ISSN 2616-5805 Vol. 5, Issue 3: 52-59, DOI: 10.25236/AJMS.2024.050308

Analysis of Drought Characteristics Based on 3-dimensional Copula Function in Yunnan Province

Shen Yurun, Huang Haiyue, Wang Junhao, Zuo Dongdong*

School of Mathematics and Physics, Yancheng Institute of Technology, Yancheng, 224051, China
**Corresponding author: dongdz_vow@163.com*

Abstract: *Based on the precipitation data from 128 stations in Yunnan Province and the drought-affected area data in the province from 1961 to 2021, an analytical model of drought duration, drought intensity and drought-affected area rate was constructed using the principle of maximum entropy and the 3-dimensional Copula function. The results show that the distribution functions of drought duration, drought intensity and drought area rate can be well derived based on the maximum entropy principle, and pass the K-S distribution test. The calculation results of drought characteristics in Yunnan Province show that the mean values of drought duration, drought intensity and drought area rate are 5.7 months, 3.1 and 0.12, respectively. Among the three Archimedean Copula, the calculation results of Gumbel are relatively better, which is more suitable to be used in the construction of distribution functions. The level of drought duration and drought intensity of 3 to 4 are the main reasons for the drought area rate greater than 0.18.*

Keywords: *3-dimensional Copula; Maximum entropy principle; drought duration; drought severity; drought area ratio*

1. Introduction

Drought is a result of prolonged water scarcity, and the process of drought causation is relatively slow compared to disasters caused by other meteorological hazards, e.g., heavy rainfall, floods, and typhoons. Due to the complexity of drought [1-4], the description of drought as well as the quantification of drought disaster impacts are difficult. China is a drought-prone country, and the affected area due to drought accounts for nearly 50% of the affected area due to all meteorological disasters every year [5]. Therefore, strengthening drought-related research can help us better understand the causes and processes of drought, the impacts of drought on the ecosystem and agricultural production, and ways to cope with drought.

The analysis of drought cannot be separated from the analysis of water supply characteristics, and precipitation and runoff are the main sources of water for agriculture, because these quantities are random variables, so the corresponding drought characteristics are also stochastic, and should be described by using probabilistic models, and carrying out the statistical analysis of drought characteristics is a typical example. For stochastic problems, the first problem to be solved is to find the distribution function of the random variable. Drought duration, drought intensity, drought range are common random variables in drought characterisation [6], and these variables are often not independent, and practical calculations also show that there is a strong correlation between the variables. For non-independent random variables with inconsistent distribution function types, it is difficult to solve the joint distribution function of the variables. Copula function is an effective solution, and more scholars have used Copula function to carry out multivariate statistical characterisation, e.g., Shiau [7] used a bivariate Copula function to establish a joint distribution function of drought duration and drought intensity to analyse the drought frequency characteristics and return period characteristics in southern Taiwan. Ping Liu et al [8] applied a two-dimensional Copula function to carry out the probabilistic characterisation of extreme precipitation.Vergni L et al [9] discussed agricultural drought in central Italy using a bivariate Copula function. Li Qian et al [10] applied Copula function for drought disaster risk assessment in Beijing-Tianjin-Hebei region. Mirabbasi et al [11] applied two-dimensional Copula function for meteorological drought analysis in northwestern Iran.

At present, the application of Copula function is mostly focused on the analysis of two-variable problems, and there are relatively few three-dimensional Copula application analyses. Three-dimensional Copula related to drought analysis are: Hao et al [12] used Copula to construct a drought frequency analysis model of drought duration, drought intensity and drought kurtosis, and analysed the spatial characteristics of drought risk in southwest China. Wang Ying et al [13] used 3D Copula function to establish the joint distribution function of multi-tributary runoff, and simulated and predicted the annual maximum flow of the mainstem using the joint Monte Carlo method. The small number of 3D Copula studies is more related to the lack of sample size, especially when it comes to the analysis of disaster characteristics, the lack of sample size is especially obvious. The article of Hou et al [14] discusses the application of the 3D Copula method in the analysis of disaster early warning on the county scale, and obtains a better application effect, but there is still the problem of the small sample size in this analysis. Considering the relative completeness of data on a provincial scale, this paper attempts to carry out drought disaster characterisation based on 3D Copula function on a provincial scale based on the study of Hou et al.

2. Data

The data used in this paper consist of two parts: 1) month-by-month precipitation data from 128 stations in Yunnan Province from 1961-2021 provided by the China Meteorological Administration; and 2) year-by-year data on the total sown area of crops and the area affected by drought in Yunnan Province from 1961-2021 downloaded from the National Bureau of Statistics. Figure 1 shows the average annual precipitation and the spatial distribution of the 128 stations in Yunnan Province.

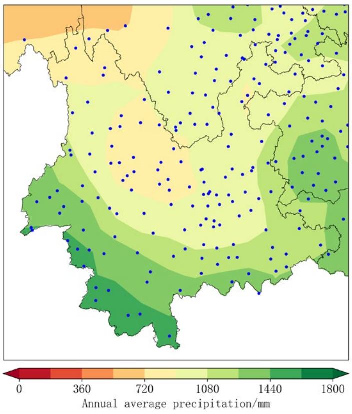


Figure 1: The spatial distribution of the annual average precipitation and meteorological site

As can be seen from Fig. 1, the precipitation in the southern region of Yunnan Province is greater than that in the northern region, with annual precipitation of up to 1600 mm in some areas and less than 1000 mm in some areas in the north.

3. Methodology

3.1 Characterisation of exposure rates and drought characteristics

3.1.1 Rate of drought-affected areas

The crop rate of drought-affected areas is a disaster index, which is used to express the magnitude of drought impacts. It is calculated with reference to the National Standard of the People's Republic of China: Drought Hazard Levels, and is expressed mathematically as follows.

$$I = \frac{A_d}{A_p} \times 100\% \tag{1}$$

Where A_d denotes the area of drought-affected crops and A_p denotes the total sown area of crops.

3.1.2 Drought characteristics

In this study, SPI developed by McKee [15] and widely used will be used as an indicator for drought description. SPI has different time scales such as 1 month, 3 months, 6 months, 12 months, etc., and the indicators of different scales are often used in different drought problem analyses, and the SPI with 3 months time scale is often used in the analysis of agricultural drought. The drought analysis in this paper belongs to the agricultural drought problem, so the SPI indicator with a time scale of 3 months is selected. In addition, in order to be consistent with the spatial scale of the drought area rate, the SPI indicator of each station is calculated first, and then the provincial average SPI is calculated from the stations included in each province. Using SPI equal to 0 as the threshold for drought event identification, the number of months in each year with SPI below 0 was identified and the value was used as the drought duration. The cumulative value of SPI below 0 was used as the drought intensity magnitude for the year. The calculation formula is as follows.

$$S = -\sum_i SPI_i \tag{2}$$

Thus, a drought duration and drought intensity can be obtained for each year.

3.2 Derive distribution function based on maximum entropy principle

Usually, when the distribution type of a random variable is unknown, it is often necessary to compare multiple distribution functions and select the best-fitting one as the distribution function of the variable, and this kind of comparative analysis requires a large amount of computation. In this paper, the distribution function of each variable will be derived based on the principle of maximum entropy, which has been shown to have excellent application in many articles. According to the maximum entropy principle and Shannon's entropy formula, the probability density function can be derived using the Lagrange multiplier method as [16, 17]

$$f(x) = \exp\left[-\sum_{i=0}^{m} \lambda_i g_i(x)\right] \tag{3}$$

The distribution function is the

$$F(x) = \int_a^x \exp\left[-\sum_{i=0}^{m} \lambda_i g_i(x)\right] dx \tag{4}$$

3.3 Copula function

The Copula function is an efficient way to construct the joint distribution function of different random variables using their marginal distribution functions, and the random variables can be non-independent. According to Sklar's theory [18], the joint distribution can be decomposed into the marginal distribution function $F(x_1), F(x_2), ..., F(x_n)$ and a Copula function. The mathematical expression is as follows

$$F(x_1, x_1, ..., x_n) = C\left(F(x_1), F(x_2), ..., F(x_n)\right) \tag{5}$$

In this paper, Archimedean Copula functions are also used and Table 1 shows the commonly used 3D Archimedean Copula functions [19, 20].

Table 1: Archimedes Copula function

Name	Copula function $C_\theta(\mu, \nu, \omega)$
Clayton	$(\mu^{-\theta} + \nu^{-\theta} + \omega^{-\theta} - 2)^{-1/\theta}$
Frank	$-\frac{1}{\theta}\ln[1 + \frac{(e^{-\theta\mu}-1)(e^{-\theta\nu}-1)(e^{-\theta\omega}-1)}{(e^{-\theta}-1)^2}]$
Gumbel-Hougaard	$\exp\{-[(-\ln\mu)^\theta + (-\ln\nu)^\theta + (-\ln\omega)^\theta]^{1/\theta}\}$

Note: μ, ν, ω in the table indicates the distribution function of the variable, θ is the Copula function parameter.

ISSN 2616-5805 Vol. 5, Issue 3: 52-59, DOI: 10.25236/AJMS.2024.050308

In this paper, the maximum likelihood estimation method is used to solve the parameter θ in the Copula function, while the root mean square error is used to select the three types of Copula function. The calculations are as follows

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Pe_i - P_i\right)^2} \tag{6}$$

where Pe_i is the empirical probability value derived directly from the sample, and P_i is the theoretical probability value under different Copula function types.

3.4 Joint distribution

From the marginal distribution function of drought duration $F_D(d)$, drought intensity $F_S(s)$ and rate of drought-affected areas $F_I(i)$, and the Copula function, the joint distribution function can be obtained as follows

$$F(d,s,i) = P(D \le d, S \le s, I \le i) = C\left[F_D(d), F_S(s), F_I(i)\right] \tag{7}$$

The joint distribution is the probability of occurrence of a drought event where the drought duration, drought intensity and rate of drought-affected areas are all less than or equal to a given value.

4. Results and analyses

4.1 Drought characterisation and Copula modelling

The average value of the SPI indicator in Yunnan Province was derived from the SPI series of each site, and the drought duration and drought intensity of each year were identified with the threshold value of SPI equal to 0. At the same time, the rate of drought-affected area was calculated with the agricultural data released by the National Bureau of Statistics, and the results of the calculation for Yunnan Province are shown in Fig. 2.

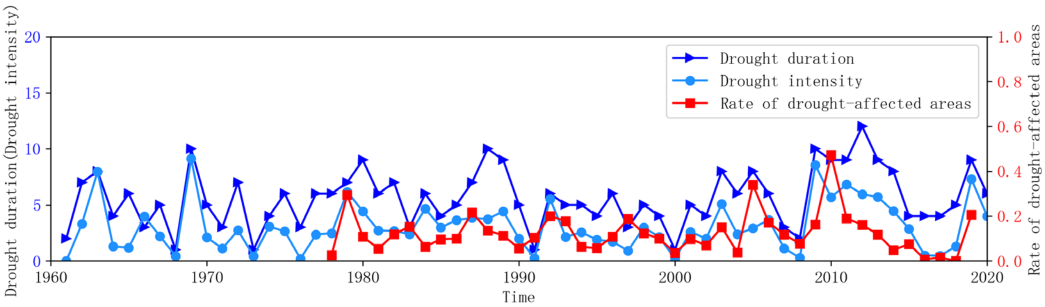


Figure 2: The time sequence of drought duration, drought severity and drought area ratio in Yunnan Province

Due to some missing disaster data, the drought-affected area rate from 1961-1977 is missing in Figure 2, and most provinces have drought-affected area data from 1978. As can be seen in Figure 2, there is some correlation between the variables, and the Pearson correlation coefficient calculation shows that the correlation coefficient between drought duration and drought intensity in Yunnan Province is 0.82, the correlation coefficient between drought duration and the rate of drought-affected area is 0.42, and the correlation coefficient between drought intensity and the rate of drought-affected area is 0.51, and all of them pass the test of a significance level of 0.05. The average drought duration was 5.7 months, the average drought intensity was 3.1, and the average drought area ratio was 0.12.

In the establishment process of Copula analysis model, the distribution function of each random variable should be given firstly, and the distribution function of each variable is derived by the principle of maximum entropy, and the result is shown in Fig. 3.

ISSN 2616-5805 Vol. 5, Issue 3: 52-59, DOI: 10.25236/AJMS.2024.050308

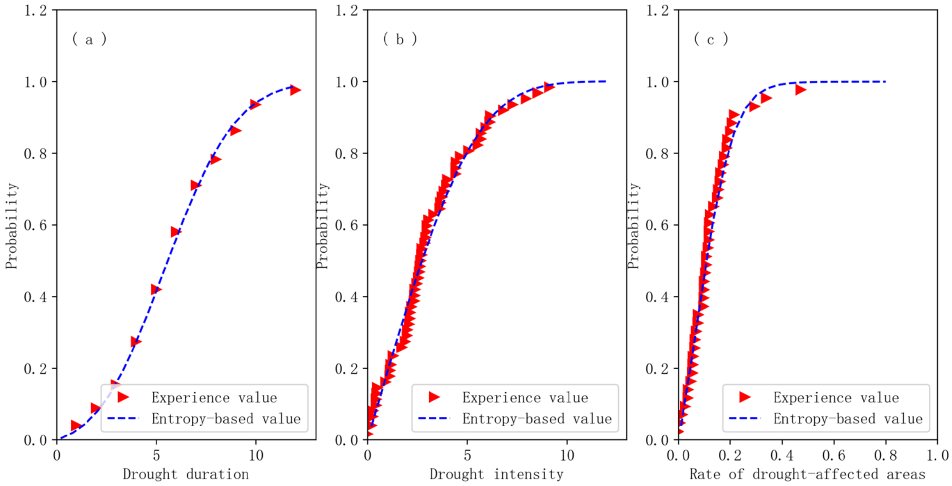


Figure 3: The comparative analysis of fitting (a) drought duration, (b) drought intensity, (C) rate of drought-affected areas

As can be seen in Figure 3, the distribution function derived based on the principle of maximum entropy fits the empirical distribution well, indicating an excellent fit, and the distribution test indicates that the three variables passed the K-S distribution test with a significant level of 0.05.

Based on the theoretical distribution function of each variable, and the Copula function types in Table 1, the maximum likelihood estimation method was used to calculate the θ value of each Copula function type, and the root mean square error between the distribution function values obtained based on the Copula function and the empirical distribution function values was found, and the results are shown in Figure 4.

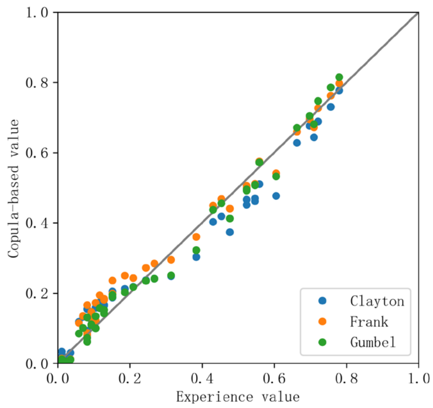


Figure 4: The selection of the Copula function in Yunnan Province

Figure 4 shows the calculation results for Yunnan Province, and the scatter indicates the comparison between the probabilities obtained using different Copula functions and the empirical probabilities, and the closer to the diagonal line indicates the more consistent the theoretical and empirical calculations are. From the position of the scatter points, it can be seen that there is not much difference between the results of the three types of Copula calculations, and all of them are near the diagonal line. The calculation of the root mean square error shows that the error of Gumbel is the smallest and the results are relatively better, and all the subsequent models will be analyzed based on the calculation results of Gumbel.

4.2 Joint probabilistic character analysis

Copula based on Gumbel type was calculated out of the distribution function in three dimensions, and Figure 5 shows the distribution function under three different drought area rates in Yunnan Province.

ISSN 2616-5805 Vol. 5, Issue 3: 52-59, DOI: 10.25236/AJMS.2024.050308

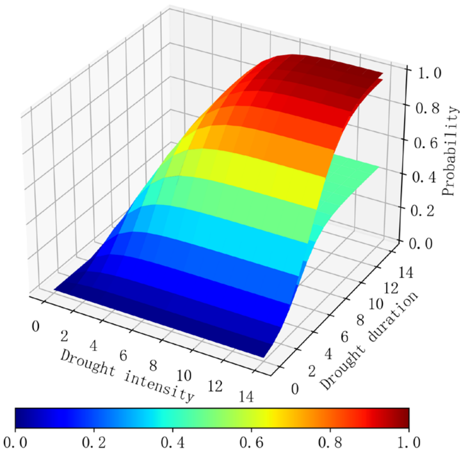
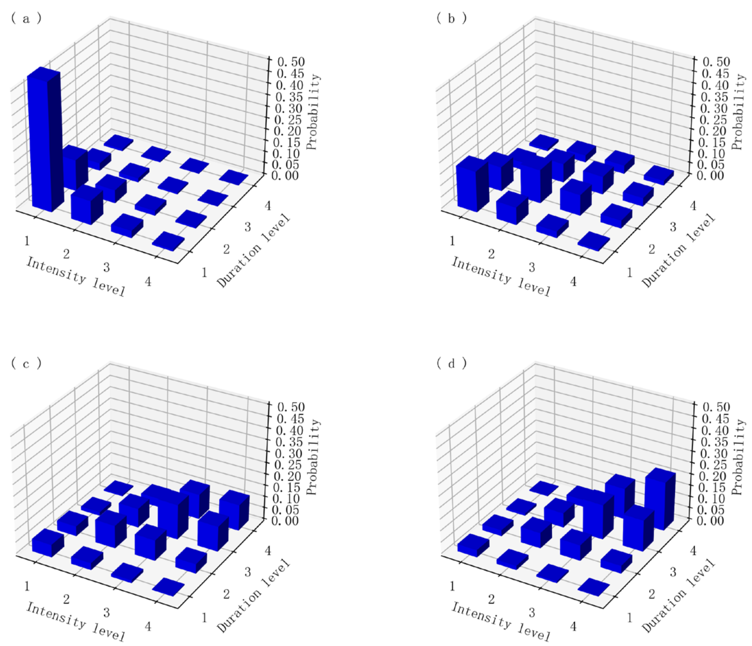


Figure 5: The joint distribution function

According to the national standard "drought disaster level", the drought area rate is divided into 10 per cent, 30 per cent and 50 per cent of the critical values of level 1, 2 and 3, and the surfaces from low to high in Fig. 5 represent the distribution functions $F(d,s,10\%)$, $F(d,s,30\%)$ and $F(d,s,50\%)$, respectively. From Figure 5, it can be seen that the probability of the drought area rate of Yunnan Province being lower than 30% is close to 1, while the probability of the drought area rate being lower than 10% is lower than 0.4. Taking the drought duration of 5, 7 and 9 months as the dividing line, and based on the same probability of dividing the drought intensity and the rate of drought-affected areas, four levels are obtained, as shown in Table 2 below.

Table 2: The classification of drought

Level	Drought duration	Drought intensity	Rate of drought-affected areas,	Probability
1	$D \le 5$	$S \le 2.41$	$I \le 0.09$	0.41
2	$5 < D \le 7$	$2.41 < S \le 4.05$	$0.09 < I \le 0.18$	0.24
3	$7 < D \le 9$	$4.05 < S \le 6.01$	$0.18 < I \le 0.27$	0.20
4	$9 < D$	$6.01 < S$	$0.27 < I$	0.15



(a) $I \le 0.09$ *,(b)* $0.09 < I \le 0.18$ *,(c)* $0.18 < I \le 0.27$ *,(d)* $0.27 < I$

Figure 6: The probability of different drought type

From Table 2, we can see that the proportion of drought events corresponding to level 1 is 0.41, level 2 is about 0.24, level 3 is about 0.2, and level 4 is about 0.15. Fixing the level of the rate of drought-affected, the magnitude of the probability of the drought duration and drought intensity under different levels can be found, and the results are shown in Figure 6.

As can be seen from Figure 6(a), the main cause of the rate of drought-affected of less than 0.09 is drought with a drought duration and drought intensity level of 1. From Figure 6(b), it can be seen that the main causes of the rate of drought-affected between 0.09 and 0.18 are droughts with a drought duration and drought intensity level of 1-3. From Figure 6(c) (d), it can be seen that the main cause of the rate of drought-affected greater than 0.18 is caused by droughts with drought duration and drought intensity level of 3 to 4. Therefore, from the historical data, when the drought duration is greater than 7 months, it is very easy to have a drought with a large affected area.

5. Conclusions

In this paper, the SPI on a 3-month time scale is used as a descriptive indicator of agricultural drought, from which the annual drought duration and drought intensity are separated, and the rate of drought-affected is calculated with the provincial agricultural data published by the National Bureau of Statistics, which is used as an assessment indicator of the impact on agriculture. Then an analytical model of drought duration, drought intensity and the rate of drought-affected was established by using the three-dimensional Copula function, based on which the characteristics of drought in Yunnan Province were analysed. The results show that there is a strong positive correlation between the three variables, and the mean values of drought duration, drought intensity and the rate of drought-affected are 5.7 months, 3.1 and 0.12, respectively. The distribution function based on the principle of maximum entropy can fit the empirical distribution of the three variables very well, and all of them pass the K-S distribution test.

In this paper, the SPI on a 3-month time scale is used as a descriptive indicator of agricultural drought, from which the annual drought duration and drought intensity are separated, and the rate of drought-affected is calculated with the provincial agricultural data published by the National Bureau of Statistics, which is used as an assessment indicator of the impact on agriculture. Then an analytical model of drought duration, drought intensity and the rate of drought-affected was established by using the three-dimensional Copula function, based on which the characteristics of drought in Yunnan Province were analysed. The results show that there is a strong positive correlation between the three variables, and the mean values of drought duration, drought intensity and the rate of drought-affected are 5.7 months, 3.1 and 0.12, respectively. The distribution function based on the principle of maximum entropy can fit the empirical distribution of the three variables very well, and all of them pass the K-S distribution test. Comparative analyses of Copula function types showed that the three Copula function types yielded similar computational results, and RMSE analyses showed that Gumbel was slightly more effective than the other two types. Nearly 60% of the droughts were events with a drought area rate of less than 0.18, and the drought duration and drought intensity level of 3 to 4 were the main reasons for the damage rate of more than 0.18.

Acknowledgements

Funding: This work was funded by National Natural Science Foundation of China (Grant: 42305065), the China Desert Weather Scientific Research Fund (Grant No: Sqj2021016).

References

[1] Gumus V , Dinsever L D , Avsaroglu Y .Analysis of drought characteristics and trends during 1965–2020 in the Tigris River basin, Turkey[J].Theoretical and Applied Climatology, 2023, 151(3-4):1871-1887.DOI:10.1007/s00704-023-04363-x.
[2] Yang J , Yao J .Estimation of multivariate design quantiles for drought characteristics using joint return period analysis, Vine copulas, and the systematic sampling method[J].Journal of water and climate change, 2023, 14(5/6):1551-1568.
[3] Rani S, Kumar P, Dahiya P ,et al.Effect of biopriming and nanopriming on physio-biochemical characteristics of Cicer arietinum L. under drought stress[J].Plant Stress, 2024, 12.DOI:10.1016/j. stress.2024.100466.

ISSN 2616-5805 Vol. 5, Issue 3: 52-59, DOI: 10.25236/AJMS.2024.050308

[4] Rahmawati N , Yasvi A P .Improvement in physio-biochemical characteristics of shallot plants with nano silica at several levels of drought stress[J].IOP Publishing Ltd, 2024.DOI:10.1088/ 1755-1315/ 1302/1/012032.

[5] Wang Z , Chang J , Yangbinghe G G .Temporal and spatial propagation characteristics of meteorological drought to hydrological drought and influencing factors[J].Atmospheric research, 2024, 299(Apr.):107212.1-107212.13.

[6] Tuan N V , Hieu N V , Bang N K ,et al.Spatio-Temporal Analysis of Drought in the North-Eastern Coastal Region of Vietnam Using the Standardized Precipitation Index (SPI)[J].Atmospheric and Climate Science, 2023.DOI:10.4236/acs.2023.132011.

[7] Yang Y C, Yang Jia-zhen, Yi T ,et al.Multidimensional identification of single and regional meteorological drought considering flash-season synthesis employed daily standardized effective precipitation and drought index in Guangxi, China[J].International Journal of Climatology, 2023, 43:2843 - 2861.DOI:10.1002/joc.8004.

[8] Mirzabaev A .A Drought Dataset Based on a Composite Index for the Sahelian Climate Zone of Niger[J].Data, 2023, 8.DOI:10.3390/data8020028.

[9] Reyniers N , Osborn T , Addor N ,et al.Projected changes in droughts and extreme droughts in Great Britain strongly influenced by the choice of drought index[J].Hydrology and Earth System Sciences, 2023.DOI:10.5194/hess-27-1151-2023.

[10] Yeh H F , Lin X Y , Huang C C ,et al.A Meteorological Drought Migration Model for Assessing the Spatiotemporal Paths of Drought in the Choushui River Alluvial Fan, Taiwan[J].geosciences, 2024, 14(4).

[11] Wang L , Zhang X , Wang S ,et al.Analysis and Application of Drought Characteristics Based on Theory of Runs and Copulas in Yunnan, Southwest China[J].International Journal of Environmental Research and Public Health, 2020, 17(13):4654.DOI:10.3390/ijerph17134654.

[12] Wang L , Chen W , Haung G ,et al.Characteristics of super drought in Southwest China and the associated compounding effect of multiscalar anomalies[J].Science China Earth Sciences, 2024(7).DOI:10.1007/s11430-023-1341-4.

[13] Okunola O H , Bako A I .Exploring residential characteristics as determinants of household adaptation to climate change in Lagos, Nigeria[J].International journal of disaster resilience in the built environment, 2023.DOI:10.1108/IJDRBE-06-2021-0060.

[14] Kheyruri Y , Nikaein E , Sharafati A .Spatial monitoring of meteorological drought characteristics based on the NASA POWER precipitation product over various regions of Iran[J].Environmental science and pollution research international, 2023.DOI:10.1007/s11356-023-25283-3.

[15] Lap N V , Oanh T T K , Man D B ,et al.Sedimentary Characteristics of Shallow Aquifers and Suitability to Irrigation in the Drought Season: The Case of the Fruit Tree Area in Ben Tre Province, Mekong River Delta[J].Earth Sciences, 2023.DOI:10.11648/j.earth.20231201.13.

[16] Iida S , Kenzo T , Shimizu T ,et al.Evaluation of water use characteristics focusing on isohydry and anisohydry for an evergreen broad-leaved species, Castanopsis sieboldii (Makino) Hatus. based on sap flow measurements[J].Journal of Japanese Association of Hydrological Sciences, 2023, 53:43-54.DOI:10.4145/jahs.53.43.

[17] Zhang X , Sun W .Spatial and Temporal Characterization of the Urban Drought in the Western Cape, South Africa, from 2015 to 2017[J].Advances in Earth Science, 2023, 38(5):493-504. DOI:10. 11867/j.issn.1001-8166.2023.020.

[18] Shafiei M , Moghaddasi M , Naderi K .Projections of drought characteristics based on combined drought index under CMIP6 models[J].Water practice and technology, 2023, 18(11):2818-2833.

[19] Gond S , Gupta N , Patel J ,et al.Spatiotemporal evaluation of drought characteristics based on standard drought indices at various timescales over Uttar Pradesh, India[J].Environmental Monitoring and Assessment, 2023, 195(3).DOI:10.1007/s10661-023-10988-2.

[20] Wang Y , Wang J , Zhang Q .Analysis of ecological drought risk characteristics and leading factors in the Yellow River Basin[J].Theoretical and Applied Climatology, 2024, 155(3):1739-1757.