J. COMBIN. MATH. COMBIN. COMPUT. 127b (2025) 5049--5065

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Selection of Typical Meteorological Cycles Based on Comparison of Cumulative Distribution Functions of Different Meteorological Indicators and Assignment of Normalized Modeling Research

Zesen Wang[1,2], Shuaihao Kong[1,2], Qi Li[1,2], Jingrong Guo[1,2], Hao Liang[1,2], Tianqi Zhao[1,2], Jiayu Ding[3], Runfeng Zhang[3,] ✉

[1] *North China Electric Power Research Institute Co., Ltd., Beijing, 100045, China*

[2] *State Grid Jibei Electric Power Research Institute, Beijing, 100052, China*

[3] *Nanjing Tode Technology Co., Ltd., Nanjing, Jiangsu, 210094, China*

ABSTRACT

Liaoning region is selected as the study area, and its meteorological data from 1974 to 2024 are used as the study samples. Based on the four indicators, SPI, SPEI, EDDI and CJDI, the normalized composite drought characteristic indicators were constructed by using the CVine joint function and entropy weighted TOPSIS, so as to explore the drought calendar and drought intensity in the Liaoning region, and to analyze the spatial and temporal evolution of the drought cycle. The results showed that the Kendall and Spearman rank correlation coefficients reached above 0.60 and 0.73, respectively, and therefore, the drought duration and drought intensity were strongly correlated. The normalized composite drought characteristics index had a significant negative correlation association with SPI (P<0.01). The normalized composite drought characteristic index has a significant positive correlation association with SPIE (P<0.05). SPI and SPEI are one of the important reasons to study the spatial distribution and temporal pattern of regional drought.

Keywords: copula function, normalization, drought indicator, entropy weighted TOPSIS

1. Introduction

Typical Meteorological Year (TMY) is an outdoor meteorological dataset used for building energy simulation, and the accuracy of TMY will directly affect the accuracy of building energy simulation and building energy efficiency design [1-2]. At present, the generation theory of TMY in China mainly refers to the Sandia method in the United States, but the theory has two defects: first, it gives

✉ Corresponding author.
E-mail address: ZrunFfeng@163.com (R. Zhang).
Received 01 March 2024; Revised 25 May 2024; Accepted 10 November 2024; Published Online 16 April 2025.
DOI: 10.61091/jcmcc127b-281

solidified weights to different meteorological parameters, and it does not take into account the climatic characteristics of different regions [3-6]. Secondly, the research field of new energy buildings has been developing rapidly in recent years, and the traditional Sandia method is not well applicable, and the theory of TMY generation for new energy buildings is still blank [7-9].

Internationally, when generating TMY, day-by-day data are usually used, and some of them use hour-by-hour or month-by-month data. According to the meteorological data information in China, the continuity and comprehensiveness of the daily value data are better, and also have meteorological parameters such as daily minimum and daily maximum, which can be more comprehensive in the analysis [10-13]. Moreover, at present, when scholars at home and abroad generate TMY, the most used data are also daily value data, coupled with the fact that China has a record of basically four times the timing data, and the lack of comprehensiveness and long-term continuity, it is difficult to be used for the analysis [14-16]. The generation of TMY usually requires about 30 years of data, the advantages of using daily value data set to select the TMM are greater. Building energy consumption accounts for a relatively large proportion of the total energy consumption of Chinese society, so the development of clean energy in buildings has ushered in a golden period [17-20]. New energy buildings have appeared in large numbers, but the research on TMY for new energy buildings has always been a gap, and the traditional TMY generation methods developed for traditional buildings are not applicable to new energy buildings, and there is an urgent need to develop a set of TMY generation methods for new energy type buildings [21-24].

In this paper, the meteorological data from 25 meteorological stations in Liaoning area are taken as research samples. Based on the Copula function, a joint function characterizing precipitation, temperature and other key effects of typical meteorological cycle-drought is constructed. Four drought indices, SPI, SPEI, EDDI, and CJDI, were used to analyze the drought duration and drought intensity under different drought indices, and Kendall and Spearman were chosen to verify the correlation between them. Based on the four drought indices, the constructed normalized drought composite index is assigned by entropy weight-TOPSIS assignment to the normalized drought composite index with suitable weight combinations for Liaoning region. Finally, the spatial and temporal characteristics of drought in Liaoning region and the evolution of drought center with time are analyzed.

2. Selection of meteorological indicators and techniques related to normalized modeling

2.1. Cumulative distribution function - Copula function

Typical meteorological cycle - drought is a phenomenon of dry air, soil water shortage or even drying up due to prolonged non-precipitation or abnormally low precipitation on a regional scale, which is divided into two manifestations of arid climate and drought disaster in meteorology. China is a large agricultural country, and at the same time is a drought disaster-prone countries, throughout its history, every drought disaster occurred, have a serious impact on agricultural production, directly hindering the pace of social and economic development. The reason why the losses of drought disasters are so serious is precisely because of the high frequency of drought, the wide range of impacts, the long duration and the great impact of the aftermath.

2.1.1. Copula function definition. Suppose $X_1,\cdots,X_d$ is a random variable with marginal distribution $F_1,\cdots,F_d$ and joint distributions $F\left(x_1,\cdots,x_d\right)$, $u_1=F_1\left(x_1\right),\cdots,u_{\mathrm{d}}=F_d\left(x_d\right)$. If $X_1,\cdots,X_d$ are independent of each other, then the joint distribution $F\left(x_1,\cdots,x_d\right)=\prod_{\mathrm{i}=1}^{d}F_i\left(x_i\right)$, where

$C\left(u_1,\cdots,u_d\right)=\prod_{i=1}^{d}u_i$ is called independent. By Sklar theory, there exists CopulaC, for all $x \in R$, $R \in (-\infty,+\infty)$, the relationship between the multivariate distribution $F\left(x_1,\cdots,x_d\right)$ and the Copula distribution $C\left(u_1,\cdots,u_d\right)$ is:

$$\begin{aligned} F\left(x_1,\cdots,x_d\right) &= P\left(X_1 \le x_1,\cdots,X_d \le x_d\right) \\ &= C\left[F_1\left(x_1\right),\cdots,F_d\left(x_d\right)\right] \\ &= C\left(u_1,\cdots,u_2\right) \end{aligned} \tag{1}$$

where $F_i\left(x_i\right)=u_i$, $i=1,2,\cdots,d$; F_i are continuous functions and $U_i \sim U(0,1)$.

2.1.2. Binary Copula Functions. The binary Copula distribution and density function are expressed as:

$$C(u,v)=F\left(F_1^{-1}(u),F_2^{-1}(v)\right),u=F_1\left(x_1\right),v=F_2\left(x_2\right) \tag{2}$$

$$c(u,v)=\frac{\partial^2 C(u,v)}{\partial u \partial v}=\frac{f\left(x_1,x_2\right)}{f_1\left(x_1\right)f_2\left(x_2\right)} \tag{3}$$

There are many kinds of Copula, among which the Archimedean Copula function contains only one parameter, the function construction and solution is relatively simple, and mostly applied to binary frequency analysis, can construct a variety of forms, adaptable binary joint distribution function, this paper selects three typical Archimedean Copula functions: Gumbel Copula, Clayton Copula and Frank Copula function for the joint analysis of drought duration and drought intensity. In this paper, three typical Archimedean Copula functions are chosen: Gumbel Copula, Clayton Copula and Frank Copula functions to jointly analyze the drought calendar time and drought intensity.

2.1.3. Vine-Copula function. Vine-Copula function is a constructive multivariate Copula modeling function. It is mainly modeled based on conditional Copula function and Vine (Vine) graph, and Pair Copula is used to construct the joint distribution of binary variables, which constitutes the Vine structure, to establish the Vine-Copula function. In this paper, the commonly used binary GumbelCopula, Clayton Copula and Frank Copula functions are selected to construct Pair Copula. CVine-Copula is mostly used to fit multivariate variables where there are dominant other factors, umbrella construction, and DVine-Copula is mostly used to fit multivariate variables that are independent of each other, linear construction. Due to the relative simplicity of three-dimensional variable construction, CVine-Copula and DVine-Copula are similar in structure, and this paper analyzes the drought events in Liaoning, and considers that there exists a certain factor in the data set of precipitation, temperature, etc. that has a key factor that influences the other factors, i.e., the main factor that causes the occurrence of drought events, and the expression of its probability density function is as follows:

$$\begin{aligned} f_{1,2\cdots n} &= \prod_{k=1}^{n} f_k \prod_{j=1}^{n-1}\prod_{i=1}^{n-j} c_{j,j+1|1,\cdots,j-1} \\ &\left\{F\left(x_j \mid x_1,\cdots,x_{j-1}\right),F\left(x_{j+i} \mid x_1,\cdots,x_{j-1}\right)\right\} \end{aligned} \tag{4}$$

Where: n -Dimension of the multivariate variable; j -Levels of the vine-structured tree; i -Number of nodes in each level of the tree.

2.1.4. Correlation measures and goodness-of-fit tests. The correlation measure can be used to show the correlation between drought duration and drought intensity. Let the random variables X,

Y have a joint distribution function Copula C, then the Kendall rank correlation coefficient τ and Spearman rank correlation coefficient ρ_n are defined as follows:

$$\tau = 4\int_0^1\int_0^1 C(u,v)dC(u,v) - 1 \tag{5}$$

$$\rho_n = 12\int_0^1\int_0^1 uvdC(u,v) - 3 \tag{6}$$

In this paper, we choose the root mean square error (RMSE) and the Akaike informativeness criterion (AIC) to test the fit of the theoretical distributions. The smaller values of RMSE and AIC indicate that the theoretical and empirical distributions are better fitted with the formula:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left[F_e\left(x_1,\cdots,x_d\right) - F\left(x_1,\cdots,x_d\right)\right]^2 \tag{7}$$

$$RMSE = \sqrt{MSE} \tag{8}$$

$$AIC = N\log\left(MSE\right) + 2k \tag{9}$$

Where: MSE -mean square error; N -number of samples; $F_e\left(x_1,\cdots,x_d\right)$-empirical distribution; $F\left(x_1,\cdots,x_d\right)$-joint distribution; k -number of Copula C parameters.

2.2. Monitoring and analysis of normalized typical meteorological cycles

2.2.1. Determination of Normalized Drought Composite Index (SDCI). The occurrence and development of drought is a complex coupled process that is difficult to monitor directly, and vegetation growth conditions, soil moisture supply and demand, and atmospheric precipitation are all indispensable factors.

Comprehensive monitoring characteristics of each parameter index and previous experience, integrating intrinsic precipitation to calculate the standardized precipitation index (SPI), precipitation and ET0 to calculate the standardized temperature (SPEI), ET0 to calculate the evapotranspiration-demand drought index (EDDI), and the Copula function of the CJDI to calculate the normalized drought composite index (SDCI) of the Liaoning area, and choosing the entropy weight-TDI, which is a highly credible, objective and simple to calculate, for decision-making results, is used. The entropy weight-TOPISI, which is strong and simple to calculate, was chosen to determine the weights of each index, and the weight coefficients of SPI, SPEI, EDDI and CJDI indices were set to a,b,c,d and $a+b+c+d=1$, respectively, and the specific formulas were as follows:

$$\mathrm{SDCI} = \mathrm{SPI}\times\mathrm{weight} + \mathrm{SPEI}\times\mathrm{weight} + \mathrm{E}DDI\times\mathrm{weight} + \mathrm{C}JDI\times\mathrm{weight} \tag{10}$$

1) Meteorological drought index - Standardized Precipitation Index (SPI) Meteorological drought refers to the phenomenon of shortage of surface water due to the imbalance between evapotranspiration and precipitation during a certain period of time, where water expenditure is greater than water income. Standardized precipitation index (SPI) is used to characterize the probability of occurrence of precipitation in a certain period of time, according to the national standard of Meteorological Drought Level (GB/T 20481-2017), the calculation principle, method and division of drought level of the SPI index are specified as follows: assuming that the amount of precipitation in a certain period of time is a random variable x, then the probability density function equation of its Γ distribution:

$$\mathrm{f}(x) = \frac{1}{\beta^y\Gamma(y)}x^{y-1}e^{\frac{x}{\beta}}, x>0 \tag{11}$$

2) Standardized Precipitation Evapotranspiration Index (SPEI) is an index standardized by normalizing the cumulative probability of the difference between precipitation and potential evapotranspiration, which is one of the more ideal indexes for drought monitoring, taking into account the precipitation and temperature factors, the response of drought to evapotranspiration as well as the spatial consistency and multi-temporal scale, etc. Therefore, the cumulative distribution function was used to standardize the SPEI. Therefore, the cumulative distribution function was used to standardize and finally calculate the monthly-scale SPEI index, and then the spatial distribution of SPEI in Liaoning was obtained with the help of Anusplin software, combined with the interpolation of the digital elevation data (DEM), with the specific equations as follows:

$$SPEI = W - \frac{C_0 + C_1W + C_2W^2}{1 + d_1W + d_2W^2 + d_3W^3} \tag{12}$$

Where: W is the value of the cumulative probability function of the evapotranspiration precipitation derivative function, $W = \sqrt{-2\ln P}$, P is the cumulative probability density, C_0, C_1, C_2, d_1, d_2, d_3 is a constant term, $C_0 = 2.515517$, $C_1 = 0.802853$, $C_2 = 0.010328$, $d_1 = 1.432788$, $d_2 = 0.189269$, $d_3 = 0.001308$. The five drought classes of the SPEI are: exceptional drought, severe drought, moderate drought, mild drought, and as no drought.

The calculated EDDI can be compared with other standardized indices such as SPI and SPEI. The formula is as follows:

$$\mathrm{EDDI} = \mathrm{W} - \frac{\mathrm{C}_0 + \mathrm{C}_1\mathrm{W} + \mathrm{C}_2\mathrm{W}^2}{1 + \mathrm{d}_1W + \mathrm{d}_2W^2 + \mathrm{d}_3W^3} \tag{13}$$

The CJDI is a drought index that combines the SPI, SPEI and EDDI, which represent drought from different perspectives (precipitation and evaporation). Empirical Copula was chosen when sufficient data were available because of its advantages of being intuitive and simple to use.The CJDI was calculated by applying this cumulative probability to the inverse function of the cumulative standard normal distribution:

$$\mathrm{CJDI} = \mathrm{N}^{-1}(\mathrm{K}_c(q)) = \mathrm{N}^{-1}(P[u_1, u_2(u_1, u_2) \le q]) \tag{14}$$

2.2.2. Determination of drought indicator weights - entropy weight TOPSIS：

1) Construct the evaluation matrix. Assuming that there is m object to be evaluated, each evaluation object has n evaluation indicators, then X_{ij} is the i th evaluation object of the j th indicator data, which $i = (1, 2, \cdots, m)$, $j = (1, 2, \cdots, n)$. The original evaluation matrix R_x is:

$$R_x = \begin{bmatrix} X_{11} & \cdots & X_{1n} \\ \vdots & X_{ij} & \vdots \\ X_{m1} & \cdots & X_{mn} \end{bmatrix} \tag{15}$$

2) Data standardization. The purpose of data standardization processing is to solve the problem of inconsistency of the magnitude in the original matrix, and this step will treat the positive and negative evaluation indicators differently, eliminating the impact of inconsistent magnitude. In addition, since in some cases, the normalized data may fall between 0 and 1. Therefore, in order to eliminate the effect of 0, this step can also shift the matrix uniformly to the right by 0.001 units to finally obtain the new normalized matrix R_y. Specific expression:

Positive Indicator:

$$Y_{ij} = \frac{X_{ij} - \min X_{ij}}{\max X_{ij} - \min X_{ij}} + 0.001 \tag{16}$$

Negative indicators:

$$Y_{ij} = \frac{\max X_{ij} - X_{ij}}{\max X_{ij} - \min X_{ij}} + 0.001 \tag{17}$$

The new matrix R_y that is:

$$R_y = \begin{bmatrix} Y_{11} & \cdots & Y_{1n} \\ \vdots & Y_{ij} & \vdots \\ Y_{m1} & \cdots & Y_{mn} \end{bmatrix} \tag{18}$$

3) Calculate the entropy value and weight of the indicators. First calculate the weight P_{ij} that each item Y_{ij} accounts for;

$$P_{ij} = \frac{Y_{ij}}{\sum_{i=1}^{m} Y_{ij}} \tag{19}$$

Then use the original formula of entropy weighting method to calculate the entropy value of each of the n evaluation indexes of m evaluation object in the matrix e_j. When the information entropy value of the indexes is smaller, the information utility contained in them is larger, and the weighting coefficient is larger. Conversely, when the information entropy value of the indicator is larger, the information utility contained is smaller, and the weight coefficient is smaller:

$$e_j = -\frac{1}{\ln n}\left(\sum_{i=1}^{m} P_{ij} \times \ln P_{ij}\right) \tag{20}$$

Finally, the weight w_j for each indicator is derived and is calculated using the following formula:

$$w_j = \frac{1 - e_j}{\sum_{j=1}^{n}\left(1 - e_j\right)} \tag{21}$$

In Eq:

w_j - the weight corresponding to each evaluation object.

e_j - entropy value.

4) Construction of weighted standardized matrix. Here, the matrix obtained from the previous standardization process will be weighted, the weight given is the weight obtained by the entropy weight method, and the relevant calculation is the corresponding weight w_j multiplied by the corresponding value of the standardized matrix Y_{ij}. This process will result in a new weighted standardized matrix R_z:

$$Z_{ij} = w_j \times Y_{ij} \tag{22}$$

In Eq:

Z_{ij} - the j th indicator data for each i evaluation objects in matrix R_z.

w_j - the weight corresponding to each evaluation object.

Y_{ij} - data of the j th indicator of the i th object of evaluation in matrix R_y.

$$R_z = \begin{bmatrix} Z_{11} & \cdots & Z_{1n} \\ \vdots & Z_{ij} & \vdots \\ Z_{m1} & \cdots & Z_{mn} \end{bmatrix} \tag{23}$$

5) Determine the positive and negative ideal solutions. Here the positive and negative ideal solutions of the indicators will be determined, the positive ideal solution Z^+ is the maximum value of the n indicator in the m evaluation object, and the negative ideal solution Z^- is the minimum value of the n th indicator in the m th evaluation object:

$$Z^+ = \left\{\left(\max_i Z_{ij} \mid j \in \text{Positive Indicators}\right), \left(\min_i Z_{ij} \mid j \in \text{Negative indicators}\right)\right\} \tag{24}$$

$$Z^- = \left\{\left(\min_i Z_{ij} \mid j \in \text{Positive Indicators}\right), \left(\max_i Z_{ij} \mid j \in \text{Negative indicators}\right)\right\} \tag{25}$$

6) Calculate the Euclidean distance. Here the distance formula will be used to calculate the Euclidean distances D^+ and D^- for each evaluation program and indicator from the positive and negative ideal solutions:

$$\begin{gathered} D^+ = \sqrt{\sum_{j=1}^{n}\left(Z_{ij} - Z^+\right)^2} \\ D^- = \sqrt{\sum_{j=1}^{n}\left(Z_{ij} - Z^-\right)^2} \end{gathered} \tag{26}$$

Eq:
$D^{+/-}$ - Euclidean distance of positive/negative ideal solutions.
Z_{ij} - the j th indicator data for each i evaluation objects in the matrix Rz.
$Z^{+/-}$ - positive/negative ideal solution.

7) Deriving the relative closeness. Here, the obtained distance values will be used to calculate the relative closeness C_j of each evaluation object to the optimal object and to rank the indicators according to their respective closeness, when the larger the value of C indicates that the closer the evaluation object is to the optimal object, i.e., the closer the corresponding indicator is to the optimal solution, and vice versa:

$$C_j = \frac{D^-}{D^+ + D^-} \tag{27}$$

In Eq:
C_j - the relative closeness of the evaluation object to the optimal object.
$D^{+/-}$ - Euclidean distance of positive/negative ideal solutions.

3. Research on normalized modeling of typical weather cycle-drought cycle

3.1. Trends in the duration and intensity of meteorological cycles in the Liaoning region

Drought is frequent in Liaoning Province, and the frequency of larger and higher droughts has increased from 20% in 1961-1980 to 53% in 2001-2024. The cumulative area of affected crops has reached 0.18 billion hm², and the cumulative economic losses have amounted to 42 billion, making the task of combating drought very arduous. Agriculture, as the area most seriously affected by drought, is closely related to agricultural production, agricultural life and rural ecological environment. According to statistics, from 1999-2024, the cause of drought is low precipitation, but there is a slow and gradual process for it to develop into a disaster, which depends not only on meteorological conditions such as precipitation and temperature.

1) Meteorological data sources. The meteorological data used in the study were selected from 25 meteorological stations with long series of meteorological data in Liaoning area. The meteorological data (including average humidity, sunshine hours, temperature, wind speed and precipitation) are on a daily scale, sourced from the China Meteorological Data Network (http://data.cma.cn/), and the data have passed strict quality control. The meteorological data cover the period of 1974-2024, and

the data records are relatively complete, with the missing ratios of average humidity, sunshine hours, temperature and wind speed less than 1%, and the missing ratio of precipitation less than 10%, and the missing data are made up by interpolating and extending the data of the nearby stations on the same day, which are used as the basis for calculating the drought indices.

2) Data pre-processing. The day-by-day observation data of 25 basic meteorological stations in Liaoning Province from 1974 to 2024 are sorted and processed to calculate the precipitation on monthly, seasonal and annual scales, and the day-by-day meteorological data are used to calculate the daily ET0, the standardized precipitation index (SPI) is calculated based on the precipitation, the standardized temperature (SPEI) is calculated based on the precipitation and the ET0, and the evapotranspiration-demand drought index (EDDI) is calculated based on the ET0, and the standardized temperature is calculated based on the SPI and the EDDI. SPI and EDDI were used to calculate CJDI based on Copula function, so as to get the drought index values in different time scales. Day-by-day meteorological data organization: in the original meteorological data, there are pure fog dew frost, snow, trace precipitation, rain plus snow and other special marking values of climate data, and Matlab software is used to eliminate the abnormal values to get the climate data without abnormal values. The index drought class division is shown in Table 1.

Table 1. SPI, SPEI, EDDI and CJDI exponential drought classification

Exponential range	SPI	SPEI	EDDI	CJDI
[-0.5,+∞)	Duration			
[-1,-0.5)	Frequency			
[-1.5,-1)	Strength			
[-2,-1.5)	Peak			
[-∞,-2)	Intensity			

3.1.1. Correlation coefficient test. In this paper, the Kendall rank correlation coefficient and Spearman rank correlation coefficient were used to measure the correlation between drought duration and drought intensity calculated by SPI-3, SPEI-3, EDDI-3 and CJDI-3 at all stations, and the larger the rank correlation coefficient is, the stronger the correlation between the two variables is. The results are shown in Table 2.

It can be seen that the Kendall rank correlation coefficient reaches 0.60, and the Spearman rank correlation coefficient reaches more than 0.73. The maximum value of the Kendall rank correlation coefficient of drought duration and drought intensity of SPI is 0.8557 at Xingcheng station, and the maximum value of the Spearman rank correlation coefficient of 0.9762 also appears at Xingcheng station. The maximum values of the two rank correlation coefficients of SPEI also appeared in Xingcheng station, which was the same as that of SPI. The maximum values of the Kendall rank correlation coefficient and Spearman rank correlation coefficient of EDDI appeared in Jianping station, which were 0.9223 and 0.9957.The maximum values of Kendall rank correlation coefficient and Spearman rank correlation coefficient of CJDI appeared in Dalian station, which were 0.9209, 0.9937, respectively, indicating that the drought duration and intensity of drought were the same. It shows that drought duration and drought intensity are strongly correlated, and the Copula function can be applied to establish a joint distribution model in Liaoning.

Table 2. The correlation coefficient of each station

	SPI		SPEI		EDDI		CJDI	
	K	S	K	S	K	S	K	S
Chanwu	0.7004	0.8763	0.8239	0.9565	0.8183	0.9629	0.8343	0.96

								27
Chang diagram	0.763	0.9306	0.7616	0.9205	0.7988	0.9457	0.7922	0.9391
Progenesis	0.716	0.8993	0.7298	0.8967	0.8385	0.9661	0.8467	0.9683
Qing yuan	0.7756	0.9294	0.7842	0.9346	0.7993	0.946	0.8383	0.9702
Morning sun	0.6984	0.8862	0.7958	0.9492	0.8076	0.9506	0.8761	0.984
Jianping	0.6747	0.8554	0.8703	0.9814	0.9223	0.9957	0.87	0.9779
Xinmin	0.8157	0.9592	0.8421	0.9653	0.8814	0.9836	0.8714	0.9825
Anshan	0.7983	0.9429	0.9049	0.9888	0.8393	0.9721	0.9024	0.9887
Shenyang	0.7372	0.7362	0.7907	0.9465	0.7438	0.9112	0.8612	0.979
Benxi	0.7821	0.9418	0.8343	0.9676	0.8302	0.9621	0.8601	0.9786
Fushun	0.7973	0.9423	0.8202	0.9613	0.7526	0.9172	0.8543	0.9716
New bin	0.7499	0.9127	0.8018	0.9448	0.8556	0.9736	0.8061	0.9484
Jianchang	0.6765	0.8569	0.7996	0.9436	0.8305	0.9684	0.8422	0.9663
Suijin	0.8131	0.951	0.7832	0.9424	0.7087	0.8968	0.8679	0.9771
Xingcheng	0.8557	0.9762	0.8892	0.9873	0.8432	0.9682	0.8998	0.9908
Dais	0.7241	0.8944	0.8365	0.9685	0.7385	0.9075	0.8398	0.9652
Camp	0.7373	0.9141	0.8193	0.9609	0.7095	0.8843	0.8619	1.0008
Urine	0.8171	0.953	0.8362	0.9625	0.7049	0.8806	0.8696	0.9777
Jade	0.7979	0.9411	0.8513	0.9694	0.846	0.9748	0.8863	0.9837
Meadow	0.6837	0.8629	0.765	0.9227	0.8113	0.9526	0.8094	0.9502
Danton	0.7578	0.9274	0.7301	0.8988	0.8307	0.9684	0.8795	0.9851
House shop	0.7664	0.9327	0.8337	0.9673	0.8367	0.971	0.8948	0.9865
Long sea	0.799	0.9509	0.8085	0.9485	0.7634	0.9334	0.8758	0.98
Zhuang river	0.7805	0.9324	0.8111	0.9485	0.8792	0.9866	0.8128	0.952
Dalian	0.7726	0.9363	0.8021	0.945	0.8488	0.9707	0.9209	0.9937

3.1.2. Comparison and selection of Copula functions. The results of the squared Euclidean distance, optimal Copula function and its parameters for each station of SPI are shown in Table 3.There are 12 stations of SPI whose optimal function is GumbelCopula, and all of them show strong upper tail correlation (Upper), and these stations tend to have long drought ephemeral time and strong drought intensity occurring at the same time. The optimal Copula function for 12 stations is FrankCopula, and these stations have no significant tail correlation. Only the Cuiyan station, whose optimal function is the ClaytonCopula function, shows a Lower tail correlation (Lower), which favors short drought durations and smaller drought intensities at the same time.

Table 3. Copula functions and corresponding parameters (SPI)

	Square o 's distance d	Optimal copula function	Parameter
Chanwu	0.0133	Gumbel	3.2831
Chang diagram	0.0044	Frank	14.6789
Progenesis	0.0064	Frank	11.9337
Qing yuan	0.0063	Gumbel	4.3598
Morning sun	0.0058	Frank	11.1158
Jianping	0.0057	Gumbel	3.0278
Xinmin	0.0047	Frank	19.334
Anshan	0.0054	Gumbel	4.8388
Shenyang	0.0059	Frank	13.0545
Benxi	0.0045	Frank	16.1147
Fushun	0.0047	Gumbel	4.8161
New bin	0.006	Gumbel	3.9211
Jianchang	0.009	Gumbel	3.0443
Suijin	0.0046	Gumbel	5.2121
Xingcheng	0.003	Frank	25.0278
Dais	0.0087	Gumbel	3.5602
Camp	0.0057	Frank	13.0585
Urine	0.0062	Gumbel	5.3211
Jade	0.0066	Clayton	7.6554
Meadow	0.0075	Gumbel	3.1131
Danton	0.0043	Frank	14.3261
House shop	0.0045	Frank	14.9174
Long sea	0.0064	Frank	17.6081
Zhuang river	0.0069	Gumbel	4.4549
Dalian	0.0047	Frank	15.3671

The results of the squared Euclidean distance, optimal Copula function and its parameters for each station in SPEI are shown in Table 4.There are 13 stations in SPEI whose optimal function is Gumbel Copula, and all of them show a strong upper tail correlation (Upper), and these stations tend to have long drought ephemeral time and strong drought intensity occurring at the same time. There are 10 stations with Frank Copula as the optimal function, and these stations have no obvious tail correlation. Only Kaiyuan and Zhuanghe stations, where the optimal function is Clayton, show Lower tail correlation (Lower), and tend to have short drought duration and smaller drought intensity at the same time.

Table 4. Copula functions and corresponding parameters (SPEI)

	Square o 's distance d	Optimal copula function	Parameter

Chanwu	0.0038	Gumbel	5.5228
Chang diagram	0.0041	Gumbel	4.1091
Progenesis	0.0056	Clayton	5.2682
Qing yuan	0.0075	Gumbel	4.5295
Morning sun	0.0058	Frank	17.3056
Jianping	0.0038	Frank	27.9487
Xinmin	0.0052	Gumbel	6.1408
Anshan	0.0028	Gumbel	9.995
Shenyang	0.0066	Frank	16.8413
Benxi	0.0052	Frank	21.6492
Fushun	0.0031	Frank	19.8526
New bin	0.0053	Gumbel	4.9229
Jianchang	0.0035	Gumbel	4.8691
Suijin	0.0063	Frank	16.2048
Xingcheng	0.0025	Frank	32.797
Dais	0.005	Frank	21.9672
Camp	0.0033	Frank	19.7494
Urine	0.007	Gumbel	5.925
Jade	0.002	Gumbel	6.5068
Meadow	0.0072	Gumbel	4.1664
Danton	0.0169	Gumbel	3.6376
House shop	0.0036	Frank	21.5648
Long sea	0.0077	Gumbel	5.0895
Zhuang river	0.0112	Clayton	8.3129
Dalian	0.0112	Gumbel	4.9296

The results of squared Euclidean distance, optimal Copula function and its parameters for each station in EDDI are shown in Table 5, it can be seen that there are 14 stations in EDDI where the optimal function is Gumbel Copula, and all of them show a strong Upper tail correlation (Upper), and these stations tend to have long drought ephemeral time and strong drought intensity occurring at the same time. There are 9 stations whose optimal Copula function is Frank Copula, and these stations have no obvious tail correlation. Only Yingkou and Xiongyue stations, where the optimal Copula function is Clayton Copula function, show Lower tail correlation (Lower), and these stations tend to have short drought duration and less drought intensity at the same time.

Table 5. Copula functions and corresponding parameters (EDDI)

	Square o 's distance d	Optimal copula function	Parameter
Chanwu	0.0048	Frank	19.7477
Chang diagram	0.0071	Gumbel	4.874
Progenesis	0.0031	Gumbel	6.0423
Qing yuan	0.004	Gumbel	4.8865
Morning sun	0.007	Gumbel	5.0936
Jianping	0.0021	Gumbel	12.2418
Xinmin	0.0038	Gumbel	8.1564
Anshan	0.0032	Frank	22.5092
Shenyang	0.018	Gumbel	3.8447
Benxi	0.0052	Gumbel	5.754
Fushun	0.0099	Gumbel	3.979
New bin	0.0032	Gumbel	6.7418

Jianchang	0.0069	Frank	21.2792
Suijin	0.0123	Frank	11.633
Xingcheng	0.0026	Gumbel	6.2192
Dais	0.007	Gumbel	3.7671
Camp	0.0069	Clayton	4.7934
Urine	0.0099	Clayton	4.6868
Jade	0.0041	Frank	23.5544
Meadow	0.006	Gumbel	5.1912
Danton	0.0052	Frank	21.2947
House shop	0.0052	Frank	22.1288
Long sea	0.0098	Frank	14.7729
Zhuang river	0.0041	Frank	30.3043
Dalian	0.0128	Gumbel	6.4456

The squared Euclidean distances, optimal Copula functions and their parameters for each station in CJDJ are shown in Table 6. 19 of the 25 stations in CJDI have the optimal function as GumbelCopula, and all of them show a strong upper-tailed correlation (Upper), and these stations tend to have long drought ephemeral times and strong drought intensities occurring at the same time. The optimal Copula function for five stations is FrankCopula, with no obvious tail correlation. Only the optimal function of Changtu station is ClaytonCopula function, which shows the Lower tail correlation (Lower), and tends to have short drought duration and smaller drought intensity at the same time.

Table 6. Copula functions and corresponding parameters (CJDJ)

	Square o 's distance d	Optimal copula function	Parameter
Chanwu	0.018	Gumbel	5.9284
Chang diagram	0.0137	Clayton	7.4873
Progenesis	0.0053	Gumbel	6.3979
Qing yuan	0.0052	Frank	22.5058
Morning sun	0.0925	Frank	29.7895
Jianping	0.0072	Gumbel	7.5226
Xinmin	0.0044	Gumbel	8.6207
Anshan	0.0037	Gumbel	9.9375
Shenyang	0.0054	Gumbel	6.6269
Benxi	0.0032	Frank	26.2284
Fushun	0.003	Gumbel	6.7245
New bin	0.0059	Gumbel	5.0797
Jianchang	0.0079	Gumbel	6.2191
Suijin	0.0064	Gumbel	7.4027
Xingcheng	0.0056	Frank	37.025
Dais	0.0068	Gumbel	6.1298
Camp	0.0052	Gumbel	6.2535
Urine	0.0029	Gumbel	7.4962
Jade	0.0038	Gumbel	8.5682
Meadow	0.0096	Gumbel	5.1666
Danton	0.0052	Frank	30.6512
House shop	0.0049	Gumbel	9.2423
Long sea	0.0042	Gumbel	7.8626
Zhuang river	0.0089	Gumbel	5.2585
Dalian	0.0018	Gumbel	12.1758

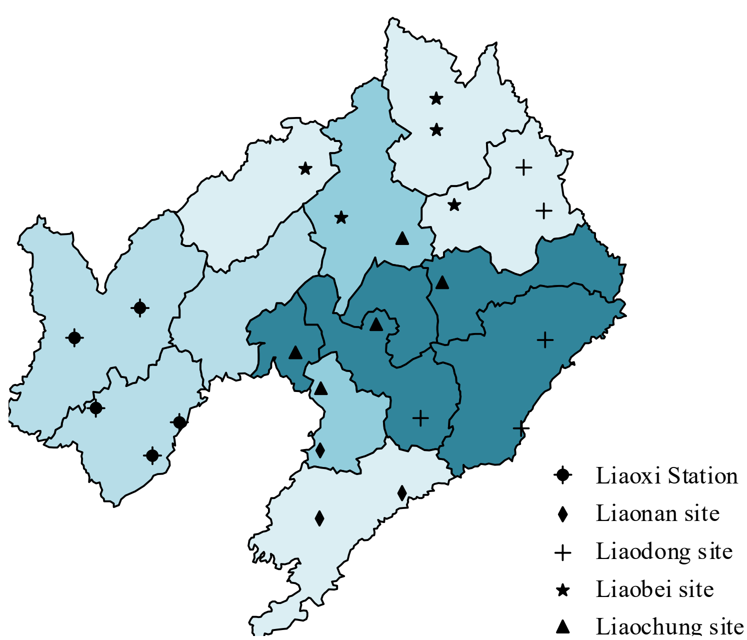
3.2. Normalized Integrated Drought Cycle Signature Analysis

Based on the entropy weight method-Topsis, the drought characteristic weight values of 1974-2024 were calculated as shown in Table 7. Based on SPI index, the study area is dominated by low and medium grade droughts, with low grade accounting for 62% and medium grade accounting for 38%. The identification results based on SPEI index are closer to the above results, and the study area is also mainly dominated by low and medium grade droughts, with 62% of low grade and 38% of medium grade.

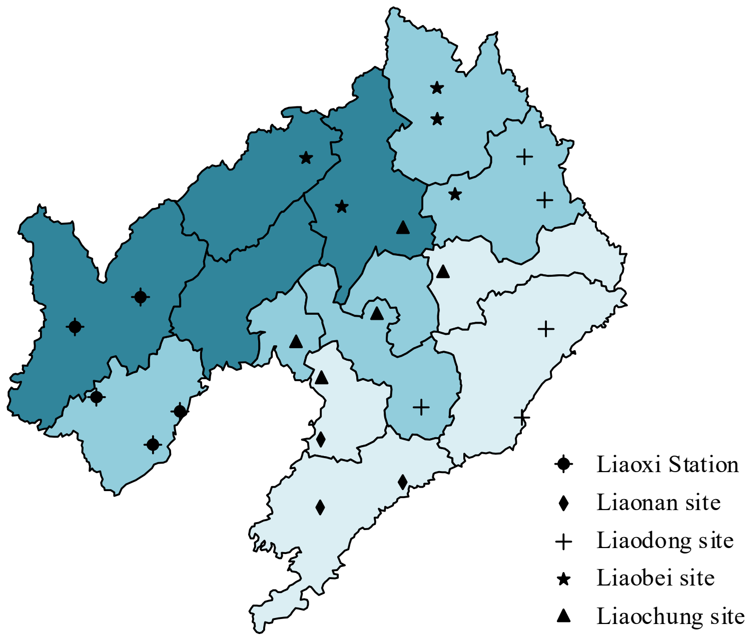
Table 7. Drought characteristics and their weights

Drought index	Dry characteristic weight				
	Duration	Frequency	Strength	Peak	Intensity
SPI	0.23	0.11	0.28	0.20	0.18
SPEI	0.17	0.15	0.30	0.22	0.16
EDDI	0.12	0.18	0.22	0.25	0.23
CIDI	0.21	0.14	0.16	0.23	0.26

3.2.1. Spatial variability of normalized typical meteorological cycles. Meteorological indicators and their spatial distribution affect the degree of drought class and spatial distribution [93,101]. The Pearson correlation coefficients of normalized composite drought characteristic indicators with precipitation and temperature are shown in Figure 1. According to the SPI index, the normalized composite drought characteristic index showed a significant negative correlation with precipitation (significance level < 0.01), with correlation coefficients ranging from -0.58 (light blue) to -0.43 (blue). According to the SPEI index, the normalized composite drought characteristic index showed a significant positive correlation with the temperature coefficient ($P<0.05$, passing the significance test except for a few areas), with correlation coefficients of 0.18 (blue) to 0.54 (light blue). Therefore, the results of SPEI are used as an example to explain the spatial distribution and evolution results of the drought center in the region using the spatial distribution and trend changes of precipitation and temperature. Precipitation is higher than that in the plains and western mountains, while the temperature in the plains is higher than that in the mountains; low precipitation and high temperature make the drought severe in the plains.



(a)Correlation analysis(SPI)

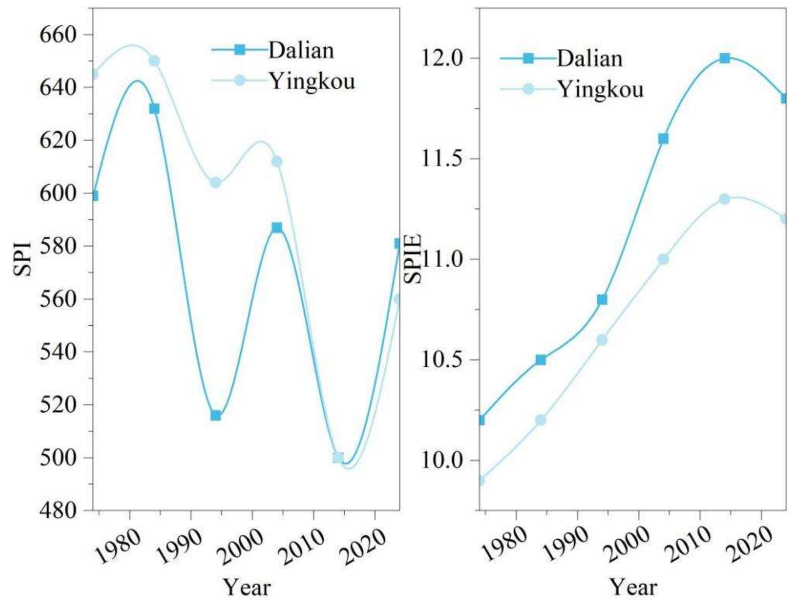


(b)Correlation analysis(SPEI)

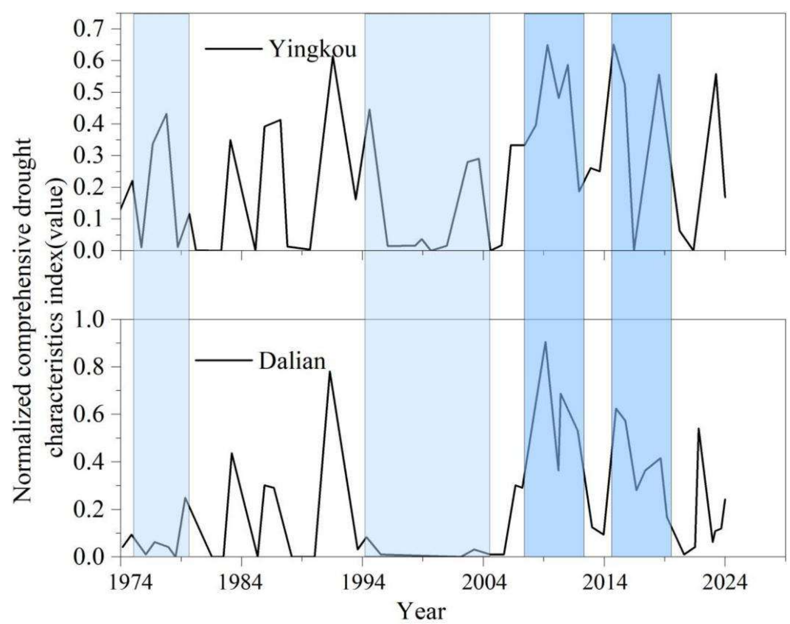
Fig. 1. The correlation between normalized drought characteristics indexes

3.2.2. Temporal variations in normalized typical meteorological cycles. To further verify the above inferences, Yingkou and Dalian were selected for specific analysis. The long-time precipitation and temperature series of different ages in the two areas, with linear trends and the corresponding normalized drought and flood composite characteristic index series of the two areas, are represented in the figure, and the results are shown in Fig. 2.

As can be seen from the figure, 1974s~2024s the average multi-year precipitation in Yingkou is 595.2 mm, which is higher than that in Dalian (569.2 mm), and the average multi-year air temperature in Dalian (10.7°C) is lower than that in Yingkou (11.5°C), and the normalized composite drought characteristic value of Yingkou is larger than that of Dalian in most of the years of the time period under the influence of the factors of precipitation and air temperature (the light blue area), i.e., drought conditions in Yingkou are stronger than those in Dalian. From 2004s to 2024s, the multi-year average precipitation in Dalian decreases less than that in Yingkou, and its drought continues to worsen with the decrease in precipitation, especially from 2004s to 2014, when the normalized composite drought eigenvalue in Dalian exceeds that in Yingkou (blue area), and the center of drought migrates from Yingkou to Dalian. Through the above analysis, it can be found that the uneven spatial distribution of precipitation and temperature and the variability of changes are one of the important reasons for the regional and migratory evolution of drought in the study area.



(a)The trend of precipitation and the change of temperature sequence



(b)The sequence of normalized and comprehensive drought characteristics

Fig. 2. DVI change trend of Liaoning Province from 2000 to 2016

4. Conclusion

Based on the meteorological data of Liaoning area from 1974 to 2024, four drought indices, SPI, SPEI, EDDI, and CJDI, are integrated, and Copula function and entropy weight TOPSIS are introduced to calculate the normalized drought composite index. The migration pattern and time pattern of meteorological drought centers in the historical period are explored. The study shows that:

1) Calculated by Kendall's rank correlation coefficient and Spearman's rank correlation coefficient, drought calendar time and drought intensity are strongly correlated.

2) According to the SPI index and SPEI index, the normalized composite drought characteristic index showed a significant negative correlation with precipitation ($P<0.01$) and a significant positive

correlation with the temperature coefficient (P<0.05). Therefore, low precipitation and high air temperature made the drought in Liaoning Plain Region serious.
3) Taking Dalian and Yingkou in the Liaoning region as an example, it further proves that the uneven spatial distribution of precipitation and temperature and the variability of changes are one of the main factors in the evolution of drought periodicity and migration in the study area.

Funding

Technology for Continuous Simulation and Quantitative Evaluation of the Adequacy of Supply-Regulation Capability in New Power Systems across Multiple Scenarios (KJZ2023117).

References

[1] Wu, Y., An, J., Gui, C., Xiao, C., & Yan, D. (2023, June). A global typical meteorological year (TMY) database on ERA5 dataset. In Building Simulation (Vol. 16, No. 6, pp. 1013-1026). Beijing: Tsinghua University Press.

[2] Kambezidis, H. D., Psiloglou, B. E., Kaskaoutis, D. G., Karagiannis, D., Petrinoli, K., Gavriil, A., & Kavadias, K. (2020). Generation of typical meteorological years for 33 locations in Greece: Adaptation to the needs of various applications. Theoretical and Applied Climatology, 141, 1313-1330.

[3] Putra, I. D. G. A., Nimiya, H., Sopaheluwakan, A., Kubota, T., Lee, H. S., Pradana, R. P., ... & Karnawati, D. (2024). Development of typical meteorological years based on quality control of datasets in Indonesia. Renewable Energy, 221, 119699.

[4] Li, H., Yang, Y., Lv, K., Liu, J., & Yang, L. (2020). Compare several methods of select typical meteorological year for building energy simulation in China. Energy, 209, 118465.

[5] Fan, X., Chen, B., Fu, C., & Li, L. (2020). Research on the influence of abrupt climate changes on the analysis of typical meteorological year in China. Energies, 13(24), 6531.

[6] Zhang, W., Li, H., Wang, M., Lv, W., Huang, J., & Yang, L. (2024). Enhancing typical Meteorological Year generation for diverse energy systems: A hybrid Sandia-machine learning approach. Renewable Energy, 225, 120369.

[7] Chan, A. L. S. (2016). Generation of typical meteorological years using genetic algorithm for different energy systems. Renewable Energy, 90, 1-13.

[8] Gai, S., Zhang, X., Xie, J., & Yin, K. (2022, July). A Method for Generating Typical Meteorological Year with Simplified Types of Meteorological Parameters. In International Conference on Building Energy and Environment (pp. 2999-3002). Singapore: Springer Nature Singapore.

[9] Realpe, A. M., Vernay, C., Pitaval, S., Lenoir, C., & Blanc, P. (2016). Benchmarking of five typical meteorological year datasets dedicated to concentrated-PV systems. Energy Procedia, 97, 108-115.

[10] Sun, J., Li, Z., Xiao, F., & Xiao, J. (2020). Generation of typical meteorological year for integrated climate based daylight modeling and building energy simulation. Renewable energy, 160, 721-729.

[11] Bigtashi, A., Papakyriakou, A., & Lee, B. (2024). Defining generation parameters with an adaptable data-driven approach to construct typical meteorological year weather files. Energy and Buildings, 303, 113781.

[12] Ohunakin, O. S., Adaramola, M. S., Oyewola, O. M., Fagbenle, R. O., Adelekan, D. S., Gill, J., & Abam, F. I. (2018). Photovoltaic performance prediction in Northern Nigeria using generated typical meteorological year dataset. African Journal of Science, Technology, Innovation and Development, 10(5), 579-591.

[13] Cebecauer, T., & Suri, M. (2015). Typical meteorological year data: SolarGIS approach. Energy Procedia, 69, 1958-1969.

[14] Bre, F., & Fachinotti, V. D. (2016). Generation of typical meteorological years for the Argentine Littoral Region. Energy and Buildings, 129, 432-444.

[15] Yu, Y., Chou, J., Yao, X., & Ma, N. (2024). Generation and application of typical meteorological year data for PV system potential assessment: A case study in China. Journal of Building Engineering, 86, 108831.

[16] Yuan, J., Huang, P., & Chai, J. (2022). Development of a calibrated typical meteorological year weather file in system design of zero-energy building for performance improvements. Energy, 259, 125031.

[17] Huld, T., Paietta, E., Zangheri, P., & Pinedo Pascua, I. (2018). Assembling typical meteorological year data sets for building energy performance using reanalysis and satellite-based data. Atmosphere, 9(2), 53.

[18] Ernst, M., & Gooday, J. (2019). Methodology for generating high time resolution typical meteorological year data for accurate photovoltaic energy yield modelling. Solar Energy, 189, 299-306.

[19] Huang, K. T. (2020). Identifying a suitable hourly solar diffuse fraction model to generate the typical meteorological year for building energy simulation application. Renewable Energy, 157, 1102-1115.

[20] Hosseini, M., Bigtashi, A., & Lee, B. (2020). A systematic approach in constructing typical meteorological year weather files using machine learning. Energy and Buildings, 226, 110375.

[21] Kulesza, K. (2017). Comparison of typical meteorological year and multi-year time series of solar conditions for Belsk, central Poland. Renewable Energy, 113, 1135-1140.

[22] Rudniak, J. (2020). Comparison of local solar radiation parameters with data from a typical meteorological year. Thermal Science and Engineering Progress, 16, 100465.

[23] Farah, S., Saman, W., & Boland, J. (2018). Development of robust meteorological year weather data. Renewable Energy, 118, 343-350.

[24] Salehi, S. S. S., Kalamees, T., Kurnitski, J., & Thalfeldt, M. (2024). New typical meteorological year generation method based on long-term building energy simulations. Building and Environment, 256, 111504.