Alinteri J. of Agr. Sci. (2021) 36(2): 22-37
e-ISSN: 2587-2249
info@alinteridergisi.com

http://dergipark.gov.tr/alinterizbd
http://www.alinteridergisi.com/
DOI:10.47059/alinteri/V36I2/AJAS21112

RESEARCH ARTICLE

Method for Measuring the Similarity of Multiple Metrological Sequences in the Key Phenological Phase of Rice-based on Dynamic Time

Z.A. Khan[1*] • Mousa Mohammed Khubrani[2] • Shadab Alam[3] • S.J. Hui[4] • Y. Wang[5]

[1*]*College of Information Science and Technology, Nanjing Agricultural University, Nanjing, China. E-mail: jmi.amu1@gmail.com*

[2]*College of Computer Science and IT, Jazan University, Jazan, KSA. E-mail: mmkhubrani@jazanu.edu.sa*

[3]*College of Computer Science and IT, Jazan University, Jazan, KSA. E-mail: snafis@jazanu.edu.sa*

[4]*School of Computer Science and Engineering, Nanjing University of Science & Technology, China.*

[5]*School of Software, East China Jiaotong University, Nanchang, China. E-mail: wangyann@189.cn*

ARTICLE INFO

Article History:
Received: 27.04.2021
Accepted: 05.06.2021
Available Online: 12.07.2021

Keywords:
Metrological Data
Rice Phenological Model
Dynamic Time Warping Algorithm
Metrological Indicators
Data Mining for Agriculture

ABSTRACT

The automatic classification of historical data of myriad diverse meteorological sequences in the annual period can help to find the climate differences through key phenology of rice. In this paper, a hybrid gradients-shape dynamic time warping (HGSDTW) algorithm is proposed to measure the similarity of meteorological data during the diverse rice growth period at various locations. The weighting calculation of Euclidean distance uses the form factor in the rice jointing and heading stage. The distance matrix constructs first & second-level gradient single-factor transformation sequences during the period. The dynamic programming method obtains the similarity distances of single and multiple meteorological factors. The results show that the classification accuracy rate from HGSDTW of the heading & jointing stage is higher than that of other similar algorithms. Furthermore, it can observe that the clustering number increases the classification accuracy, and the HGSDTW algorithm maintains the accuracy of 14% for varieties of rice at diverse locations to multiple years of jointing. Besides, the automatic classification experiment of sequence period shows that the classification accuracy of this method is higher than that of another similarity measure. The classification accuracy rate of the heading stage sequence is 10%~14% higher than that of a similar previous standard measurement algorithm, and the jointing period is 1%~9% higher. In this case, the cluster number increasing the classification accuracy, and the HGSDTW maintain the overall accuracy of 14%. Thus, this method can be effectively combined with the classification algorithm to improve the efficiency of the automatic classification of multi-weather sequence data in key phenological periods of rice.

Please cite this paper as follows:

Khan, Z.A., Khubrani, M.M., Alam, S., Hui, S.J. and Wang, Y. (2021). Method for Measuring the Similarity of Multiple Metrological Sequences in the Key Phenological Phase of Rice-based on Dynamic Time. *Alinteri Journal of Agriculture Sciences, 36*(2): 22-37. doi: 10.47059/alinteri/V36I2/AJAS21112

Introduction

The traditional method of similarity analysis of sequence meteorological data is to re-describe the characteristics of meteorological data using average temperature, standard deviation and equal meteorological statistical indicators [1] [2] [3] [4] [5].

But, the similarity compares the approximate degree of statistical indicators. This kind of method uses statistical indicators to reflect the characteristic changes in meteorological conditions. However, the distribution characteristics of meteorological data over a while ignoring the trend of the data itself, weakening the abnormal value of temperature change and the variation of the cumulative value of rainfall.

* Corresponding author: jmi.amu1@gmail.com

The scientist's optimism to define the agro-meteorological sequence data using time series calculates the similarity and degree of separation between inter-annual or different regions by distance metrics [6-12]. The similarity distance metric commonly used in agricultural meteorology is the Euclidean distance [10-13]. Xu Fan et al. [4] used the Euclidean distance to analyze the similarity between the floating cloud and the previous year's data of the contiguous areas. The statistical similarity frequency divided the regional climate types, and finally, the similarity of agricultural meteorological conditions estimates the regional environmental degree. Although this method is effective and straightforward, it only considers the difference of the corresponding time points. It does not take into account the fluctuation characteristics of meteorological data in different seasons. Bannayan et al. & Chen Shang et al. [10-11] used the Euclidean distance combined with KNN to analyze the similar years of meteorological data between years and screened similar earlier meteorological data to replace unknown meteorological data during the target growth period, but the similarity to meteorological data.

In the different phenological stages of crops, there are apparent differences in agro-meteorological elements such as temperature, light and rainfall [14-15]. The rice is affected by extremely high temperature and sudden changes in rainfall at the heading stage, which often causes an abnormality of the sequence shape fluctuation curve and the occurrence of peak-valley misalignment between the inter-annual sequences. Therefore, it brings difficulties to the similarity analysis of sequence data. The use of the dynamic time warping (DTW) metric algorithm provides a new way to solve such problems. The DTW is a method for solving a similar matching problem by solving the linear drift caused by the time series on the similar sample sequence data with uneven distortion or bending [16]. The application of this method in agricultural remote sensing data analysis is obvious. Guan et al. [17] used a DTW algorithm to analyze the similarity of NDVI time-series data throughout the year and divided the areas where rice grows. Mondal et al. [18] used DTW to analyze the similarity of NDVI in different farming modes such as double cropping and single cropping to determine the agricultural planting patterns in different regions. However, during the phenological period of rice growth, the inter-annual meteorological sequence data not only has the peak-valley misalignment phenomenon. It also has morphological change characteristics such as the direction of change, the magnitude of change and uses the DTW method to measure the sequence meteorological data. These similarities bring difficulties.

In this paper, for the problem of mismatch matching between "crest-peak" and "valley-trough" in the matching of rice temperatures, sunshine time series and rainfall sequence data in the phenological period is considered. Also, it comprehensively considers the direction and variation range of meteorological sequence in rice phenology. A newly proposed morphological and mixed gradient-based multivariate dynamic time warping algorithm, namely hybrid gradients-shape dynamic time warping (HGSDTW). The First and second-order gradient transformation represents the Multilevel meteorological sequences consisting of max & min temperature, rainfall and sunshine hours for every day. The Multilevel gradient is used to explore more morphological information and trend information within the meteorological sequence. Further, using the morphological factor to calculate the Euclidean distance weighting and construct the distance matrix for the single-factor transformation sequence of the original. Both level gradients are effectively preventing the occurrence of excessive phase difference at the matching point in the jointing & heading periods. Also, avoid the problem of peak-to-valley mismatching. Finally, dynamic programming obtains the similarity distances of single and multiple meteorological factors.

The rest of the paper is organized in different sections: section 2 describes the research significance and background. Section 3 presents the DTW metric multivariate time series similarity and its variants. In section 4, describe the morphological characteristics of meteorological time series data during rice phenology and proposed new hybrid gradients shape dynamic time wrapping algorithm framework. In section 5 explains the test and results analysis. The discussions are carried out in section 6. Finally, section 7 concludes the paper.

Research Significance and Background

The prediction of the daily maximum temperature can adequately grasp the daily temperature threshold. It assistances the peoples to comprehend whether the daily temperature outstrips the threshold range which rice growth period can bear and whether there is a high-temperature threat.

The ground observation method is also improving and making the number of observation stations exploding as well as the frequency is becoming more and denser [10]. The massive meteorological data generated by the development of these technologies not only has a large data scale but also the form of data will be varied. Contemporary, the method for meteorological forecasting sequence divides into two steps. Initially, constructed the prediction model and then the historical meteorological sequence is used to train the model through the error loss function. So, the predicted meteorological sequence is similar to the real weather sequence approximation. Consequently, meteorological sequence prediction can be divided into two aspects: meteorological sequence prediction and meteorological sequence similarity measurement.

Research Progress on Meteorological Sequence Prediction Methods

The accurate prediction of future meteorological conditions can provide an essential theoretical basis for the rational arrangement of rice cultivation and the prevention of extreme weather to cause natural disasters caused by rice growth.

From the step size of meteorological element prediction, it can divide into a short-term forecast, medium-term forecast and long-term forecast [11]. A short-term forecast is generally the future three intra-predictions, while more than ten days is the medium-term forecast, and more than one month is called a long-term forecast. As far as the current meteorological prediction techniques are concerned, they are mainly divided into four types [12]: forecasting basis based on weather graphics and atmospheric motion systems, statistically-based probability estimation methods, atmospheric model-based model predictions, and different Forecast method integration forecast. However, since the meteorological data itself has the time dimension feature, the use of meteorological time series for meteorological data prediction is more in line with the real scene. Therefore, many scholars have studied meteorological time series prediction. At present, the mainstream divides into two categories: meteorological data prediction based on traditional time series prediction methods and meteorological time series prediction based on machine learning.

Problem Formulation

In the prediction of the highest temperature series during the growth period of rice, the current method is usually only suitable for short-term prediction, and medium and long-term temperature prediction will be a significant error. However, for rice production guidance, its significance is not great. So, how to improve the time step and accuracy of prediction becomes the key to predicting the highest temperature forecast during the rice growth period. In the current prediction methods mostly uses a stepwise prediction. The prediction results only focus on the numerical value of the data but ignore the overall distribution and morphological characteristics of the prediction results, and error transmission is inevitable. So this prediction modelling method may not be the most suitable for predicting the highest temperature sequence during rice growth. The most important thing to construct this model is the loss function construction problem in the training process. In the prediction process, constructing the loss function that accords with the temperature distribution during the phenophase and emphasizes the extreme value is the key to predictive model training. Most of the research on such problems lies in the similarity measure between the predicted sequence and the real sequence. The primary issue is the meteorological sequence similarity measure based on the Euclidean distance among the point-by-point pairs. It cannot adapt to the time drift and peak-trough dislocation characteristics during the rice growth period. The time series similarity measure can be used to calculate the degree of approximation between meteorological data. However, since the growth period of rice is subject to changes in external meteorological conditions, the time dimension is advanced or delayed, and the number of days in the growth period will also be different. So, only measuring the Euclidean distance of an equal length sequence. It may not be the most appropriate choice. Although the dynamic time warping algorithm can effectively solve this problem, the morphological characteristics of the meteorological data during the growth period and the magnitude of the change often cause the peak and trough error matching in the matching process of the similarity measure. It also causes the dynamic time warping algorithm to have defects in measuring the similarity of meteorological data during the growth period.

Based on Dynamic Time Warping (DTW) Metric Multivariate Time Series Similarity

The popularity of the DTW algorithm stems from its high efficiency as a "time-series similarity measure." It reduces the impacts of time-shifting and distortion by allowing time series to be transformed "elastically" in order to identify identical forms with various phases [8]. The Dynamic Time Warping algorithm (DTW) determines the best path for warping two-time series. The algorithm estimates both the warping path values and the distance between the two series [9].

Dynamic Time Bending Distance

Dynamic time warping was initially applied to the field of speech recognition [19] and is often used to compare the similarity of two-time series. For the convenience of representation, a one-dimensional time series representation is being used here. Two time series S = (s1, s2, ..., sm) and Q = (q1, q2, ..., qn), si and qj represent the ith time sum in the sequences S and Q, respectively. The data at the jth time point, 1 ≤ i ≤ m, 1 ≤ j ≤ n. Construct a distance matrix D of n*m using the distance between any two points: 2, as follows:

$$D_{\mathrm{m*n}} = \begin{bmatrix} d(1,1), & d(2,1), & ...., & d(m,1) \\ d(1,2), & d(2,2), & ..., & d(m,2) \\ ... & ... & d(i,j) & ... \\ d(1,n), & d(2,n), & ..., & d(m,n) \end{bmatrix}$$

Where d (i, j) represents the distance between the time series data points for si and qj, i.e.

$$d(i,j) = (s_i - q_j)^2 \tag{1}$$

The basic idea of DTW calculation is to find the path with the smallest cumulative distance between two sequences S and Q from the distance matrix, which demonstrates as DTW (S, Q). The specific calculation equation is:

$$\mathrm{DTW(S,\ Q)} = \frac{\min}{\mathrm{K}}\left(\sum_{k=1}^{K} d(i,j)_k\right) \tag{2}$$

First, construct a cumulative distance matrix D (m, n) of m * n, each element defines in the matrix given below:

$$\mathrm{D}(i,j) = d(i,j) + \min\begin{cases} D(i-1,j) \\ D(i,j-1) \\ D(i-1,j-1) \end{cases} \tag{3}$$

Where i = 1,2....,m. j = 1,2, ...,n. The sum of the alignment distance d (i, j) of the current square with coordinates (i, j) can demonstrate as the D (i, j), and the minimum distance of the cumulative distance of three adjacent elements of the previous segment, and D (0,0) = 0. Then D (m, n) is the cumulative sum of the DTW distances between S and Q.

Multiple Time Series Similarity Measure

Multivariate time series, i.e. multivariate time series [18], defined as:

Where K is the number of variables, and N is the time dimension. For the time dimension

$$X_{K*N} = \begin{bmatrix} \chi_{11}, & \chi_{12}, & \cdots, & \chi_{1N} \\ \chi_{21}, & \chi_{22}, & \cdots, & \chi_{2N} \\ \cdots & \cdots & \chi_{ij} & \cdots \\ \chi_{K1}, & \chi_{K2}, & \cdots, & \chi_{KN} \end{bmatrix}$$

Xij ($1 \le i \le K, 1 \le j \le N$) represent the observed value of the ith variable at a jth time.

For the multivariate time series SK*M and QK*N, the DTW similarity the measure is performed and the multivariate time series divided into several one-dimensional time series, and then the time series of the same variable dimension is subjected to DTW metric distance to obtain each dimension set:

$$DTW(S_{K\times M}, Q_{K\times N}) = \{DTW(S_{k\times M}, Q_{k\times N}), k = 1,2,\cdots,K\}^T$$

Finally, the similarity distances of the multiple time series are combined by weighting, as follows:

$$DTW(S_{K\times M}, Q_{K\times N}) = \sum_{k=1}^{K} w_k * DTW(S_{k\times M}, Q_{k\times N}) \quad (4)$$

Variants of Dynamic Time Warping

Derivative Dynamic Time Warping (DDTW)

Instead of matching the original signals, this form of DTW matches the first derivatives of the sequences. This matching is more resilient to outliers in practice. In plain English, the Dynamic Time Warping (DTW) technique is beneficial for aligning two signals that are comparable except for local accelerations and decelerations along the time axis. When the two sequences are different on the Y-axis, such as different scaling (amplitude scaling) [13] or linear trends can be efficiently deleted, the method has difficulty.

In Derivative Dynamic Time Warping, the distance between two signals is not Euclidean, but rather the square of the difference between the calculated derivatives of the two signals. The DDTW algorithm formula is shown below:

DA (dai) = $\frac{(a_i - a_{i-1}) + ((a_{i+1} - a_{i+1})/2)}{2}$, 1< i< m

Where m is the length of sequence A, because the first and last estimates are not defined, it is considered that da1 = da2 and dam = dam-1, and ai is the first input signal. This technique has also been used in weighted variants of DTW (WDTW), known as Weighted Derivative Dynamic Time Warping (WDDTW).

Weighted Derivative Dynamic Time Warping

One prominent DTW version is the weighted derivative dynamic time warping method, which is a variation of DTW and its extension, derivative dynamic time warping (DDTW) [11]. The DTW may attempt to explain Y-axis variability by warping the X-axis; this may result in unanticipated singularities, which are alignments of a point in one series with several points in another series, as well as unintuitive alignments. DDTW turns the original points into higher-level features, which contain the shape information of a sequence, to address the drawbacks of DTW. [11] gives the approximated equation for changing data point ai in sequence A. The following is the weighted version of DDTW:

WDDTWp (DA, DB) = $\sqrt[p]{\xi^*(i,j)}$

Where ξ*(I, j) = |w|i-j| (dai – dbj) |p + min {ξ* (i-1, j-1), ξ* (i-1, j), ξ* (i-1, j-1)}, and DA and DB are the transformed sequences from sequence A and B, respectively.

To put it another way, the WDDTW with optimum weights offers a lot of potential for boosting time series classification and grouping accuracy. The WDDTW defines a weight to determine the distance between two sequences. On the other hand, the distance between two sequences is not calculated directly from the original value. Instead, the original sequences are subjected to neighbour point averaging, resulting in the creation of two new sequences. The slopes of a line running through the test point and its left neighbor, as well as the slopes of the left neighbor and the right neighbor, are averaged.

Weighted Hybrid Dynamic Time Warping (WHDTW)

The WHDTW is a sophisticated statistical-based tool for modelling generative sequences with an underlying approach that produces an obvious sequence [9]. The WHDTW has been used in a variety of signal processing applications, including speech processing. It has, however, been effectively used for low-level NLP (Natural Language Processing) applications such as phrase unitization, document extraction, and part-of-speech tagging. First, the methodology is used to train a model that should, in our example, represent a word utterance. After that, this model is used to test an utterance and determine the likelihood that the model generated the sequence of vectors.

Multivariate Meteorological Sequence Similarity Algorithm based on Dynamic Time Bending Distance Metric of Shape Coefficient and Mixed Gradient

Morphological Characteristics of Meteorological Time Series Data during Rice Phenology

The meteorological elements, including max & min temperature and rainfall & sunshine hours, are strictly related to rice's whole growth process. There are obvious influences on the meteorological elements in different ecological points and different phenological stages. Xiao Hui et al. [20], through the analysis of the trend of the main influencing factors and the volatility characteristics, the practical physical significance shows that sunshine hours in the tillering and jointing periods are the most critical and the key influencing factor in the heading and the flowering period is precipitation. Tan Jiang et al. [15] verified the seed setting rate of rice at different temperatures by the high-temperature stress test, and the quality of rice would be different. Therefore, this paper analyzes the morphological characteristics such as the highest temperature, the lowest temperature, the rainfall, the sunshine hours and other weather sequences affecting the jointing and heading stages of rice which lays a foundation for further algorithm design demonstrate in Figure 1.

The following four cases can be summarized from the trend phenomenon in Figure 1, as described in detail below:

Case 1 (a): The numerical gap is gradually reducing, but the direction of change is reversing. As shown in Figure 1 (a), the BC segment and B'C' overall trend are similar, But the AB and A'B' segments change in the opposite direction.

Case 1 (b): A peak or trough that is too large in phase difference. In Figure 1 (b), the DG segment and D'G' are generally similar, but in fact, the EG segment and D'F' segment does the best match.

Case 1(c): The trend of change is the same, but there are vertices with different amplitudes of local changes. Although the overall trend of the HJ segment and the H'J' segment in Figure 1(c) is decreasing, there is a turning point at the I point and I' point, which results in different speeds of subsequent changes.

Case 1(d): The peak-trough shift caused by the sequence distortion. Figure 1 (d) shows peaks of the two rainfall values of K and K', but the position is precisely symmetrically misaligned.

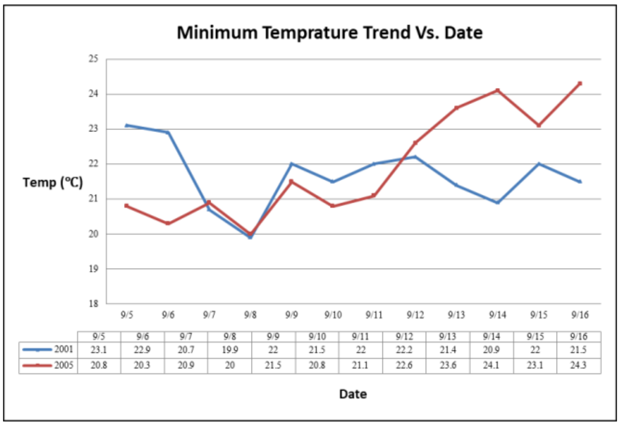


Fig. 1(a). The minimum temperature trend

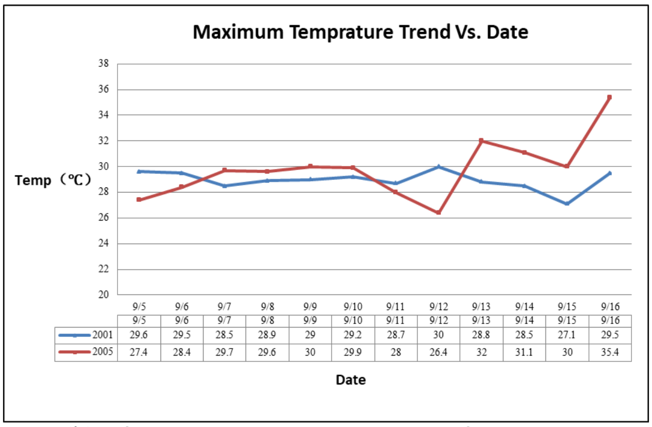


Fig. 1(b). The maximum temperature trend

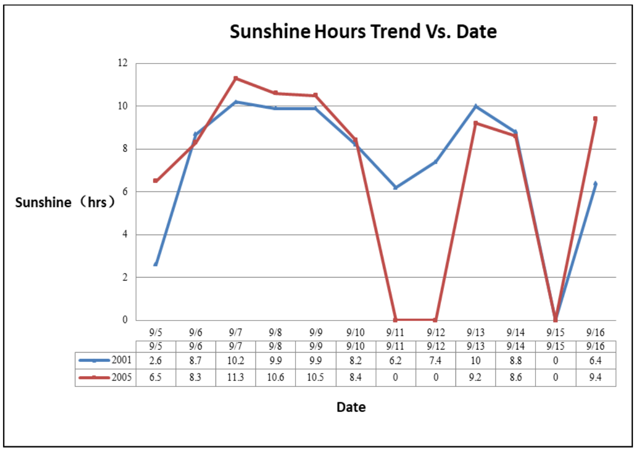


Fig.1(c). The sunshine hours trend

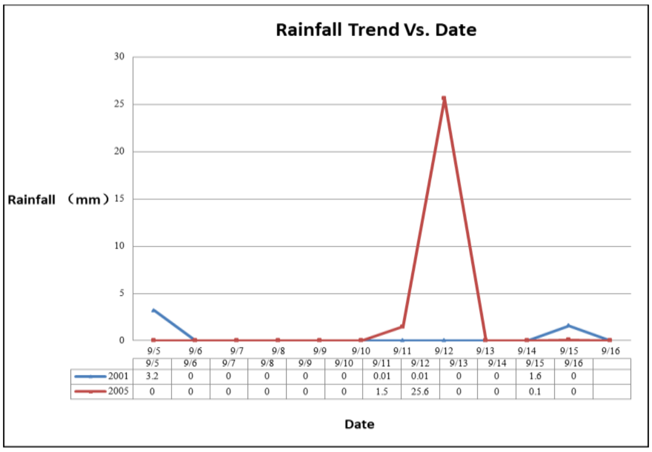


Fig.1(d). The rainfall hours trend

Figure 1. Checking the trends in minimum temperature 1 (a), maximum temperature 1 (b), sunshine hours 1 (c), and rainfall hours 1 (d) for the sequence of heading period of Yixing 'Wuyu' in 2001 and 2005.

Proposed Hybrid Gradients Shape Dynamic Time Wrapping (HGSDTW) Algorithm Framework

The meteorological sequence data has the problem of peak matching of the wave trough described in Case 1(d) during the heading period. The proposed dynamic time warping algorithm capable of solving the distortion drift phenomenon. The HGSDTW algorithm uses the dynamic programming method to optimize the shortest path of the distance matrix formed by the sequences S and Q to find the shortest distance.

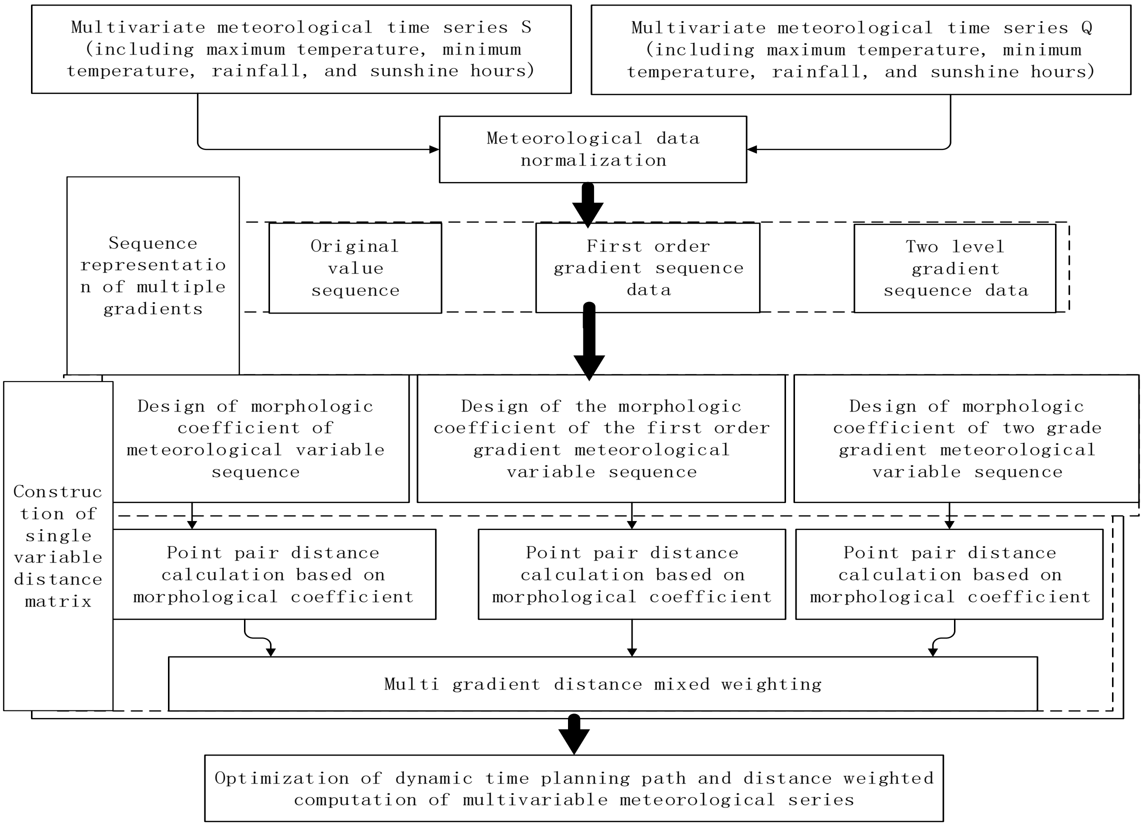


Figure 2. The Framework of the proposed hybrid gradients-shape-dynamic time warping algorithm

The similarity measure of the multivariate meteorological event sequence decomposes into a univariate meteorological sequence to construct the distance matrix D and then use the dynamic programming method to find the shortest path to obtain the distance. Finally, weigh the distance between every meteorological variable sequences to get the last two. Whether the distance in the distance matrix between univariate sequences accurately represent the distance between pairs of sequences becomes the focus of this work. According to the analysis of Cases 1(a) & 1(c), the morphological factors, including the direction of change and magnitude of the changing area a vital feature of meteorological data and consider these characteristics in similarity measure; the sequence is re-described by introducing a first and a second-order gradient that can represent the direction of change and fast and slow magnitude.

Moreover, in the case of case 1(b), the distance between the pair of points needs to take into account the phase difference between the two points. Therefore, as the phase difference increases, the proportion of the distance will become smaller and smaller. Subsequently, the key to this study is how to mix the pairs of points in a multi-gradient sequence? The process of a proposed hybrid gradient shape dynamic time warping algorithm is shown in figure 2, and Table 1 represents the pseudo-code of the proposed HGSDTW algorithm.

Sequence Representation of Normalization and Multilevel Gradients

The meteorological data has a front-to-back dependence on the time series, and the direction & magnitude of the change can describe its morphological characteristics. This study performs Z-score standardization for each meteorological factor:

$$z_i = \frac{x_i - \mu_i}{\sigma_i} \tag{5}$$

Where z_i represents the normalized meteorological variable; xi, μi and σi represent the original value, mean and variance of the meteorological variable in the entire data space, respectively. The first-order gradient uses to represent the sequence of the change direction and, based on it after that, calculate the gradient and explore the information sequence of change speed. The derivative gradient calculation method used as the slope in the calculation process [21].

This calculation method is beneficial to smooth the abnormal fluctuations of the sequence generated by the outliers. The specific calculation is as shown in equation (6).

$$q'_i = \frac{q(i,j) - q(i,j-1) + (q(i,j+1) - q(i,j-1))/2}{2} \tag{6}$$

Table 1. The Pseudo of HGSDTW

HGSDTW Algorithm
1: Input sequence q , r
2: Using formula（3-2）Calculate the first derivative of each sequence separately q' 、 r'、Second Derivative q'' 、 r''
3:Dist = zeros(q Number of rows in the sequence, r Number of rows in the sequence)
4:for i=1 To q The end of the sequence：
5: for j=1 To r The end of the sequence：
6: total = zeros(N,1);total_1 = 8:zeros(N,1);total_2 =zeros(N,1);%N for q Number of columns in the sequence
7: for k=1 To the end of the sequence：%K The number of columns representing the sequence also refers to the attribute dimension of the sequence.
8: Using formula（2-2）Calculate the original sequence separately q,r Attribute value distance total, First derivative distancetotal_1、Second order
9： Number distance_total_2
10: end
11: Using formula（3-5）Calculating the distance matrix_Dist
12: end
13:end
14: Initialize cumulative distance matrix D=Inf (size(Dist))
15: Using formula（2-3）Calculation_D
16: Use the dynamic programming method to find the slave D(1,1) To D(q Number of rows in the sequence, r sequence 21：Number of rows) Shortest path, and
17：The number of points that need to pass to record the shortest path_t
18: Final sequence q,r the distance between _Distval =D (q Number of rows in the sequence, r Number of rows in the sequence) / t
20: Return output _Distval

Single Variable Mixed Distance Weighting and Matrix Construction

Sequence Shape Coefficient Design

The irregular fluctuations in weather sequences due to abnormal values such as extreme temperatures and accumulated rainfall. In order to avoid the calculation error caused by the matching peak or trough matching in the sequence matching metric, a constant form factor "g" is introduced to control the weight function wi-j, which can effectively control the maximum interval of matching and the difference of sequence shape. There are enormous phase difference, smaller weight and distance. The equation for the weight function wi-j is:

$$w_{i-j} = \frac{1}{1+\exp(-g(i-c))} \tag{7}$$

Where i-j represents the position of the corresponding points of the two sequences i and j, and c represents the distance of i from its sequence centre position [m/2].

Morph-Based Point-to-Distance Calculation and Multi-Gradient Distance Hybrid Weighting

The key to the dynamic time warping algorithm lies in the construction of the distance matrix. The algorithm uses the original value obtained in section 2.2.1 and the distance between the corresponding pairs of points in the two-level gradient to calculate d(i,j) and fill it [22]. The distance matrix equation (8) is given below.

$$d（i,j）= w_0 * d_0(i,j) + w_1 * d_1(i,j) + w_2 * d_2(i,j) \tag{8}$$

Combined with the above description, the equation (5) and (6) of the superposition equations are used to derive the final weighted-based mixed gradient calculation distance metric equation:

$$d(i,j) = w_{i-j}(w_0 * d_0(i,j) + w_1 * d_1(i,j) + w_2 * d_2(i,j))$$
$$= \frac{1}{1+\exp(-g(i-c))}（w_0 * d_0(i,j) + w_1 * d_1(i,j) + w_2 * d_2(i,j)）\tag{9}$$

According to Nandyala et al. [23], w0 =1, w1= w2 = 2 is the optimal design of mixed gradient weights and the study by Jeong et al. [24] found that the shape coefficient would have a good effect when going to 0.1~0.6. According to a combined experiment of Ye Yanqing [25], the parameter set separately of w_0, w_1, w_3 and g, and the classification accuracy of different parameter combinations related to 1NN. The w_0, w_1, w_3 are randomly changed between 0 and 5 in steps of 1, and g is the constant coefficient of the logical weight function. According to this study, when g is between 0.01~ 0.6 obtains good results. So, the values are respectively taken rendering to 0.01, 0.1, 0.2, 0.3, 0.4, 0.5, and 0.6, and in the meantime tested the combinations of w0, w1, and w3. The subsequent experiments are carried out and always maintain the accuracy of the values at the forefront. The combination of w0, w1, w3 and g is 1, 2, 2, 0.2, where g has a large difference between 0.1 and 0.3 and 0.2. So this study determines the values of parameters like w_0 =1, w_1 = w_2 = 2, g = 0.2.

Dynamic Planning Path Optimization and Multivariate Meteorological Sequence Distance Calculation

The distance matrix of each meteorological variable can be obtained by Section 3.2.2.2. Equation (2) calculates the dynamic time-bending distance between the sequences of each meteorological variable, and equation (4) calculates the distance between multivariate meteorological sequences. The weight between each weather variable is set to 1.

Tests and Result Analysis

The meteorological time series data of rice jointing and heading stage were applied to verify the effectiveness of the proposed HGSDTW algorithm. The algorithm program is created using the MatLab R2014b.

Test Data Acquisition

The meteorological data acquired from the National Meteorological Data Sharing Center and the phenological period information contained in the pilot experiment. Table 2 shows the meteorological data years as per the locations for the jointing stage.

Table 2. Basic information of meteorological data

Location	Start and End Time	Phenological Phase
Jiangsu Yixing	1994—2011	Jointing stage
Anhui Hefei	1999—2011	Jointing stage
Guangdong Gaoyao	2004—2009	Jointing stage
Jilin Tonghua	1998—2010	Jointing stage

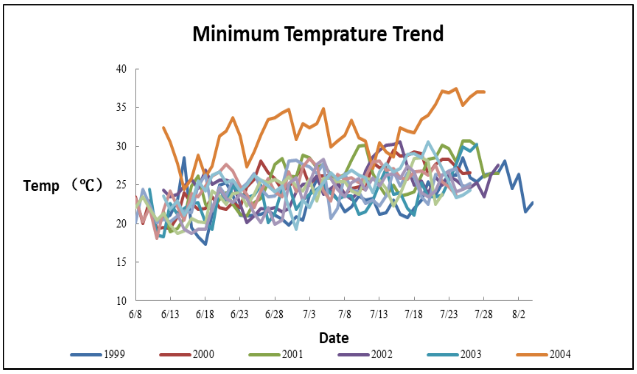


Figure 3 (a). The trend of the minimum temperature

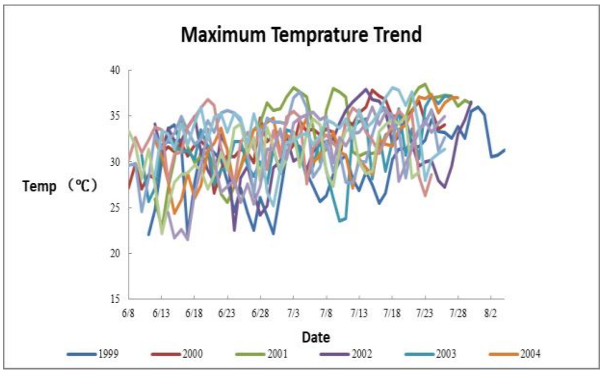


Figure 3 (b). The trend of the maximum temperature

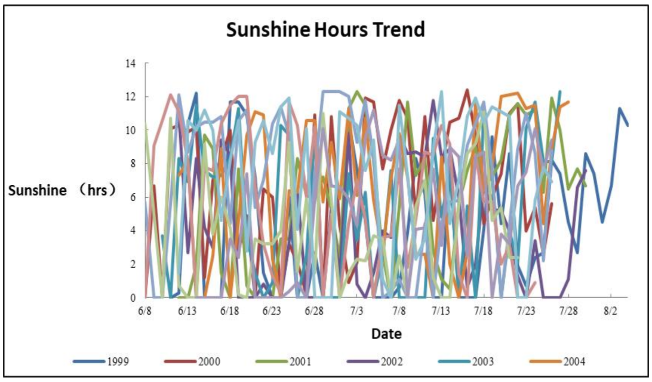


Figure 3 (c). The Trend of Sunshine Hours

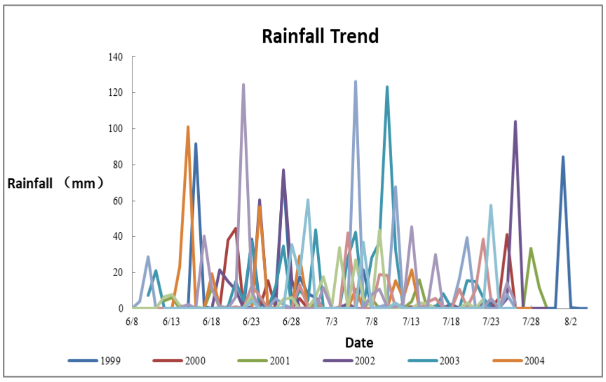


Figure 3 (d). The Trend of Rainfall

Figure 3. Meteorological data of jointing period from 1999 to 2004 in Hefei area (4(a): maximum temperature; 4 (b): minimum temperature, 4 (c): rainfall, 4(d): sunshine hours).

Figure 3 demonstrated the trends for all meteorological indicators like maximum and minimum temperatures, rainfall and sunshine hours at the jointing stage. The trend demonstrates low temperature in 2003 and high in 2001 as compared to other years. The sunshine is not constant as per hours varying every day, and the rainfall trend shows more rain at Hefei Area in 2002 and 2003. The actual collection of data for the Hefei area from 1999 ~ 2011 to all meteorological indicators. However, figure 3 shows only 1999~2004 data in pictorial representation for the Hefei area to better understate the trends.

Table 3 demonstrated the meteorological data as per the dates of months for the heading and jointing stage to all locations. It represents the starting and ending date for the collection of metrological data. Every location has the same pattern to the collection of data based on starting and ending date.

Table 3. Actual date range corresponding to rice growth periods

Jiangsu Yixing				Anhui Hefei		Guangdong Gaoyao		Jilin Tonghua	
Start of HS	End of HS	Start of JS	End of JS	Start of JS	End of JS	Start of JS	End of JS	Start of JS	End of JS
1994/9/2	1994/9/15	1994/6/25	1994/7/26	1999/6/11	1999/8/4	2004/5/27	2004/6/13	1998/5/20	1998/7/18
1995/9/4	1995/9/14	1995/6/28	1995/8/2	2000/6/8	2000/7/26	2005/6/4	2005/6/13	1999/5/25	1999/7/18
1996/9/3	1996/9/16	1996/6/25	1996/8/4	2001/6/12	2001/7/30	2006/6/6	2006/6/15	2000/5/28	2000/7/18
1997/9/3	1997/9/16	1997/6/23	1997/8/3	2002/6/12	2002/7/30	2007/6/6	2007/6/16	2001/5/28	2001/7/18
1998/9/5	1998/9/15	1998/6/26	1998/8/4	2003/6/10	2003/7/27	2008/6/6	2008/6/16	2002/5/28	2002/7/18
1999/9/4	1999/9/18	1999/7/5	1999/8/8	2004/6/12	2004/7/28	2009/6/4	2009/6/14	2003/5/26	2003/7/18
2000/9/1	2000/9/16	2000/7/1	2000/8/9	2005/6/8	2005/7/26			2004/5/26	2004/7/18
2001/9/5	2001/9/16	2001/7/2	2001/8/7	2006/6/8	2006/7/24			2005/5/29	2005/7/20
2002/9/5	2002/9/17	2002/6/29	2002/8/7	2007/6/8	2007/7/22			2006/5/28	2006/7/20
2003/9/5	2003/9/16	2003/6/30	2003/8/6	2008/6/14	2008/7/26			2007/5/28	2007/7/20
2004/9/4	2004/9/16	2004/7/1	2004/8/6	2009/6/12	2009/7/26			2008/5/28	2008/7/20
2005/9/5	2005/9/16	2005/6/26	2005/8/5	2010/6/8	2010/7/24			2009/5/28	2009/7/20
2009/9/5	2009/9/14	2009/6/27	2009/8/2	2011/6/12	2011/7/26			2010/6/3	2010/7/22
2010/9/4	2010/9/16	2010/7/1	2010/8/8						
2011/9/4	2011/9/16	2011/7/1	2011/8/8						

Whereas JS = Jointing Stage & HS = Heading Stage

Data Labeling and Evaluation

Data Annotation Processing

To measure the algorithm's accuracy for the similarity measurement of meteorological data in rice phenology, it is necessary to obtain meteorological data during the phenological period marked with the category label. The data division and labelling are shown in figure 3. First, extract the corresponding phenological period data, and after that, obtains the tagged earlier phenological meteorological data. Further, the method is processed according to the method in FIG for this purpose. Since the meteorological sequence itself contains multiple meteorological factors such as max & min temperature and rainfall & sunshine hours. These types of labelling of meteorological data will have a problem of class division granularity and the differences between the categories acceptable for the experimental test scenario. The degree will have a significant impact on the classification. In order to make the experimental method of this study conforms to the similarity measure under different demand granularities.

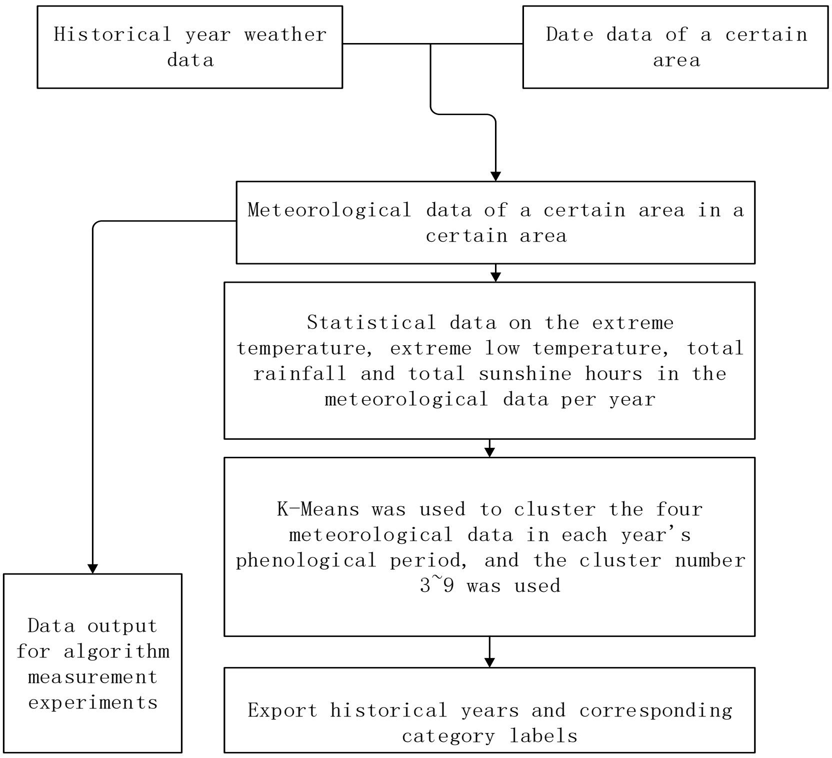


Figure 3. Data Division and Labeling

Table 4 shows that the first row in the table indicates the years, and the first column indicates the number of different clusters. Besides, the other rows represent the meteorological data category under the cluster number, and the same number indicates the same category. The cluster 3~9 in the data category labelling using to perform the K-means clustering on the statistical data of the original meteorological data [27]. It covers numerous granularity necessities as much possible as. Due to the extreme high-end, extreme low temperature, cumulative rainfall value and the cumulative value of sunshine hours, rice growth is more visible [15] [21]. For this reason, the indicators used in this study are the maximum & minimum temperature, rainfall and sunshine hours.

Table 4. Results of hand marking under different clustering data of heading dates over the years in Yixing

Number of Clusters	Cluster Data Years														
	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2009	2010	2011
3	0	2	1	0	1	1	0	0	2	1	0	1	0	2	1
4	3	0	1	3	1	1	3	3	2	1	3	1	0	2	1
5	4	0	1	4	1	3	4	4	2	1	0	3	0	2	3
6	0	4	3	0	3	1	0	3	2	3	4	1	4	5	1
7	1	4	3	1	3	2	1	1	5	3	1	6	4	0	2
8	1	7	3	1	3	6	1	1	4	3	2	0	2	5	6
9	6	4	8	6	8	1	6	3	7	8	0	5	3	2	1

Test Methods and Evaluation Indicators

To verify the accuracy of the algorithm similarity measure, the experiment combines the 1NN algorithm to match the similarity of the target a year with other years, find the nearest one year and then judge whether the categories are consistent and count the jointing and heading periods in the same area. The exact number of matches within the experiment and the evaluation of an index used in the experiment is the accuracy rate. The calculation formula is given below:

$$\text{Score} = \frac{\textit{Number of Right Discriminate}}{\text{Total Number}}$$

Comparative Experiment Algorithm

The proposed HGSDTW algorithm and the Weighted Hybrid dynamic time warping (WHDTW) [23] used in this paper only consider the phase difference. The weight-based dynamic time of the mismatch matching problem is assessed only by considering the phase difference. The weighted dynamic time warping (WDDTW) is a derivative-based dynamic time warping algorithm that only considers the direction of change of the first-order gradient. The DDTW and DTW based on the original value [21] [29] [30] were tested according to the method in Section 4.2.2, respectively and to calculate the classification accuracy.

Experimental Test Design

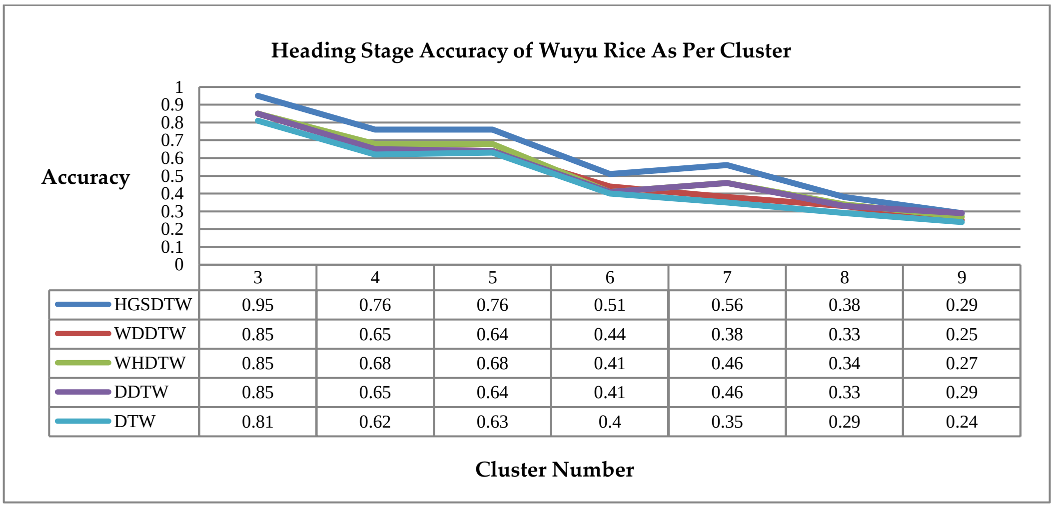
This paper mainly proposes a similarity measure algorithm for meteorological data in rice phenology during complicated time and space scenarios. The two groups designed to verify the ecumenicity of the algorithm and spatial distribution in the time dimension. It has described in detail below:

- The ecumenicity of different similarity measure algorithms in the time dimension. The experiments verify the accuracy of the time dimension by comparing the algorithm's accuracy with other algorithms in different phenological periods under the same species in the same region. For example, this group of experiments used the phenological dates of jointing and heading of the rice variety 'Wuyu'. The corresponding meteorological data will be mapped to verify the algorithm measurement of this study. Also, it can apply to different phenological meteorological conditions.
- Comparing the different similarity measure algorithms in spatial distribution and the experiments compares the measurement algorithm's accuracy under the same phenological period with various varieties in different regions. The experimental design of this group based on many climatic characteristics of different ecological zones and under the premise of ensuring the same phenological period. The meteorological data of Jiangsu Yixing, Guangdong Di Gao, and Anhui Hefei, Jilin Tonghua, are shown in Table 1. The internal meteorological data fragments of the heading period during rice growth were extracted and compared the accuracy of the different algorithms. As a result, it has verified that the proposed HGSDTW algorithm can apply to different varieties in different locations to better outcomes.

Experimental Test Results and Analysis

Analysis of Experimental Results under Different Phenological Periods of the Same Variety in the Same Area

Figure 4(a) demonstrates that the HGSDTW algorithm is 10% ahead of the similar algorithm in the different clustering numbers of 'Yixing Wuyu' at the heading stage, and its highest accuracy rate HGSDTW algorithm is 95%. The standard DTW algorithm has a 14% advantage. HGSDTW algorithm enhances the adaptability of sudden changes in the time dimension compared to the comparison series's dynamic time warping similarity measure algorithm. Besides, the accuracy of various algorithms decreases when several clusters increase because the data is distributed on various clusters and executing data simultaneously. The algorithms perform in the same pattern of setup experiment. In other words, the reason for this situation is mainly because the number of clusters does not appear in the same category as the number of clusters increases. When the number of categories is 9, the HGSDTW algorithm low accuracy is also 29%, which is 5% higher than the other dynamic time warping algorithm based on the actual value. It also shows that the proposed HGSDTW algorithm can effectively distinguish between metrics when there is a slight difference in inter-annual meteorological data.

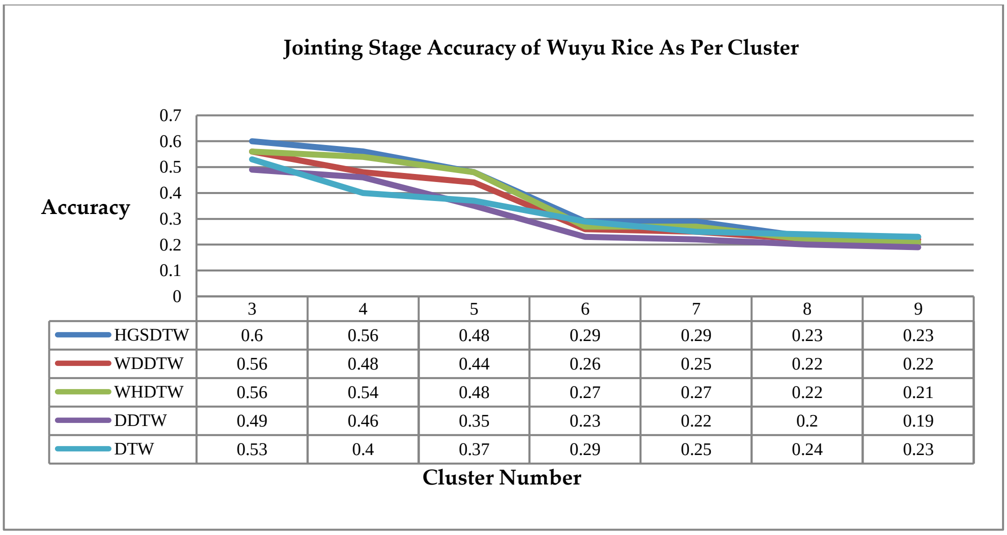


	3	4	5	6	7	8	9
HGSDTW	0.95	0.76	0.76	0.51	0.56	0.38	0.29
WDDTW	0.85	0.65	0.64	0.44	0.38	0.33	0.25
WHDTW	0.85	0.68	0.68	0.41	0.46	0.34	0.27
DDTW	0.85	0.65	0.64	0.41	0.46	0.33	0.29
DTW	0.81	0.62	0.63	0.4	0.35	0.29	0.24

Figure 4 (a). Classification Accuracy of Rice Heading Stage in Wuyu Japonica rice, Yixing Area

In Figure 4 (b), the HGSDTW is almost equivalent to the WHDTW algorithm. Combined with the trend of each meteorological element sequence in Figure 2, it can find that this is mainly because the jointing period belongs to the rainy season. The sequence fluctuates significantly in each dimension, and the peak-trough itself is not much different. So, the metric algorithm increases the morphological coefficient does not have a significant advantage. However, when the number of clusters is 4, it is 16% higher than the DTW algorithm. It also reflects the performance of the HGSDTW algorithm and is always kept in the first echelon sequence in the complex meteorological sequence similarity measure.

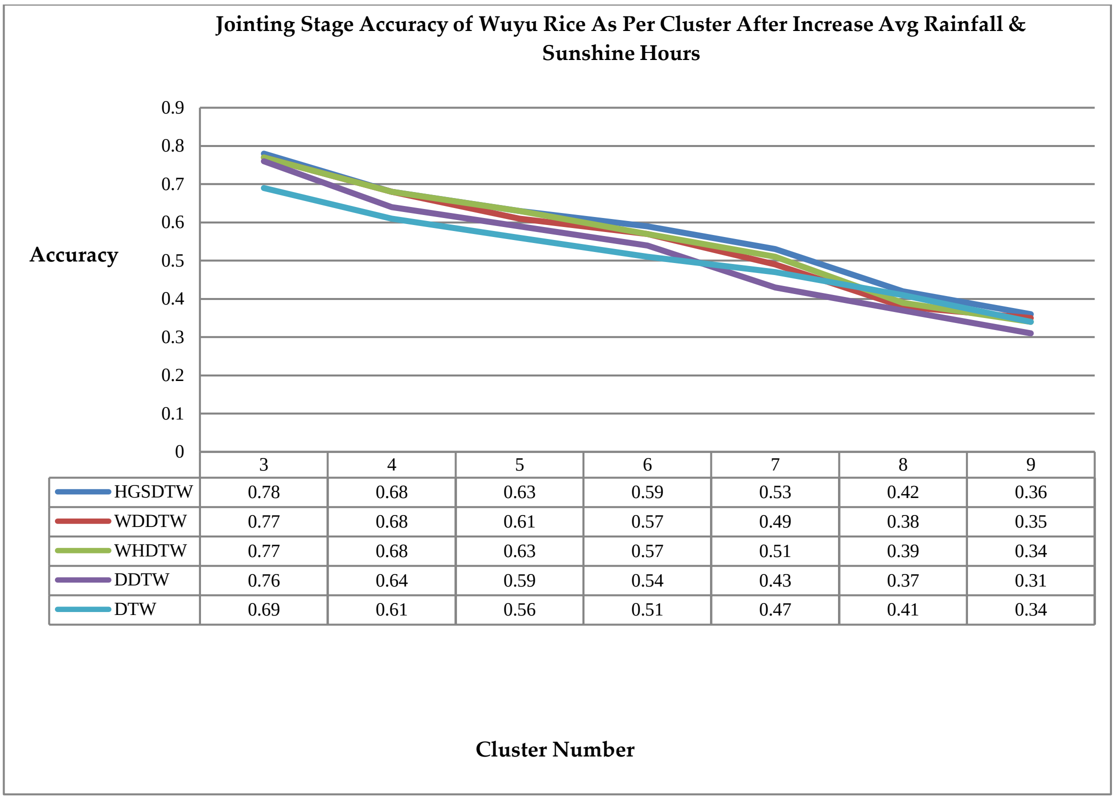
In the Yixing jointing experiment, the accuracy of all algorithms was significantly lower than that measured during the heading period. However, the reason for the analysis data was not significantly different during the jointing period, and the cluster centres were relatively concentrated. Therefore, the need to increase the clustering indicators may improve the accuracy of the classification. In order to prove this hypothesis, increase the average rainfall and the average sunshine hours and two indicators for data annotation to generate new samples for experiments.



	3	4	5	6	7	8	9
HGSDTW	0.6	0.56	0.48	0.29	0.29	0.23	0.23
WDDTW	0.56	0.48	0.44	0.26	0.25	0.22	0.22
WHDTW	0.56	0.54	0.48	0.27	0.27	0.22	0.21
DDTW	0.49	0.46	0.35	0.23	0.22	0.2	0.19
DTW	0.53	0.4	0.37	0.29	0.25	0.24	0.23

Figure 4 (b). Classification Accuracy of Rice Breeding Stage in Wuyu, Yixing Area

Figure 4. Classification accuracy of rice 'Wuyu' heading stage 5(a) and jointing stage 5(b) in the same area of Yixing, Jiangsu province experimental results under different phenological periods of the same variety



	3	4	5	6	7	8	9
HGSDTW	0.78	0.68	0.63	0.59	0.53	0.42	0.36
WDDTW	0.77	0.68	0.61	0.57	0.49	0.38	0.35
WHDTW	0.77	0.68	0.63	0.57	0.51	0.39	0.34
DDTW	0.76	0.64	0.59	0.54	0.43	0.37	0.31
DTW	0.69	0.61	0.56	0.51	0.47	0.41	0.34

Figure 5. Classification accuracy of rice 'Wuyu' rice at the jointing stage in the Yixing area after increasing average rainfall and average sunshine hours

With the help of this above experiment, it can provide analysis that the accuracy of various classification algorithms has significantly improved, and the HGSDTW algorithm still maintains superiority. Furthermore, the proposed HGSDTW algorithm has distinct advantages in accuracy compared with other dynamic time warping algorithms under different phenological periods in the same place, proving the accuracy of an algorithm in time attributes.

Results of Actual Test under Different Phenological Periods of Different Varieties in Different Regions

There is no significant difference in the accuracy of various algorithms from the experimental analysis of this group in the Hefei location. Applying the DTW series algorithm to the similarity measure of meteorological data in the rice jointing period is useful. However, in the WDDTW algorithm, the accuracy rate is still about 2% higher than other algorithms. It maintains this advantage under different cluster numbers, which indicate the algorithm's stability. Although experiments were carried out in Jilin Tonghua and Guangdong Gaoyao, the WDDTW algorithm had a higher accuracy than the HGSDTW algorithm under partial clustering; despite its apparent volatility, stability was not high. This phenomenon is mainly because that Jilin and Guangdong belong to different climate types. However, it's all belonging to the early summer and rainy weather. The temperature is relatively stable; the trend and magnitude are not significant. The information represented by the increased multilevel gradient does not play a role.

The WHDTW algorithm uses the cluster numbers of 3 to 5 in the Yixing to experiment. The cluster number 3 and 7 in Hefei and the cluster number are 3 to 5 in the Gaosong area Guangdong. The case of 7 is equivalent to the accuracy of the HGSDTW algorithm; the number of clusters to the experiment in Hefei is 8, and the number of clusters in the Jilin regional trial is 5 & 8 and in the high test area in Guangdong When the class number is 9, the accuracy of the WHDTW algorithm is 1% higher than that of the HGSDTW algorithm. However, there is a 4% reduction in backwardness in most experiments which means that the HGSDTW algorithm has higher stability than the previous standard algorithms. It has verified the ecumenicity of the proposed algorithm in spatial distribution and analysis from this group of experiments that the proposed algorithm can use to whole of China.

Fig.6(a). Classification Accuracy of 898 Rice As Per Cluster Number at Jointing period in Hefei Area.

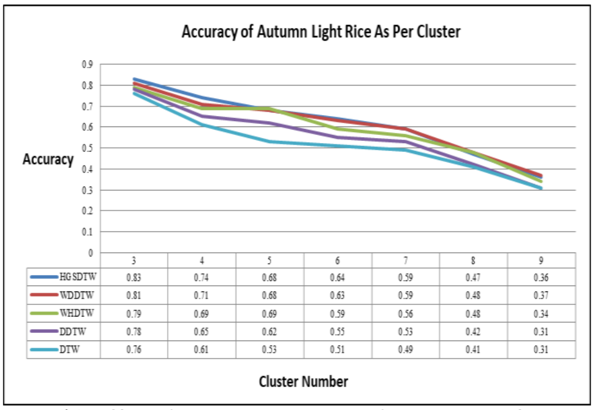


Fig.6(b). Classification Accuracy of Autumn Light Rice As Per Cluster Number at Jointing period in Tonghua Area.

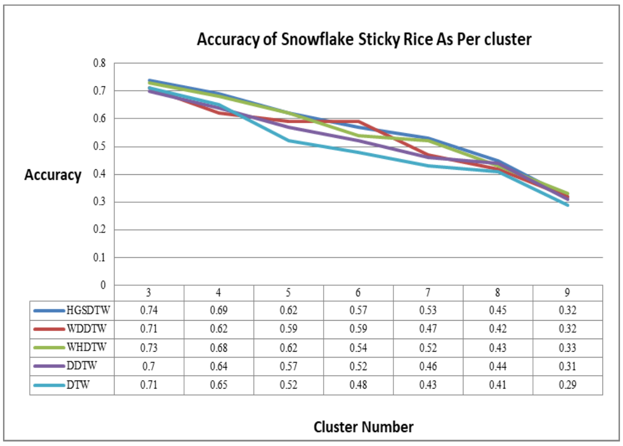


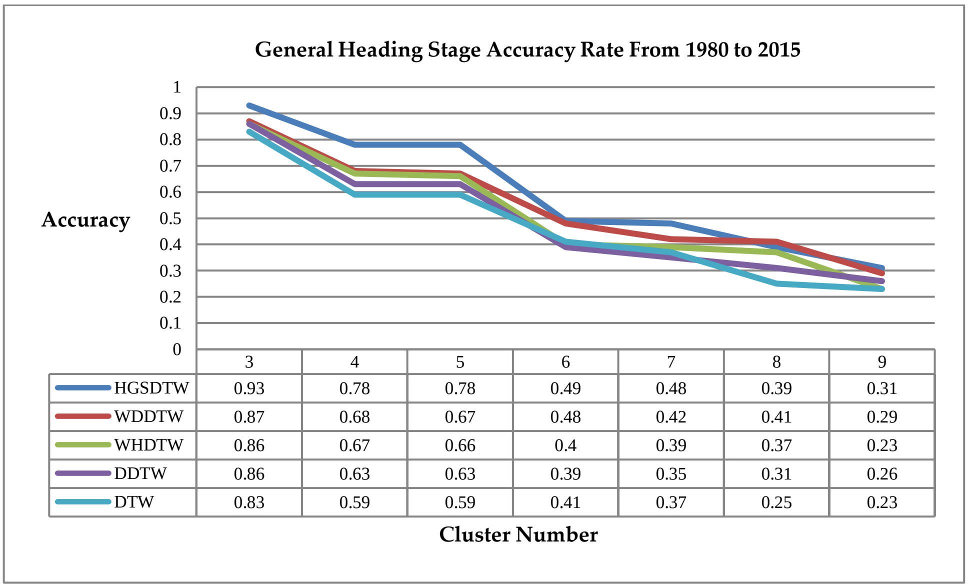
Fig6(c). Classification Accuracy of Snowflake Sticky Rice As Per Cluster Number at Jointing Stage in Gaoyao Area

Figure 6. Experimental results under the different phenological period for diverse varieties in different regions

Discussions

In this study, the mismatch matching of "peak-to-peak" and "valley-trough" in matching rice temperature, sunshine hours and rainfall sequence data during the growth period is considered. The direction and change of meteorological sequence during the rice growth period are considered comprehensively. It is based on the morphological characteristics of amplitude. A dynamic time warping algorithm based on morphology and mixed gradient (Hybrid Gradients-Shape Dynamic Time Warping, HGSDTW) is proposed to measure the similarity of meteorological data during the rice growth period. First, the meteorological sequence with the highest temperature, the lowest temperature, the rainfall, and the sunshine hours is normalized. After the primary meteorological sequence is expressed using the first-order gradient and the second-order gradient, the Euclidean distances using the additional form factors are different. The first-level gradient meteorological sequence and the second-gradient sequence data point pair of various meteorological factors in the same growth period are calculated. The weighting and constructed distance matrix of different meteorological factors obtain the distance obtained by a dynamic programming method. The distance between various meteorological factors in the year, and finally, the distance between the meteorological sequences in the same growth period in different years according to the overlap of the various meteorological factors. In this paper, according to cluster number 3~9, K-means clustering is applied to the meteorological sequence to assign category labels to the highest temperature maximum value, minimum temperature minimum value, and cumulative rainfall and accumulated sunshine hours. The multi-year meteorological sequence classification of rice at the jointing stage and heading stage showed that when the number of clusters was 3, the classification accuracy of the proposed method at the heading stage and jointing stage reached 95% and 78%, respectively. The classification accuracy rate of the heading stage sequence is 10%~14% higher than that of the similar measurement algorithm, and the jointing period is 1%~9% higher. In the case that the clustering number increases the classification accuracy, the algorithm can maintain the accuracy of 14%; for the Yixing Wuyu, Hefei 898, Guangdong Gaoyao, Jilin Tonghua Qiuguang varieties, automatic classification experiment of many years of jointing stage It has shown that when the number of clusters is 3, the classification accuracy of this method is 78%, 80%, 83%, and 74%, respectively. Compared with the similarity measure, the classification accuracy is 2%~9% higher. When the number of clusters increases the classification accuracy, the algorithm can maintain the accuracy of 1%~8%.

The data used for the experiments from before 2011 in section 4.3. In order to prove that the algorithm is also applicable to the similarity measure of rice phenological meteorological data in recent years, this hypothesis supports a set of experiments. The experimental data obtained from the National Meteorological Data Sharing Network and the Ministry of Agriculture. The meteorological data contains the Max & Min temperature, rainfall & sunshine hours from 1st Jan 1980 to 31st Dec 2015 in Nanjing. The date of the phenological period is a date range of expected heading date in the Nanjing area, such as Nanxun 44, Ningjing No. 3, Wuyunyu No. 24, Yangshuo 4038, Zhen Dao 11, Yanyou No. 8, and Changnongyu No. 5. The above experiment method obtains the following results:



	3	4	5	6	7	8	9
HGSDTW	0.93	0.78	0.78	0.49	0.48	0.39	0.31
WDDTW	0.87	0.68	0.67	0.48	0.42	0.41	0.29
WHDTW	0.86	0.67	0.66	0.4	0.39	0.37	0.23
DDTW	0.86	0.63	0.63	0.39	0.35	0.31	0.26
DTW	0.83	0.59	0.59	0.41	0.37	0.25	0.23

Figure 7. The Most Similar Detection Accuracy Rate in the General Heading Period from 1980 to 2015 in Nanjing

From figure 7, it can observe that the dynamic time-bending metric algorithm of morphology and the mixed gradient is the most similar. The highest accuracy reaches 93%, and the maximum is 19% better than other algorithms. The worst case is only 2% lower than the WDTTW algorithm. When compared with the same year detection experiment of rice heading stage in the Yixing area in section 4.3, this experiment amplifies the year of the data, including the more meteorological data changes. Notably, it also added the recent year's meteorological data. In this study, the generality of the proposed HGSDTW algorithm is extreme.

The meteorological data have typical spatial and temporal dual attributes, and the climatic conditions of specific regions have distinct characteristics [29]. There are also some differences in different years [30]. Based on the above two characteristics, the agro-climatic zone can be divided [31] and not only can the similarity retrieval of the inter-annual meteorological data of the same region be carried out. However, also, the climate region can be divided according to the meteorological similarity analysis. The Euclidean distance is used to calculate similarity, but this method ignores the time-scale characteristics of the meteorological data itself. There is also a significant fluctuation in the meteorological data itself, which causes the drift of the sequence on the time scale. Therefore, when performing meteorological data clustering or classification, only the combination of the trend on the time scale and the morphological distortion caused by the outliers can effectively perform the matching metrics of the multivariate meteorological time-series data. This study attempts to introduce a multivariate time series for meteorological data description and the DTW used for the process of similarity measurement. The mixed weighted representation of the original value and used the first and second level gradient. So that the original sequence and the meteorological are considered respectively in the sequence of the measurement. The information on the direction of change of data and the magnitude of the change of the meteorological data make the characteristics of the morphological change more critical when performing similarity matching. Simultaneously, increasing the morphological factor allows different meteorological data to match a specific phase difference [32] accurately. However, there is a strong correlation between the various variables of agricultural meteorological data. How to increase the coupling of the dimensions of each variable makes the measurement method more characteristic in the agricultural field has become a problem to be studied.

Moreover, the 1NN method used in the experiments in this paper performs the most similar year search. Although this method is a general method for distance measurement evaluation, there are many other uncertain factors in the agro-meteorological data. For example, there may be multiple similar years. At present, it is not a good explanation, which is also worth considering in the follow-up study.

Conclusion

This paper proposed a multivariate meteorological sequence similarity measurement method based on dynamic time warping, namely HGSDTW. It is considering the direction and range of variation of multiple sequences in rice phenology. The first-level gradient and the second-level gradient express the multilevel meteorological sequence consisting of the daily maximum & minimum temperature and rainfall & sunshine hours. The morphological factors are used to weight the Euclidean distance, respectively, for the jointing period and the initial period of the heading period. The single-factor gradient sequence of the first-level gradient and the second-level gradient constructs the distance matrix. The similarity distance between the single meteorological factor and the dynamic programming method obtains the multi-meteorological factor. The 1NN algorithm is used to classify and calculate the classification accuracy automatically. The multi-year meteorological sequence classification experiment of Yixing Wuyu rice at jointing & heading stage and the automatic classification experiment of Yixing Wuyu, Hefei 898, Guangdong Gaoyao, Jilin Tonghua Qiuguang varieties, and many years of jointing stage sequence.

The following conclusions obtained from the experiment outcomes and description given below:

- The proposed (HGSDTW) method in this study can effectively solve the similarity measurement problem of meteorological sequences in multiple growth periods in the same place. The experiment verifies its applicability in time. The experiment verifies its relevance in time. The multi-year meteorological sequence classification of Yixing "Wuyu" rice at the jointing stage and heading stage showed that when the number of clusters was 3, the classification accuracy by the proposed method in the heading and jointing stages reached 95% and 78%, respectively. The classification accuracy rate of the heading stage sequence is 10%~14% higher than that of a similar previous standard measurement algorithm, and the jointing period is 1%~9% higher. In this case, the cluster number increasing the classification accuracy, and the HGSDTW proposed algorithm can maintain the overall accuracy of 14%.
- The HGSDTW method uses the same growth period in multiple locations and proves its universality in space. The HGSDTW proposed an algorithm for multi-year meteorological sequence classification and multiple locations such as Yixing Wuyu, Hefei 898, Guangdong Gaoyao, Jilin Tonghua Qiuguang varieties with the help of the same setup experiment as same as a single location. The automatic classification experiment for many years of jointing stage shows that when the number of clusters is 3, the classification accuracy of this proposed method reaches 78%, 80%, 83% and 74%, respectively. The similarity measure method is 2%~9% higher in classification accuracy. The proposed algorithm can maintain the accuracy of 1%~8% when the clustering number increases the classification accuracy. The best advantage of this algorithm for multiple locations.

Future work

The design of weights of different variables in the process of similarity measurement of multivariate meteorological sequences. There are strong correlations among the various variables of agricultural meteorological data. During the different growth stages of rice, different meteorological factors play different roles. For example, in the tillering stage, the number of sunshine hours is the most critical. On the other hand, the critical factor affecting the flowering stage is precipitation. How to increase the coupling of the dimensions of each variable and make the measurement method more characteristic in the agricultural field has become a problem to be studied.

Acknowledgement

This research is supported by grants from the National Natural Science Foundation of China (No.31872847).

References

Yu, H., Lu, W. T., & Cao, S. N. (2015). Effects of changes in Temperature and Precipitation Resources on Rice & Wheat Production in Jiangsu Province in Recent 20 Years. *Resources and Environment in the Yangtze Basin.*, 24 (11), 1876 –1883.

Liu, X., He, X. M., & You, S. C. (2017). Characteristics of Main Agricultural Meteorological Disasters in China from 1976 to 2015. *Chinese Journal of Agro meteorology.*, 38 (8), 481-487.

Sitaram, D., Dalwani, A., & Narang, A. A. (2015). Measure of Similarity of Time Series Containing Missing Data using the Mahalanobis Distance. *2nd International Conference on Advances in Computing and Communication Engineering.,* 622-62.

Xu, F. (2017). A Study on the Environmental Meteorological Similar it between Yunfu and Its Surrounding Area. *Guangdong Chemical Industry.*, 44 (11), 209-211.

Yao, J. M., Feng, L. P., & Huang, W. L. (2012). Development of Common Agricultural Climate Index Calculation Software. *Chinese Journal of Agrometeorology.,* 33 (4), 595-602.

Krzyszczak, J., Baranowski, P., & Zubik, M. (2017). Temporal Scale Influence on Multiracial Properties of Agro-Meteorological Time Series. *Agricultural & Forest Meteorology.,* 239, 223-235.

Wang, X. E., Decker, & W. L. (1989). The Use of Distance Coefficient in the Research on Agroclimatological Resemblance. *Journal of Nanjing Institute of Meteorology.,* 12 (2), 187-199.

Senin, P. (2008). Dynamic Time Warping Algorithm Review. *University of Hawaii at Manoa Honolulu,* USA.

FURTUNĂ T, F. (2008). Dynamic Programming Algorithms in Speech Recognition. *Revista Informatica Economică.*

Bannayan, M., & Hoogenboom, G. (2008). Weather Analogue. A Tool for Real-Time Prediction of Daily Weather Data Realizations Based on a Modified K-Nearest Neighbor Approach. Netherlands: *Elsevier Science Publishers B. V.,* 703-713.

Chen, S., Dou, Z., & Jiang, T. (2017). DSS AT-CERES-Maize Prediction of Maize Yield Based on Historical Similar Meteorological Data Collected by Cluster Method. *Journal of Agricultural Engineering.,* 33(19), 147-155.

Elżbieta, R., & Katarzyna, R. (2015). Multi-Trait Analysis of Agroclimate Variations during the Growing Season in East-Central Poland (1971-2005). *International Agrophysics.*, 29 (2), 213-219.

Gao, S., Wang, H., & Sang, X (2016). Time Series Analysis of Rainfall and Runoff in Yuanjiang-Red River Basin. *System Engineering.*, (3), 153-158.

Liu, J., Liu, Y., & Wang, L. (2017). Simulating the Effect of Different Nursery Period and Greenhouse Temperature on Rice Growth in Hetao Irrigation Area. *Chinese Journal of Agrometeorology.*, 38 (8), 496-506.

Tan, J., Li, Y., & Pan, X. (2013). Effects of High-Temperature Weather on the Flowering, Fruiting and Quality of Rice. *Journal of Applied and Environmental Biology.*, 19 (6), 935-940.

Weng, Y., & Zhu. Z. (2004). Research on Time Series Data Clustering Algorithm Based on Dynamic Time Warping. *Computer Simulation.*, 21(3), 37-40.

Guan, X., Huang, C., & Liu, G. (2016). Mapping Rice Cropping Systems in Vietnam Using an NDVI-Based Time-Series Similarity Measurement Based on DTW Distance. *Remote Sensing.*, 8 (1), 19.

Mondal, S., & Jeganathan, C. (2018). Mountain Agriculture Extraction from Time-Series MODIS NDVI using Dynamic Time Warping Technique. *International Journal of Remote Sensing.*, 39 (11), 3679-3704.

Brown, J. C., & Miller, P. J. O. (2017). Automatic Classification of Killer Whale Vocalizations Using Dynamic Time Warping. *Journal of the Acoustical Society of America.*, 122 (2), 1201-1207.

Xiao, H., Shen, J., & Liu, J. (2011). An Analysis of Meteorological Impact Factors of Rice Yield in Lianyungang Region. *Journal of the Meteorological Sciences.*, 31: 81-86.

Luczak, M. (2016). Hierarchical Clustering of Time Series Data with Parametric Derivative Dynamic Time Warping. Oxford, UK: Pergamon Press, Inc.

GãRecki, T. Å.,¤& Uczak, M. (2014). First and Second Derivatives in Time Series Classification Using DTW. *Communications in Statistics - Simulation and Computation.*, 43 (9), 2081-2092.

Nandyala, S. P., & Kishore, K. T. (2014). Hybrid HMM/DTW Based Speech Recognition with Kernel Adaptive Filtering Method. *International Journal on Computational Sciences & Applications.*, 4(1), 11-21.

Jeong, Y. S., Jeong, M. K., & Omitaomu, O. A. (2011). Weighted dynamic time warping for time series classification. *Pattern Recognition.*, 44(9).

Ye, Y., Yang, K., & JIANG, J. (2017). Multivariate Time Series Similarity Matching Method Based on Weighted Dynamic Time Bending. *Pattern Recognition & Artificial Intelligence.*, 30 (4), 314-327.

Belgiu, M., & Csillik, O. (2017). Sentinel-2 Cropland Mapping using Pixel-Based & Object-Based Time-Weighted Dynamic Time Warping Analysis. *Remote Sensing of Environment.*, 204, 509-523.

Liu, R., Yang, T., & Chen, P. (2016). Loss Criteria for Loss of Continuous Rain in Flowering Period of Rape in Anhui Province. *Chinese Journal of Agrometeorology.*, 37(4), 471-478.

Vavrek, J., Viszlay, P., & Lojka, M. (2018). Weighted Fast Sequential DTW for Multilingual Audio Query-by-Example retrieval. *Journal of Intelligent Information Systems.*, (2), 1-17.

Li, H., & Liang, Ye. (2016). A dynamic Time Warping Method for Segmented Aggregate Approximation and Numerical Derivatives. *Journal of Intelligent Systems.*, 11(2), 249-256.

Orozco, A. M., CastroC, P. A., & Bicego, M. (2015). The DTW-Based Representation Space for Seismic Pattern Classification. *Computers & Geosciences.*, 85, 86-95.

Wang, X. (1992). Discussion on the Methods of Regionalization and Similar Types of Eco-climatic Suitability of Maize in China. *Chinese Journal of Agrometeorology.*, 13(5), 25- 27.

Zhang, Y., Zhang, X., & Guo, T. (2015). Effects of Temperature and Precipitation Change on Climate Productivity in Northwestern Liaoning. Chinese Journal of Agrometeorology., 36(2), 203-21.

Wang, L., Chen, H., & Li, Qi. (2010). Progress in Research on Agricultural Climate Zoning Methods. *Chinese Journal of Agrometeorology.*, 31(2), 277-281.

Zhao, J., & Itti, L. (2017). Shape DTW: shape Dynamic Time Warping [J*]. Pattern Recognition.*, 74, 171-184.