this way, the probability distribution of the former can also be obtained.

Take the hydrological station in the Jialingjiang Watershed as an example. The flow control area of the hydrological station is 82,391 km². According to the 35-year maximum flow probability distribution map, the distribution of the flow data matches *f*-3. Suppose another hydrological station, with a control area of 725 km², has no data available. Then, according to the scale index θ found in *Table 2*, the peak discharge probability value of the area with no data can be obtained based on the existing data of the area with data. *Table 3* shows the calculated peak discharge values and the true ones, which are almost consistent.

Table 3. Calculations for the application example

P(%)	Q(Aref)(m ² /s)	Q(A)(m ² /s)	
		True value	Calculated value
0.1%	52264	4562	4412
0.2%	42142	3244	3135
0.5%	38263	2806	2709
1%	36021	2623	2645
2%	35682	2453	2346
10%	30347	2074	2012
15%	27803	1678	1789
20%	23649	1241	1231
31%	16323	954	924

Discussion

According to the three-parameter lognormal distribution model, the six parameters (n_1 , n_2 , m_1 , m_2 , s_1 and s_2) are used to determine the three parameters μ , σ^2 and b of the lognormal distribution, and thus the distribution of the watershed flood peaks with different confluence area S in the Jialing River Basin. Finally, the design peak discharge in the area with no data can be estimated based on the existing hydrologic data of the area with data.

According to the model, relevant data from 31 hydrological stations in the Jialingjiang Watershed are selected, and the six parameters are all obtained: $s_1 = 2.341$, $s_2 = 0.573$; $n_1 = 2.453$, $n_2 = 0.756$; and $m_1 = 0.273$, $m_2 = 0.026$. Based on this, the mean μ , the dispersion σ^2 and the parameter b in the three-parameter lognormal distribution model are determined. Then the distribution of the flood peaks in the Jialingjiang Watershed is obtained according to the analysis in Section 3 above, and the effects of

flood peaks on the hydrology and water resources in the watershed are determined, together with preliminary evaluation on the water environment. The results also show that the lognormal distribution model selected in this paper can reasonably predict the impact of the flood peaks in the Jialingjiang Watershed on the conditions of hydrology and water resources in this region.

Conclusion

Taking the Jialingjiang Watershed as an example, this paper collects the peak discharge data and historical data in the watershed and then studies the effects of watershed flood peaks on the hydrology and water resources and water environment in the Jialingjiang Watershed using three-parameter lognormal distribution model. The main conclusions are as follows:

(1) The three-parameter lognormal distribution model is more applicable to the analysis and evaluation of hydrologic conditions in river watersheds than the two-parameter one. The former can reasonably characterize the scale effect law of the confluence area in the watershed peak discharge distribution.

(2) The peak discharge in the Jialingjiang Watershed changes with the confluence area in a scale-invariant way; the values of the 6 parameters in the lognormal distribution model can be determined according to the confluence area S , and then based on this, the distribution of watershed flood peaks can be determined.

(3) The study on the effects of peak discharge in Jialingjiang Watershed on the hydrological and water environments in this area further deepens our understanding of the development patterns of and influencing factors to the hydrological conditions in this area and provides technical materials and theoretical support for the rational development of hydrology and water resources in this area in the next steps.

Acknowledgements. The authors would like to thank all of the researchers who participated in the project. The authors would also like to thank the Research and Extension of Water Conservancy Science and Technology in Shanxi Province (201625) for financial support. Furthermore, the authors would like to thank the anonymous reviewers for their helpful and constructive comments.

REFERENCES

- [1] Ademila, O., Saloko, B. (2018): Hydrogeoelectrical evaluation of groundwater flow pattern in a basement complex terrain, Southwest Nigeria. – *Environmental and Earth Sciences Research Journal* 5(1): 7-14.
- [2] Al-Rawas, G. A., Valeo, C. (2010): Relationship between wadi drainage characteristics and peak-flood flows in arid northern Oman. – *Hydrological Sciences Journal* 55(3): 377-393.
- [3] Burton, T. A. (1997): Effects of basin-scale timber harvest on water yield and peak streamflow1. – *JAWRA Journal of the American Water Resources Association* 33(6): 10.
- [4] Cannistraro, G., Cannistraro, M., Trovato, G. (2017): Islands “Smart Energy” for eco-sustainable energy a case study “Favignana Island”. – *International Journal of Heat and Technology* 35(S1): S87-S95.
- [5] Chinnasamy, P., Muthuwatta, L., Eriyagama, N., Pavelic, P., Lagudu, S. (2017): Modeling the potential for floodwater recharge to offset groundwater depletion: a case study from the Ramganga basin, India. – *Sustainable Water Resources Management* 4(2): 331-344.

- [6] Del Giudice, G., Rasulo, G., Siciliano, D., Padulano, R. (2014): Combined effects of parallel and series detention basins for flood peak reduction. – *Water Resources Management* 28(10): 3193-3205.
- [7] Drake, C. W. (2014): Assessment of flood mitigation strategies for reducing peak discharges in the Upper Cedar River watershed. – Master of Science Thesis, University of Iowa. DOI: 10.17077/etd.yu3t0xjm.
- [8] Enzel, Y., Ely, L. L., House, P. K., Baker, V. R., Webb, R. H. (1993): Paleoflood evidence for a natural upper bound to flood magnitudes in the Colorado river basin. – *Water Resources Research* 29(7): 2287-2297.
- [9] Guo, Y., Chang Huang, C., Pang, J., Zhou, Y., Zha, X., Mao, P. (2016): Reconstruction palaeoflood hydrology using slackwater flow depth method in the Yanhe River valley, middle Yellow River basin, China. – *Journal of Hydrology* 2016: 156-171. DOI: 10.1016/j.jhydrol.2016.11.017.
- [10] Gupta, V. K., Mesa, O. J., Dawdy, D. R. (1994). Multiscaling theory of flood peaks: regional quantile analysis. – *Water Resource* 30(12): 3405-3421.
- [11] Hoseini, Y., Azari, A., Pilpayeh, A. (2016): Flood modeling using WMS model for determining peak flood discharge in southwest Iran case study: simili basin in Khuzestan province. – *Applied Water Science* 7(6): 3355-3363.
- [12] Koutroulis, A. G., Tsanis, I. K. (2010): A method for estimating flash flood peak discharge in a poorly gauged basin: case study for the 13-14 January 1994 flood, Giofiros basin, Crete, Greece. – *Journal of Hydrology (Amsterdam)* 385(1-4): 150-164.
- [13] Kuntiyawichai, K., Sri-Amporn, W., Pruthong, C. (2014): Quantifying consequences of land use and rainfall changes on maximum flood peak in the lower Nam Phong river basin. – *Advanced Materials Research* 931-932: 791-796.
- [14] Lanfredi, M., Macchiato, M., Ragosta, M., Serio, C. (1998). Time correlation structure in hourly concentration time series of CO, NO_x, and O₃ in urban areas. – *Fractals* 6(2): 151-158.
- [15] Martinezgoytre, J., House, P. K., Baker, V. R. (1994): Spatial variability of small-basin paleoflood magnitudes for a southeastern Arizona mountain range. – *Water Resources Research* 30(5): 1491-1501.
- [16] Olang, L. O., Fürst, J. (2011): Effects of land cover change on flood peak discharges and runoff volumes: model estimates for the Nyando river basin, Kenya. – *Hydrological Processes* 25(1): 80-89.
- [17] Sanjeev, R. (2017): Geophysical resistivity survey (VES) for selection of appropriate artificial recharge (Ar) structures for augmentation of groundwater resources in Gwalior, MP, India. – *Environmental and Earth Sciences Research Journal* 4(1): 7-11.
- [18] Te, L. A. H., Aerts, J. C. J. H., Bakker, A. M. R., Kwadijk, J. C. J. (2010): Simulating low-probability peak discharges for the Rhine basin using resampled climate modeling data. – *Water Resources Research* 46(46): 135-146.
- [19] Xin, J., Venkataramana, S. (2012) Impacts of climate change on hydrology and water resources in the Boise and Spokane River Basins. – *JAWRA Journal of the American Water Resources Association* 48(2): 197-220.
- [20] Zhang, Q., Gu, X., Singh, V. P., Xiao, M., Xu, C. Y. (2014): Stationarity of annual flood peaks during 1951-2010 in the pearl river basin, China. – *Journal of Hydrology* 519: 3263-3274.
- [21] Zhang, Y., Smith, J. A., Baeck, M. L. (2001): The hydrology and hydrometeorology of extreme floods in the great plains of eastern nebraska. – *Advances in Water Resources* 24(9): 1037-1049.