Calculation of Hydrological Random Variables and Probability Distribution Based on Copula Function

Song Songbai [1] ,Wang Xiaojun[2]

1. School of Water Conservancy and Architectural Engineering, Northwest Agriculture and Forestry University, Key Laboratory of Agricultural Water and Soil Engineering in Arid Areas, Ministry of Education, Yangling 712100, Shaanxi Province;
2. State Key Laboratory of Hydrology and Water Resources and Hydraulic Engineering Science of nanjing hydraulic research institute, Nanjing 210029, Jiangsu)

Abstract: Hydrological random variables and distribution calculation are important contents for deducing design flood area composition, design flood downstream of cascade reservoirs, etc. They are of great importance for planning, design and management of water conservancy and hydropower projects downstream of reservoirs, and urban flood control risk assessment, etc. The traditional hydrological random variables and distribution are derived from the function distribution of two-dimensional variables. The marginal distribution must be of the same type, and its application is limited. According to the definition of two-dimensional random variables and probability distribution, this paper uses the Copula function and product variation to change the original principle, and strictly deduces the calculation formula of two-dimensional phase distribution according to machine variation and probability. The distribution probability calculation formula of the sum of variables under the two commonly used marginal distributions of Gamma distribution and P-Ⅲ distribution is only one-dimensional integration of conditional Copula function, thus avoiding information distortion of data conversion by probability combination discrete summation method and overcoming the requirement of traditional multivariate distribution that the marginal distribution is of the same type. Taking the 3h flood volume composition from Shuibuya Reservoir to Geheyan Reservoir in Qingjiang River Basin as an example, the calculation method of hydrological random variables and distribution is given. The model and calculation method in this paper are expected to provide theoretical support for the composition of design flood areas and the calculation of design flood downstream of cascade reservoirs in China.
Keywords: dependency; Sum of random variables; Probability distribution; Copula function; Qingjiang River Basin

Research background

The calculation of two-dimensional dependent hydrological random variables and probability distribution is an important content and core calculation technology of basin design flood area composition, cascade reservoir downstream design flood, etc. Therefore, how to improve the calculation accuracy of hydrological variables and distributions is highly concerned by many scholars. The Handbook of Design Flood Calculation for Water Conservancy and Hydropower Projects recommends the use of area composition method, frequency combination method and random simulation method for design flood

The calculation of the regional composition of Among them, the frequency combination method is recommended to use the probability combination discrete summation method proposed by Wang Ruichen and others in 1990 .

Yuan Xi was one of the first scholars to carry out the calculation of two-dimensional hydrological random variables and probability distribution in China, and deduced that the shape parameter is a positive integer.

The Gamma hydrological variable sum and difference distribution analytical calculation formula of. Huang Nong

On this basis, Zhang Yuanxi's research results were expanded and put forward.

A numerical method for calculating the sum of two independent Gamma distribution variables is presented. Since the

The discrete summation method is used to calculate the design flood area composition of cascade reservoirs.

Yan Baowei and others

Systematically summarizes the existing calculation methods

Inadequacies: The discrete summation method converts two-dimensional integration into a combination summation of two variables with a finite number of "state" frequencies, which inevitably leads to information distortion in the data conversion process; The stochastic simulation method carries out on-board simulation of upstream and downstream sections and interval flood processes by establishing a spatial multi-station stochastic mathematical model, and the retention of statistical characteristics of simulation sequences is still difficult to master. Although the calculation method of regional composition method is simple and convenient, this method will reduce the interval flood

[7]

According to a certain composition is relatively fixed, the uncertainty of human factors is greater. Li Tianyuan and others

Applying Copula Function to Obtain Explicit Table of Conditional Probability

At last, the conditional probability curve is discretized to solve the distribution probability of the sum of two-dimensional dependent hydrological random variables. This method improves the calculation accuracy of the probability combination discrete summation method, and does not need to do independent processing on variables. However, it still belongs to the probability combination discrete summation method.

The marginal distribution derived by Nadarajah must be of the same type

According to the determination of two-dimensional random variables and distribution functions

According to the definition of two-dimensional random variable and probability distribution, the calculation of two-dimensional random variable and probability distribution is converted into one-dimensional definite integral of conditional Copula by using Copula function and mathematical integral transformation principle, so as to avoid information distortion in the data conversion process of discrete summation method. On this basis, the formula for calculating the distribution probability of the sum of two commonly used marginal distribution variables, Gamma distribution and P-Ⅲ distribution, is derived. Taking the 3 h flood volume composition from Shuibuya Reservoir to Geheyan Reservoir in Qingjiang River Basin as an example, the application of the model in the exposition is explained. The model in this paper is expected to provide theoretical support for the composition of flood planning areas and the design flood calculation of the downstream of cascade reservoirs in China.

Copula Function Expressions for Two-Dimensional Random Variables and Distributions

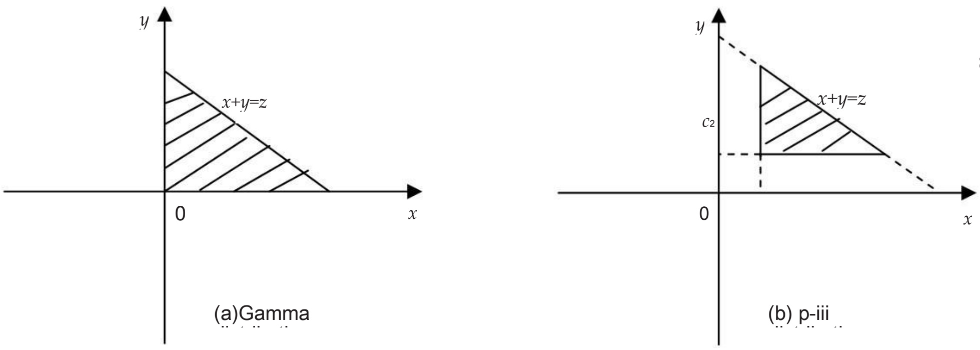
Where: F_X *x "and* x *are density functions of random variable* x *; The upper and lower limits of the integrals of random variables* x *and* y" can be determined respectively according to the value ranges of their distribution function variables.

Equation (4) is the calculation expression of the Copula function for two-dimensional random variables and probability distribution. obviously, the integral function is only the function of the variable *u*

Number. The two-dimensional Copula function is C U， V According to the dependent characteristics of the two-dimensional random variable X， Y , the function type can be selected and the corresponding parameters can be determined according to the multivariate distribution calculation method usually based on Copula function. " The upper and lower limits of integral variable and U can be determined by U = F, X and X according to the value range of *random variable* and x

Expressions of Variable and Distributed Copula Functions under Marginal Distribution Gamma and P-Ⅲ Distribution

Gamma and P-III distribution are commonly used distribution functions in water texts. The root data formula (4) is pushed to guide the distribution of variables and distribution Copula function expressions from Gamma and P-III



spectively

Figure 1 Integral Region

Given Gamma and P-Ⅲ marginal distributions, the variables and distributions $Z = X+Y$ and the integration area for probability calculation are shown in Figure 1. Gamma

distribution variables, x, and, y, respectively, $x > 0$， $y >$, 0, p-iii distribution variables, $x > c$， $y > c$， $z = x+y$ are finite and known according to "

The integration region and the two-dimensional definite integral principle shown in figure 1 , the x integral is converted into the variable *u* integral, the upper limit value of which is the *x* marginal distribution value of *z*, and the value 1 is no longer taken.

Two-dimensional Random Variables and Probability Distribution Model Solution

From equation (9), it can be seen that *z* is already known, and the two-dimensional distribution f_z and *z* with machine variation and probability are only one-dimensional fixed product of variation *u*. 当

The expression of expressiveness is generally more complicated.

In actual calculation, the continued fraction calculation generally takes I = 7-10 "to meet the accuracy requirements.

Application examples

In this paper, the 3 h flood volume sequence distribution parameters and Copula function parameters of Qingjiang cascade reservoirs adopted in reference [7] are selected to illustrate the application of two-dimensional dependent hydrological random variables and probability distribution calculation model. *X* represents the 3-h flood volume of Shuibuya Reservoir, *Y* represents the 3-h flood volume from Shuibuya Reservoir to Geheyan Reservoir (water-separation interval). The water volume of Shuibuya Reservoir plus the water-separation interval equals the water volume of Geheyan Reservoir downstream. Statistical parameters and distribution parameters of 3 sequences are shown in table 1 .

Substituting Equation (14) into Equation (9), random variables *X* and *Y* respectively select Gamma and P-III distributions, and apply the above continued fraction numerical calculation method.

Method, under two marginal distributions, $X+Y \leq Z$ and the probability calculation results are shown in columns (2) and (5) of Table 2 respectively, and the comparison is shown in Figure 2. " In order to verify the correctness of the calculation results of the model in this paper, columns (3) and (6) of Table 2 respectively list the theoretical probability values calculated by Gamma and P-III respectively for the X+Y *sequence, and columns (4) and (7) give the corresponding error calculation values.* Its error is due to x 、 y *and* different distribution parameters. Formula (4) has no explicit analytical calculation formula and can only be calculated numerically by the variable step ladder method in this paper, thus it has certain deviation. In addition "

It is also affected by the fitting precision of marginal distribution. It is not difficult to see that under the two marginal distributions, the calculation results of the sum probability distribution of the model and variables and sequences in the paper are basically consistent, which indicates that the model and calculation method in the paper are correct.

Conclusion

The calculation of two-dimensional dependent hydrological random variables and probability distribution is the important content of design flood calculation and design flood area composition at the downstream of cascade reservoirs. Based on the Copula function and the principle of mathematical integral transformation, this paper strictly deduces the probability calculation formula of the two Gamma distributions and the P-Ⅲ distributions depending on the hydrological random variables and distributions. Taking the 3 h flood volume composition from Shuibuya Reservoir to Geheyan Reservoir in Qingjiang River Basin as an example, it is said that

Application of model in plaintext. The main research conclusions are as follows: (1) The calculation formula in this paper overcomes the possible errors in data conversion by discrete summation method. (2) The calculation formula is expressed as the definite integral of one-dimensional conditional Copula function. Using computer numerical integration, it is not difficult to obtain satisfactory solution results. For P-Ⅲ distribution, it can be converted into Gamma distribution function numerical calculation by calculating sequence variable value minus distribution position parameter. (3) In this paper, the two-dimensional dependent hydrological random variables and the probability distribution are calculated by using conditional Copula function, which can conveniently combine the hydrological characteristic variables with different distribution dependencies, and can overcome the limitation that the traditional multivariate distribution function requires the same kind of marginal distribution function. (4) Under the two Gamma distributions and P-III distributions in the study area, the flood volume combination and probability distribution results show that the calculation formulas of random variables and probability distribution are correct. The model in this paper is expected to provide theoretical support for the design flood area composition of watersheds and the design flood calculation of downstream cascade reservoirs in China.

(5) Based on the fact that the downstream design section flood volume is equal to the sum of the upstream reservoir section flood volume and the section flood volume, without considering the reservoir regulation effect, the calculation formula of the downstream design section flood volume probability distribution is derived. Therefore, how to consider the downstream section design flood under the influence of reservoir flood regulation needs further in-depth study.

References

1. Ministry of Water Resources, Yangtze River Water Conservancy Commission, Hydrographic Bureau, Ministry of Water Resources, Nanjing Institute of Hydrology and Water Resources. Design Flood Calculation Manual for Water Conservancy and Hydropower Projects [M]. Beijing: China Water Conservancy and Hydropower Publishing House, "2001.
2. Zhang Yuan Jubilee. Independent Quantity GAMMA Joint Distribution Function of Sum and Difference of Distribution Random Variables [J]. Journal of Hefei University of Technology, "1983 (3) : 53-59 .
3. Huang Nong. A Combined Frequency Numerical Method for Calculating the Sum of Two Independent γ Distribution Random Variables [J]. Journal of Hefei University of Technology, "1987 , 9 (5) : 110-115 .
4. Xie Xiaoping, Huang Lingzhi, Xi Qiuyi, et al. Research on the Formation of Flood Control Zones Based on JC Method [J]. Journal of Hydrodynamics, 2006 , 25 (6) : 125-129 .
5. Huang Lingzhi, Xie Xiaoping, Huang Qiang, et al. JC Method [J]] Journal of Natural Disasters, 2006 , 15 (4) : "in Research on Design Flood Area Composition of Cascade Reservoirs 163-167 .
6. Yan Baowei, Lei Shizhong, Hu Si-yi, et al. Study on Design Flood Area Composition Based on Copula Function [J]. Journal of Hydroelectric Power Generation, 2010 , 29 (6) : 60-65 .
7. Li Tianyuan, Guo Shenglian, Liu Zhangjun, et al. Research on Method of Calculating Flood Water under Cascade Water Reservoirs [J]. Journal of Hydrology, 2014 , 45 (6) : 641-648 .
8. Liu Zhangjun, Guo Shenglian, Li Tianyuan, et al. Research on the Establishment of Flood Control Zones for Cascade Water Reservoirs Based on the Copula Function [J]. Research on Water Resources.2014 , 3 (2) : 124-135 .
9. Liu Zhangjun, Guo Shenglian, Li Tianyuan, et al. Study on the Method of Inter-district Estimation Based on Setting up Flood Control Zones [J]. Journal of Hydrology, 2015 , 46 (5) : 543-549 .
10. Liu Zhangjun, Guo Shenglian, Li Tianyuan, et al. General formula for calculating the most probable area composition of design flood of cascade reservoirs [J]. advances in water science, 2014 , 25 (4) : 549-558 .
11. Li Fei, Guo Shenglian, Li Tianyuan, et al. Study on Composition Method of Discontinuous Series Design Flood Areas [J]. Hydroelectric Energy Science, 2011 , 29 (5) : 47-49 .
12. NADARAJAH S . Products and ratios for a bivariate gamma distribution [J]. Journal of Applied Mathematics and Computing , 2005 , 24 : 581-595 .
13. NADARAJAH S . A bivariate distribution with gamma and beta marginals with application to drought data [J].Journal of Applied Statistics , 2009 , 36 : 277-301 .
14. NADARAJAH S . A bivariate gamma model for drough [t J]. Water Resources Research , 2007 , 43 : 749-759 .
15. NADARAJAH S . A bivariate Pareto model for drough [t ment , 2008 , 23 : 811-822 .J]. Stochastic Environmental Research and Risk Assess-
16. NADARAJAH S , GUPTA A K . Cherian's bivariate gamma distributioins as a model for drought data [J]. Agro- ciencia-Mexico , 2006 , 40 : 483-490 .
17. NADARAJAH S , GUPTA A K . Friday and Patil's bivariate exponential distributioin with application to drought data [J]. Water Resources Management , 2006 , 20 : 749-759 .
18. NADARAJAH S , GUPTA A K . Intensity-duration models based on bivariate gama distributions [J]. Hiroshima Mathematical Journal , 2006 , 36 : 387-395 .

19. NADARAJAH S , KOTZ S . Sums , products and ratios for Downton's bivariate exponential distribution [J] . Sto- chastic Environmental Research and Risk Assessment , 2006 , 20 : 164-170 .
20. Huang Zhenping, Chen Yuanfang. Hydrostatistics [M] . Beijing: China Water Resources and Hydropower Press, 2011.
21. Song Songbai, Cai Huanjie, Jin Juliang, et al. Copulas Function Theory and Its Application in Hydrology [M] . Beijing: Science Press, 2012.
22. NELSEN R B . An Introduction to Copulas [M] . Springer , New York , 1999 .
23. Xu Shiliang. FORTRAN Common Algorithm Assembly [M] .2 Edition. Beijing: Tsinghua University Publishing House, 1997.
24. Zou Hai. New Calculation Method in Optimal Design [M] . Beijing: New Times Publishing House, 1982.