

Article

# Characterizing the Spatio-Temporal Pattern of Land Surface Temperature through Time Series Clustering: Based on the Latent Pattern and Morphology

Huimin Liu <sup>1,2</sup>, Qingming Zhan <sup>1,2,\*</sup> , Chen Yang <sup>1,2</sup>  and Jiong Wang <sup>3</sup>

<sup>1</sup> School of Urban Design, Wuhan University, 8 Donghu South Road, Wuhan 430072, China; hmliu@whu.edu.cn (H.L.); yangchen\_cug@foxmail.com (C.Y.)

<sup>2</sup> Collaborative Innovation Center of Geospatial Technology, 129 Luoyu Road, Wuhan 430079, China

<sup>3</sup> Faculty of Geo-Information Science and Earth Observation, University of Twente, 7500 Enschede, The Netherlands; j.wang-4@utwente.nl

\* Correspondence: qmzhan@whu.edu.cn

Received: 5 March 2018; Accepted: 20 April 2018; Published: 23 April 2018



**Abstract:** Land Surface Temperature (LST) is a critical component to understand the impact of urbanization on the urban thermal environment. Previous studies were inclined to apply only one snapshot to analyze the pattern and dynamics of LST without considering the non-stationarity in the temporal domain, or focus on the diurnal, seasonal, and annual pattern analysis of LST which has limited support for the understanding of how LST varies with the advancing of urbanization. This paper presents a workflow to extract the spatio-temporal pattern of LST through time series clustering by focusing on the LST of Wuhan, China, from 2002 to 2017 with a 3-year time interval with 8-day MODerate-resolution Imaging Spectroradiometer (MODIS) satellite image products. The Latent pattern of LST (LLST) generated by non-parametric Multi-Task Gaussian Process Modeling (MTGP) and the Multi-Scale Shape Index (MSSI) which characterizes the morphology of LLST are coupled for pattern recognition. Specifically, spatio-temporal patterns are discovered after the extraction of spatial patterns conducted by the incorporation of *k*-means and the Back-Propagation neural networks (BP-Net). The spatial patterns of the 6 years form a basic understanding about the corresponding temporal variances. For spatio-temporal pattern recognition, LLSTs and MSSIs of the 6 years are regarded as geo-referenced time series. Multiple algorithms including traditional *k*-means with Euclidean Distance (ED), shape-based *k*-means with the constrained Dynamic Time Warping (*c*DTW) distance measure, and the Dynamic Time Warping Barycenter Averaging (DBA) centroid computation method (*k*-*c*DBA) and *k*-shape are applied. Ten external indexes are employed to evaluate the performance of the three algorithms and reveal *k*-*c*DBA as the optimal time series clustering algorithm for our study. The study area is divided into 17 geographical time series clusters which respectively illustrate heterogeneous temporal dynamics of LST patterns. The homogeneous geographical clusters correspond to the zoning custom of urban planning and design, and thus, may efficiently bridge the urban and environmental systems in terms of research scope and scale. The proposed workflow can be utilized for other cities and potentially used for comparison among different cities.

**Keywords:** LST; spatio-temporal; pattern recognition; time series clustering; latent pattern; morphology

## 1. Introduction

Land Surface Temperature (LST) derived from satellite remotely sensed thermal infrared (TIR) imagery is a key indicator in understanding the impact of urbanization on the urban thermal

environment [1–3]. To some extent, the pattern recognition of LST is more important than that of Urban Heat Island (UHI) [1], as conventional UHI study is significantly restricted by the “urban-rural” dichotomy, where cities are treated holistically and temperature patterns within cities are inevitably ignored [4]. Focusing on the LST brings insightful understanding of thermal patterns exclusively at the phenomenon level without being restricted by the dichotomy and can be extended to intra-city thermal pattern analysis. Previous studies on LST were inclined to apply only one snapshot to analyze the spatial pattern of LST or its relationship with surface indicators [5–8] without considering the temporal non-stationarity. However, any LST patterns or dynamics drawing from a static point view can be misleading as “there is nothing spatial that is not temporal” [9]. Atmospheric and hydrological variances can result in considerable difference in temporally adjacent remotely sensed LST images even though the land use and land cover (LULC) remains constant. As a result, applying different images at a particular area during a certain period may generate conclusions that vary from each other. Recent studies covering the time-varying behavior of LST are well documented, including the diurnal [10–13], seasonal [12,14], annual [2,15], and inter-annual or across different years’ variations [16–18]. Whereas the studies on diurnal, seasonal, and annual pattern analysis of LST concentrate on how LST varies along with the earth’s rotation and revolution [2,11,19]. The knowledge derived from such studies has limited support for the understanding of how LST varies along with urbanization. In contrast, analyzing the temporal pattern of LST across different years can generate valuable knowledge about the impact of urbanization on the urban thermal environment. However, multiple challenges hinder the advancing of related research. The key problem lies in that it is defective to directly compare images acquired by satellite sensors of different atmospheric and hydrological conditions [3,6,20]. Thus, it is essential to normalize the data and make comparison under certain standards. Conventional normalization methods are limited as LST depends on LULC types nonlinearly [20]. This, to some extent, explains the popularity of UHI study where the direct comparison of LST data acquired at different time points can be avoided.

As a result, it is urgent to seek for a temporal analysis approach to cover the time-varying thermal characteristics while settling the data comparison dilemma. Time series clustering could be a potential alternative as it can locate the geographical regions with homogeneous climate patterns with the temporal observation bias due to sensor, atmospheric, and hydrological conditions taken into consideration, hence, effectively eliminate the improper comparison among satellite data acquired at different conditions. Homogeneous LST patterns, even subject to be under or overestimated, should be pinpointed by one particular trend line along the temporal dimension and distinguishable from the others. Furthermore, the spatial pattern of the clusters which possess homogeneous temporal LST patterns can also be revealed as the LST derived from satellite sensors is geo-referenced (time evolving values observed at fixed geographical locations) [21], thus, realizing the spatial and temporal patterns recognition simultaneously. In fact, time series clustering has been extensively applied for exploratory analysis or as a pre-processing step for a data mining task in climate and other domains like biology and finance [22–24]. It explores the data structure at a higher level of abstraction without the dependence on human supervision [21,25,26]. Specifically, it identifies clusters with homogeneous temporal patterns through the comparison of similarity among the different time series [23]. The research of time series clustering applied to daily temperature pattern recognition is well documented [27–29]. All studies utilized the daily temperature data acquired from sparsely distributed in situ stations. The application to time series LST data acquired from satellite sensors to explore the spatio-temporal patterns of the whole area is limited. It has been stressed that the interpretation and visualization of the time series clustering results is difficult, especially when the input data is high-dimensional [30]. Hence, spatial clustering can be operated before the time series clustering to provide a basic understanding of the spatial patterns and the corresponding temporal variances.

However, the parameters to characterize the spatio-temporal patterns of LST are limited [1,31,32], although sufficient parameters are available for UHI study such as magnitude or intensity, spatial extent, position, and orientation. Since research in geography has suggested that continuous and

smooth surface supports pattern recognition more effectively [33,34], the parameterized surface models [35–37] and non-parametric kernel models [31,38,39] were utilized to quantify thermal characteristics and generate parameters from a smoother surface. Among them, Wang et al. [31] introduced the latent pattern of LST (LLST) and a morphological parameter named the Multi-scale Shape Index (MSSI) to the urban climate community to enrich the parameter system of LST pattern mining. LLST is a continuous and smooth surface generated from the original data which is regarded as hidden continuous patterns in noises. It is derived by the non-parametric Multi-Task Gaussian Process (MTGP) model by sharing information across different images. MSSI is a morphology parameter generated from LLST to characterize the shape of LLST at optimal scales and illustrate the LLST correlation with near spaces. By coupling LLST with MSSI in time series clustering, regions with homogeneous variability of LLST patterns but distinct variability of the correlations with the surroundings, or vice versa, can be identified into different clusters. Besides, the operation of MTGP can also fill the missing pixels by referring to the target image itself and also the temporally adjacent images, hence, solving the frequent cloud-cover [2,11,40] problem which has hampered the temporal analysis of LST. As a matter of fact, cloudy conditions actually account for more than half of the day-to-day weather around the globe [41].

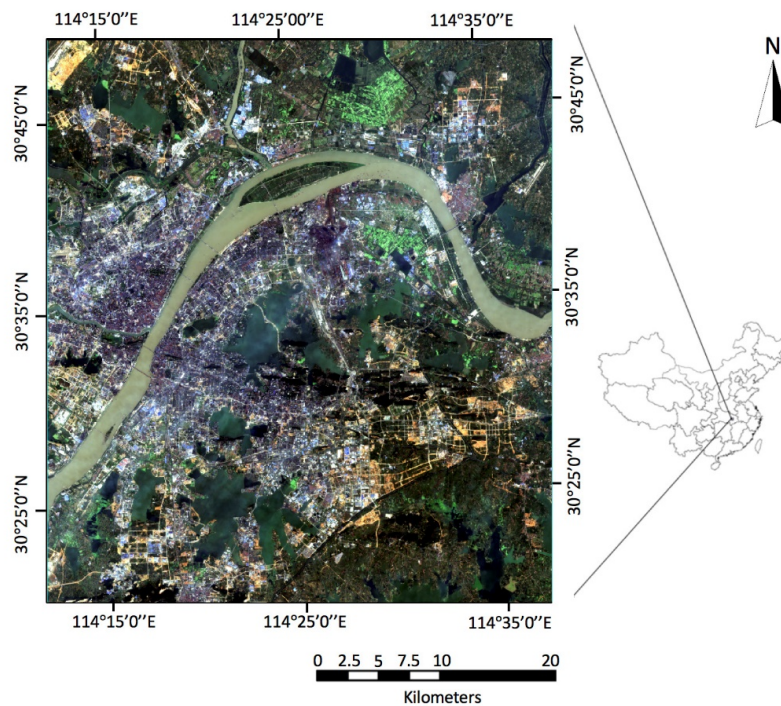
Given the above background, this study intends to incorporate the LLST with MSSI to extract the spatio-temporal pattern of LST through time series clustering. The aims of this study are to (1) apply time series clustering to discover the spatial and temporal patterns of LST simultaneously while avoiding the data comparison dilemma; (2) incorporate LLST and MSSI to characterize LST so that regions with identical LLST temporal pattern and correlation with the surroundings can be located into the same cluster; (3) generate useful information about how LST varies with the advancing of urbanization.

The paper is structured as follows. Section 2 briefly describes the study area and the data sets. Section 3 explains the methodology, including the generation of LLST and MSSI, the recognition of the spatial pattern and the spatio-temporal pattern. Section 4 presents the main results. Section 5 discusses the findings and implications. Finally, the conclusions are given in Section 6.

## 2. Study Area and Data Sets

### 2.1. Wuhan, China

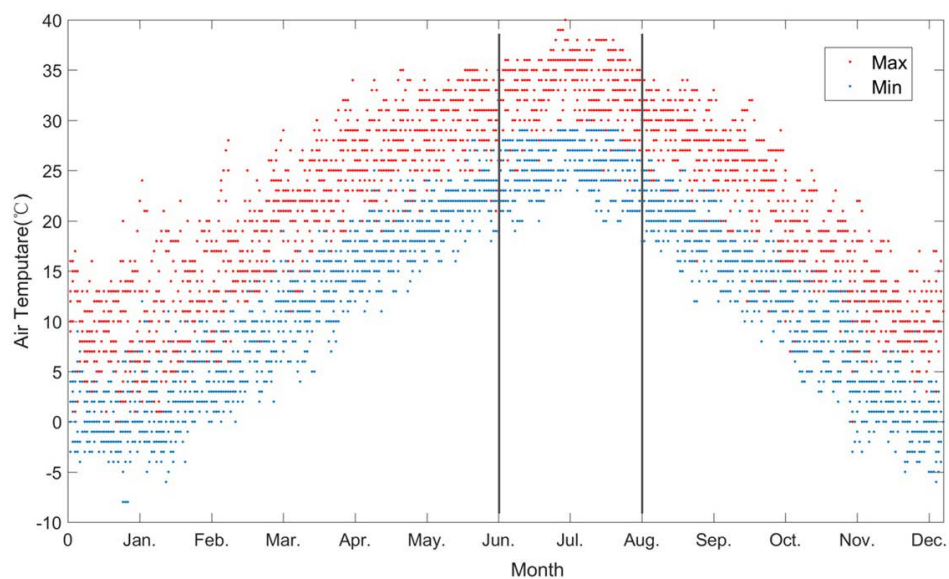
The study area is the city of Wuhan which is located in central China. It is ranked as a megacity of China in 2016 and also on its way to a national central city. Wuhan has experienced rapid expansion and population growth during the past 15 years. Statistic information reveals a 276% increase in the construction area for the whole city from 2002 to 2017. The rapid development has resulted in drastic changes in the urban thermal environment. The development trend is supposed to continue in the next decade, making the thermal environment and dynamics study of the city significant. As shown in Figure 1, a  $49 \times 44$  km area with heterogeneous land cover is selected for the specific study. The area covers most part of the central downtown, also the evident expansion during the past 15 years. The upper-left and lower-right coordinates are  $30^{\circ}47'32''\text{N}$ ,  $114^{\circ}12'01''\text{E}$  and  $30^{\circ}20'19''\text{N}$ ,  $114^{\circ}38'32''\text{E}$ . This coverage is sufficient to exhibit the land composition of the city. Figure 1 employs a Landsat 8 image collected on 6 October 2014 to demonstrate the diversified land cover of the study area.



**Figure 1.** The study area represented by the Landsat 8 image (RGB) on 6 October 2014.

## 2.2. Data Sets

July and August from 2002 to 2017 with a time interval of 3 years are chosen as the study period. According to the daily maximum and minimum temperature of Wuhan from 2011 to 2017 collected from online (<http://www.weatheronline.co.uk/>), July and August represent the hottest months over the year (Figure 2). The same result was also concluded from 2001 to 2010 based on the statistics and visualization conducted by Shen et al. [42]. The consistent selection of the hottest days of each year promotes the compatibility of the data.



**Figure 2.** The date range of the air temperature collected from 2011 to 2017. The red dots represent the daily maximum temperatures, while the blue dots represent the daily minimum temperatures.



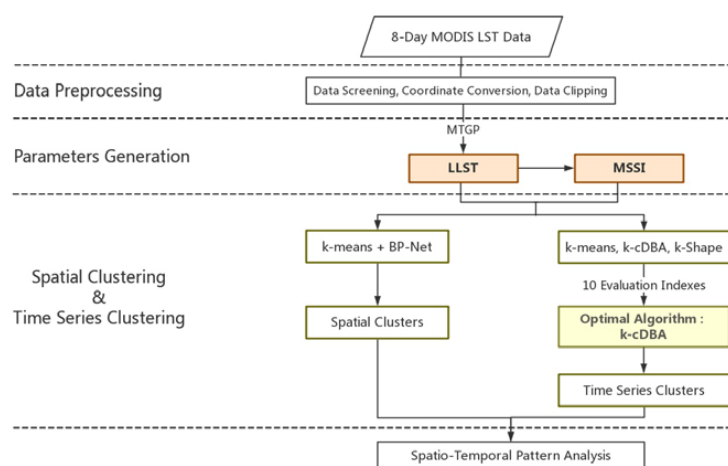
MODIS/Aqua (MYD11A2) V5 LST/E8-Day L3 Global 1 km Grid products acquired at 13:30 are used to represent the typical LST at the hottest hour in Wuhan. The high temporal resolution of the MODIS products guarantees the selection of comparable data for the 6 years and the alternative images as temporal references for MTGP operation. The LST data on a particular day is calculated by a simple average method with the daily MYD11A1 LST products from the corresponding 8 days [43]. The products are generated by utilizing the generalized split-window LST algorithm [44] with an accuracy claimed to be better than 1 °C (0.5 °C in most cases) [45,46]. Specifically, all 6 × 8 images (altogether 6 years for our study, 8 images for the July and August of each year) are downloaded to check the data deficiency degree and weather conditions. Those with the least deficiency among the 8 images of the year and the most similar weather conditions among the 6 years are chosen to represent the specific year. Two other images of the same year with the least deficiency and the nearest time adjacency are applied as temporal references to conduct the MTGP. Lacking of weather station data, the multivariate attributes representing the weather conditions of all 48 days are collected from online (<http://www.weatheronline.co.uk/>). The values are calculated by a single average method with the data collected from the corresponding 8 days. Table 1 demonstrates the information of the final 6 days selected.

**Table 1.** The MODIS data and the weather conditions.

Year	Image Data (Julian)	Average Maximum Temperature (°C)	Average Minimum Temperature (°C)	Average Relative Humidity	Average Wind Force (km/h)
2002	4 July (185)	35.3	26.6	68.6	4.3
2005	12 July (193)	34.2	26.5	65.6	5.9
2008	20 Aug (233)	32.1	25.5	81.5	7
2011	13 Aug (225)	34.5	26.6	67.5	12.8
2014	28 July (209)	35.3	25.8	74.6	10.4
2017	12 July (193)	35.4	28.1	67.3	13.2

### 3. Methodology

The methodology developed in this study follows a workflow for the spatio-temporal pattern extraction of LST. The methodological framework is presented in Figure 3 which includes four main steps: (1) employ MTGP to fill the missing pixels of the origin MODIS data and extract continuous and smooth LLST (Section 3.1). (2) generate the morphological parameter MSSI on the basis of LLST (Section 3.2). (3) incorporate *k*-means and BP-Net to discover the spatial patterns and form a basic understanding about the corresponding variances of the 6 years (Section 3.3). (4) apply and evaluate *k*-means, shape-based *k*-cDBA and *k*-shape to identify the optimal time series clustering algorithm for our study (Section 3.4).



**Figure 3.** The representation of the methodological framework used in this study.

### 3.1. The Latent Pattern of LST

This study applies the latent continuous and smooth surface of LST (LLST) to better capture its overall pattern and generate MSSSI. The MTGP model promoted by Bonilla et al. [47] is utilized to derive LLST. By using shared information across different images, LLST is generated through machine learning models according to the information of the image itself and the temporally adjacent images. The multiple benefits for the application of MTGP in LST pattern recognition were summarized sufficiently [31]. The benefits include its ability to (1) remove noises and fill vacancies, (2) smooth the surface and populate the resolution by a factor of 2 for the later morphological parameter extraction, (3) control local variations and maintain global diversity, and also to (4) fulfill the process without human intervention. Though cloud covered pixels are not considered in the MODIS LST products [48], the pixels around the cloud edges are still impacted by the cloud. The MTGP model can improve the quality of these pixels by referring to the temporally adjacent images. The robust test of MTGP applied for LLST extraction can be checked in Wang et al. [31].

According to Bonilla et al. [47] and Wang et al. [31], the observed LST data set is defined as  $D = \{(x_i, t_{ij}) | i = 1, \dots, n, j = 1, \dots, m\}$ , in which  $x_i$  denotes the  $i$ -th location index in dimensional space  $R^d$ , and  $t_{ij}$  denotes the observation at location  $x_i$  in the  $j$ -th image. Besides,  $n$  is the number of pixels on one image, and  $m$  is the number of images applied in the model. The Gaussian Process (GP) model generalizes the extract form into an infinitely long vector  $[f_1, \dots, f_n]^T$ , among which any finite set of the vector is jointly Gaussian. The model  $f(x) \sim GP(m(x), K^f K^x)$ , is completely defined by the mean function  $m(x)$  and the covariance function.  $K^f$  is the covariance function stating the inter-task information between the images, while  $K^x$  is the covariance function explaining the intra-task information inside one image. In our study, the Automatic Relevance Determinant (ARD) Squared Exponential function is employed. The optimal hyper-parameters of the ARD Squared Exponential function are decided by maximizing the log marginal likelihood.

Any mean value  $\bar{f}^*$  of LLST is predicted by

$$\bar{f}^* = m(x^*) + \left(k^f \otimes k^x(x^*, x)\right)^T \left(K^f \otimes K^x + \Delta \otimes I\right)^{-1} (t - m(X)) \quad (1)$$

where  $\otimes$  denotes the Kronecker product of matrices or vectors, and  $\Delta$  is a diagonal matrix to record noise  $\sigma^2$  [47].

### 3.2. The Morphology

As a spatial and geographical phenomenon, the morphological feature of LST is significant and thus, included in our pattern recognition. The Multi-scale Shape Index (MSSI) [49] is an extension of the Shape Index (SI) promoted by Koenderink and van Doorn [50]. The index estimates shapes after the optimal scale is selected. Wang et al. [31] utilized the index to analyze the morphology of LST in Wuhan. They introduced the morphological parameter to the community of urban climate to enrich the traditional numeric parameter system. The application of MSSI in pattern recognition can distinguish those regions with identical LLST patterns but inconsistent climatic surroundings. The advantages of MSSI over traditional SI were practically verified by Wang et al. [31].

Specifically, the LST pattern  $f(s)$  is projected to the scale space [50,51]  $S$  through the formula below:

$$S(f(s), \sigma) = f(s) * k(s - u, \sigma) = \int_0^s f(u)k(s - u, \sigma)du \quad (2)$$

in which  $k(\cdot)$  indicates the Gaussian kernel, and  $\sigma$  as the convolution scale of location  $u$  on surface  $s$ . Every point in the original image corresponding to the convolution kernel center would shift as the kernel smoothing at different scales. The fluctuation intensity of the original image thus gets

compressed. The feature scale is generated by the maximization of the normalized drift distance at a different convolution scale  $\sigma$ . The normalized distance is calculated by

$$d = \frac{D(f, \sigma)}{\sigma} = \frac{\|S(f, \sigma) - f\|^2}{\sigma} \quad (3)$$

The original SI [50] is calculated through

$$SI = \frac{2}{\pi} \arctan \frac{\kappa_2 + \kappa_1}{\kappa_2 - \kappa_1}, \quad SI \in [-1, 1] \quad (4)$$

where  $\kappa_1$  and  $\kappa_2$  ( $\kappa_1 \geq \kappa_2$ ) represents the principal curvatures generated from a noiseless continuous LLST surface. The value of  $SI$  indicates the protruding and depression of a certain pixel. Specifically, these shapes can be described as cups, ruts, saddles, ridges, and caps.

### 3.3. The Spatial Pattern

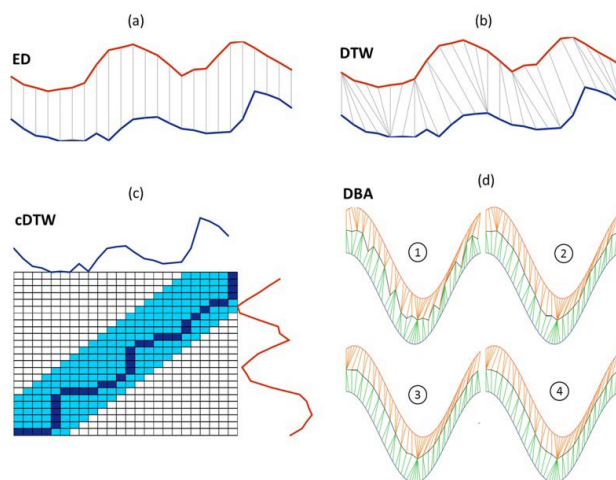
The  $k$ -means and Back-Propagation neural networks (BP-Net) are incorporated to recognize the spatial patterns of LST and form a basic understanding of the temporal variances. The two approaches are jointly applied as a combination of an unsupervised clustering algorithm and a supervised classification algorithm is beneficial for pattern recognition of static data [52,53]. The  $k$ -means algorithm is selected as it is identified to be more efficient than hierarchical, spectral, or  $k$ -medoids methods [23]. Specifically,  $k$ -means is applied to automatically explore the structure of the data without artificial interference. The cluster centroids get updated iteratively until the Euclidean distances (ED) between pixels and centroids are minimized. The BP-Net is further employed to improve the classification robustness with the results from  $k$ -means as priori knowledge.

The selection of the attributes for clustering and the number of clusters ( $k$ ) are two critical problems involved in the clustering process [25]. In this paper, we incorporate LLST with MSSSI to represent the numerical and morphological feature of LST. The optimal  $k$  is chosen where 99% of the data variance can be interpreted while more clusters lead to less significant improvement [53].

### 3.4. The Spatio-Temporal Pattern

Traditional  $k$ -means, shape-based  $k$ -means with the constrained Dynamic Time Warping ( $cDTW$ ) [54] as distance measure, and the Dynamic Time Warping Barycenter Averaging (DBA) [55] method for centroid computation ( $k$ - $cDBA$ ) and  $k$ -shape [23] are applied and compared to select the best algorithm for our study. This is due to the request that the assessment of time series clustering should involve the comparison with simple and stable metrics such as Euclidean distance [56]. Specifically,  $k$ -means with ED is proved to be a robust approach [23] and has been effective during the past 60 years [25].  $k$ - $cDBA$  is a shape-based approach as the distance measure  $cDTW$  is invariant to scale, translation, and shift. The shape-based similarity measure is devoted to discover the similar time-series in time and shape [24]. Dynamic Time Warping (DTW) is an extension of ED that offers a local (non-linear) alignment [23]. It is one of the most popular distance measurement methods for time-series data clustering and the best measure in most domains in spite of numerous alternatives [57], especially for short time series [24]. The  $cDTW$  is a slope constraint version of DTW with improved accuracy and efficiency [54]. It behaves slightly better than DTW and significantly reduces the computation time in the experiment conducted among the 9 distance measures by Wang et al. [58]. DBA [55] is a global averaging method used to extract centroids, which compute coordinates of each sequence as the barycenter by iteration. The incorporation of  $k$ -means and DBA has been testified to be the most robust and efficient of the approaches attempting to conjunct centroid computation methods with DTW [55]. The operating principles of ED, DTW,  $cDTW$ , and DBA are respectively presented in Figure 4.  $k$ -shape is also a shape-based approach which outperforms all the scalable and non-scalable partitioned, hierarchical, and spectral methods in terms of accuracy and efficiency in the experiment

operated by Paparrizos and Gravano [23] on 48 public datasets. The algorithm adopts a normalized version of the cross-correlation as the shape-based distance (SBD).



**Figure 4.** The operating principles of Euclidean Distance (ED), Dynamic Time Warping (DTW), The Constrained Dynamic Time Warping (cDTW), and the Dynamic Time Warping Barycenter Averaging (DBA). (a) Similarity computation under ED; (b) similarity computation under DTW; (c) Sakoe-Chiba band with a warping window of 5 cells (light blue cells in band) and the warping path computed under cDTW (dark blue cells in band) [23]; (d) centroid computation through the Dynamic Time Warping Barycenter Averaging (DBA) [55].

The comparison among the 3 algorithms lies in their differences in similarity measurement and centroid computation approaches as presented in Table 2. Let  $\vec{x} = (x_a, \dots, x_m)$  and  $\vec{y} = (y_a, \dots, y_m)$  of length  $m$  be the two time series for comparison. Let  $S = \{S_1, \dots, S_N\}$  be a set of sequences and  $C = \{C_1, \dots, C_T\}$  be the centroid, where  $N$  is the number of time sequences in one potential cluster and  $T$  is the potential number of the cluster. Specifically, the barycenter of one potential cluster is regarded as the centroid in  $k$ -means [59].

**Table 2.** The comparison of similarity measurement and centroid computation approaches among  $k$ -means, time series clustering algorithm with cDTW as the distance measure, and DBA for centroid computation ( $k$ -cDBA) and  $k$ -shape.

Algorithm	Similarity Measurement	Centroid Computation
$k$ -means	ED [59] compares the two time series as follows: $ED(\vec{x}, \vec{y}) = \sqrt{\sum_{i=1}^m (x_i - y_i)^2}$	The barycenter of potential cluster $j$ is calculated as follows [59]: $C = \text{barycenter}(S) = \frac{\sum_{i=1}^N S_i}{N}$ The barycenter calculated discretely using ED as follows: $C = \text{argmin} \sum_{i=1}^N ED^2(S_i, C)$
$k$ -cDBA	cDTW [54] compares as follows: $D(\vec{x}, \vec{y}) = \min_F \left[ \frac{\sum_{k=1}^K d(c(k)) \cdot w(k)}{\sum_{k=1}^K w(k)} \right]$ where $k \geq m$ . $w(k)$ is a nonnegative weighting coefficient to measure the goodness of warping function between points $c(k) = (i(k), j(k))$ .	DBA [55] optimized the calculation of the barycenter for the application of cDTW. As cDTW takes the sequences association into consideration, the barycenter extraction conducted by using cDTW is as follow: $C = \text{argmin} \sum_{i=1}^N cDTW^2(S_i, C)$
$k$ -shape	SBD compares as follows [23]: $SBD(\vec{x}, \vec{y}) = 1 - \max_w \left( \frac{CCW(\vec{x}, \vec{y})}{\sqrt{R_0(\vec{x}, \vec{x}) R_0(\vec{y}, \vec{y})}} \right)$ where $CCW(\vec{x}, \vec{y})$ is the cross-correlation sequence with length $2m - 1$ . $R_k(\vec{x}, \vec{y})$ is computed as: $R_k(\vec{x}, \vec{y}) = \begin{cases} \sum_{l=1}^{m-k} x_{l+k} \cdot y_l, & k \geq 0 \\ R_{-k}(\vec{y}, \vec{x}), & k < 0 \end{cases}$	$k$ -shape optimizes the centroid extraction by finding the maximizer squared similarities to the other time trajectories [23]. The centroid is calculated iteratively as follows: $C = \text{argmax}_{\vec{x}_i \in P_k} \sum \left( \frac{\max_w CCW(\vec{x}_i, \vec{y}_i)}{\sqrt{R_0(\vec{x}_i, \vec{x}_i) \cdot R_0(\vec{y}_i, \vec{y}_i)}} \right)^2$



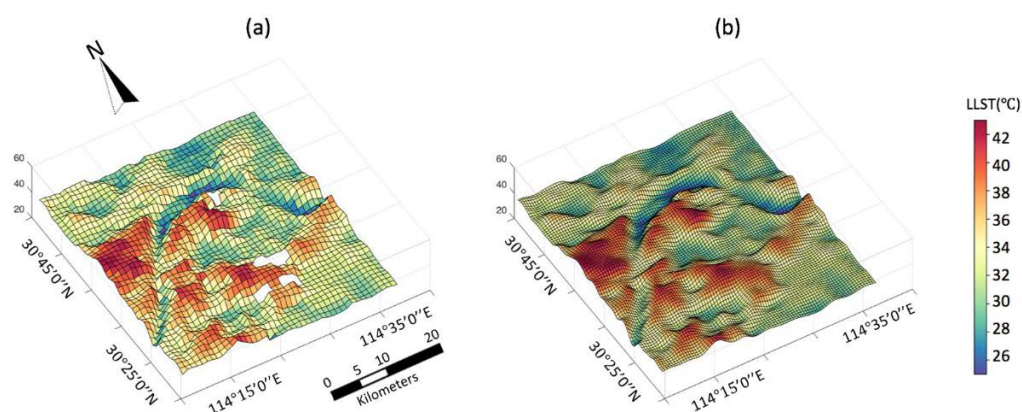
Overall, the clustering procedure includes data representation and normalization, distance measure and clustering algorithm selection, and evaluation. In this study, LLST and MSSI of each year are treated as a whole to identify the geographic regions with homogeneous temporal dynamics of LST patterns. Only those pixels with both homogeneous LLST and MSSI patterns can present similar shapes of the time series and thus, be allocated to the same cluster. Specifically, LLSTs and MSSIs are separately normalized by the Z-normalization method widely recommended for time series clustering [23,60] to make the data comparable. The normalized LLST and MSSI of each year are sorted by year. The time series clustering algorithms are utilized to compare the similarity of each temporal trend line to locate the pixels with homogeneous amplitude and phase. The optimal  $k$  is chosen where the Sum of Squared Error (SSE) reaches relative stability and become acceptable. SSE is the most commonly used measure for time series clusters evaluation [24,61]. The lower the SSE values, the better the clustering.

In the end, 10 external validation indexes are employed to evaluate the results generated by the 3 algorithms as the internal index SSE is not suitable for comparison among different models/metrics [24]. The external indexes include the Rand Index (RI), Adjusted Rand Index (ARI), Purity, Jaccard Score, F-measure, Folkes and Mallow index (FM), Cluster Similarity Measure (CSM), Normalized Mutual Information (NMI), and Entropy [24]. All the indexes range from 0 to 1, where 1 represents the best matching between clustering results and referable ground labels (except for Entropy). The ground label generation is crucial and may result in new uncertainty. The unbiased attitude and elaborate work is demanded to ensure the accuracy of the ground label values. Specifically, 5% of the total samples are randomly selected as ground labels. Those pixels with consistent cluster numbers for all 3 algorithms are chosen as benchmarks. Values of the ground labels are determined through the time series similarities compared with benchmarks and also with each other.

## 4. Results

### 4.1. The LLST Patterns

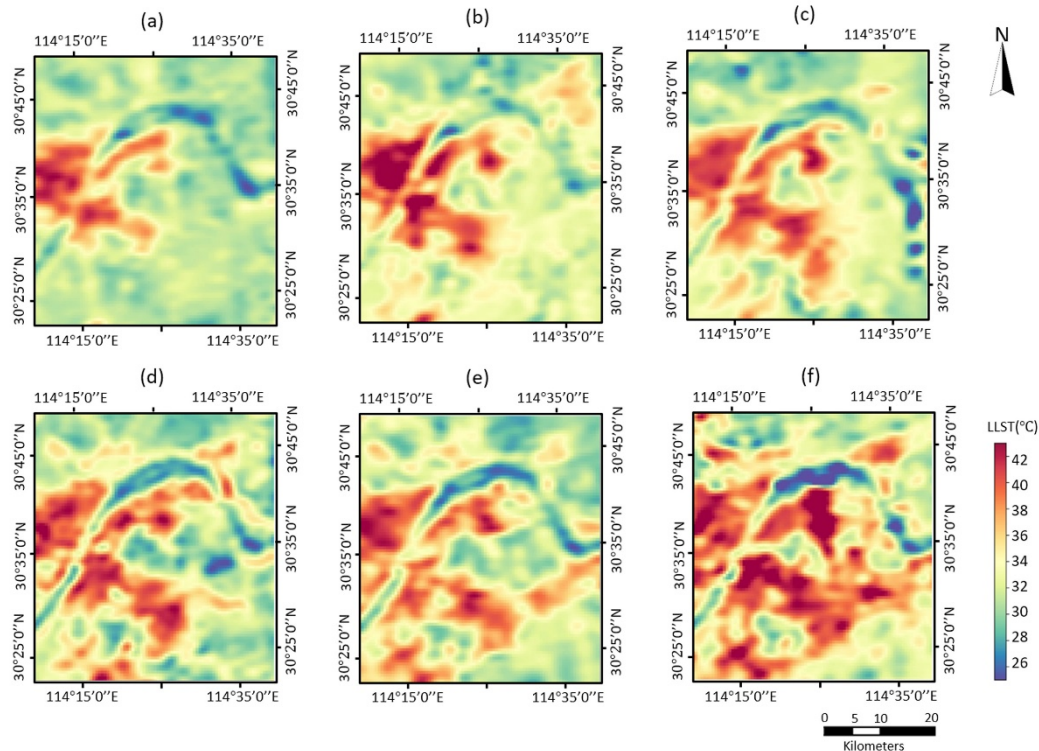
The operation of MTGP is exemplified by the extraction of LLST in 2014. The images acquired on 20 July and 5 August are applied as the temporal reference and the image acquired on 28 July as the spatial reference to fill the missing pixels and generate continuous and smooth surface temperature pattern for the 28 July image. The original and recovered LSTs are shown in Figure 5.



**Figure 5.** The Latent Pattern of LST (LLST) of 28 July 2014 generated by the Multi-Task Gaussian Process Modeling (MTGP). (a) The original image; (b) the LLST.

The MTGP model is conducted on all 6 images with appreciable accuracy. The LLSTs of the 6 years are shown in Figure 6. There is an obvious expanding and deterioration tendency of high level LLST from 2002 to 2017. Table 3 illustrates the accuracy evaluation results of the operation

with RMSE, standard deviation, and Correlation Coefficient (CC). The RMSEs of each year are all within two standard deviation from the LLSTs. The CCs are all beyond 0.95 which indicate an appreciable correlation and information preservation degree between the original LST images and the generated LLSTs.



**Figure 6.** The LLSTs of the 6 years generated by MTGP. (a) 2002; (b) 2005; (c) 2008; (d) 2011; (e) 2014; (f) 2017.

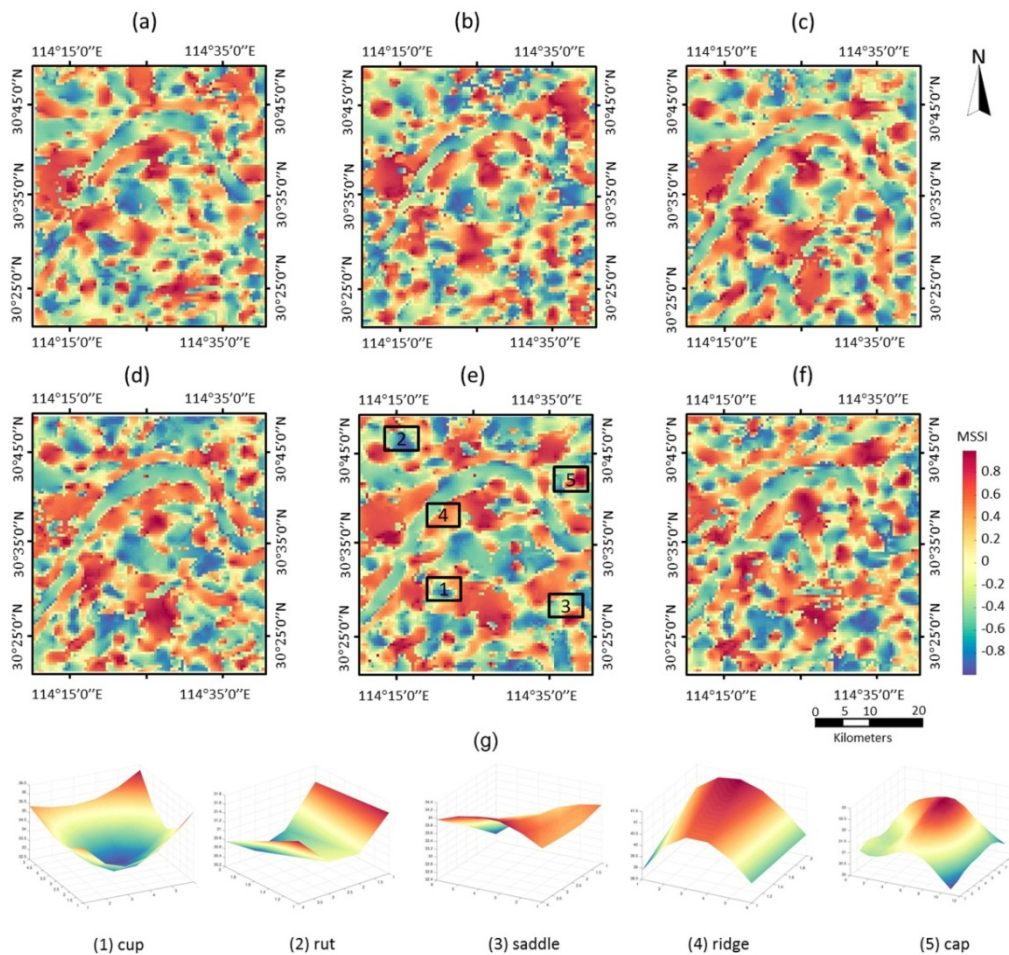
**Table 3.** The statistical summary of Latent Pattern of LST (LLST) generation for each year.

Year	Image Data (Julian)	RMSE (°C)	Standard Deviation (°C)	2 × Standard Deviation (°C)	Correlation Coefficient
2002	0704 (185)	0.23	0.27	0.54	0.99
2005	0712 (193)	0.23	0.21	0.42	0.98
2008	0820 (233)	0.14	0.12	0.24	0.99
2011	0813 (225)	0.23	0.26	0.52	0.97
2014	0728 (209)	0.30	0.30	0.60	0.96
2017	0712 (193)	0.20	0.22	0.45	0.99

#### 4.2. The Morphological Characteristics of LLSTs

The morphological characteristic of LST is further calculated as MSSI based on the continuous and smooth LLST. The MSSIs of the 6 years are illustrated in Figure 7a–f. The morphologies of the LLSTs are well described over the area. The diversified MSSI values also exert an appreciable demonstration of the landscape pattern. Typical morphological shapes such as cup, rut, saddle, ridge, and cap are illustrated in Figure 7e by taking 2014 as an example. The shapes can be illustrated by the LULC types inner and surrounding the 5 regions as research has revealed that the LST of built-up and development areas is always the highest, followed by semi-bare land, vegetation, and water areas [6,31,62]. Accordingly, location 1 in Figure 7e performs as a cup shape as a water area surrounded by vegetation and built up area. Also mainly as a water area, location 2 demonstrates as a rut shape due

to its adjacency of vegetation and built up area at the ends of the north and the south. Location 3 reveals a saddle morphology as it is a mixed land cover centered area with built up land at the northwest and southeast ends, while having vegetation areas at the other two ends. Location 4 performs as a ridge shape because it is a built up area with water area at the northwest end and vegetation area at the southeast end. Besides, being a built up land surrounded by vegetation area makes the MSSI of location 5 presents as a cap shape.



**Figure 7.** The Multi-Scale Shape Indexes (MSSIs) of the 6 years indicating the surface morphology of LLSTs. The MSSI of (a) 2002, (b) 2005, (c) 2007, (d) 2011, (e) 2014, (f) 2017, (g) the classic morphologies taken from the MSSI of 2014.

#### 4.3. The Spatial Patterns

The spatial patterns of LSTs are extracted through the combination of  $k$ -means and BP-Net. The correlation between the normalized LLSTs and MSSIs of the 6 years are checked with a maximum Correlation Coefficient (CC) of 0.47 and an average value of 0.31, which indicates a tolerable collinearity between the two parameters. The weights of 0.6 and 0.4 are respectively assigned to LLST and MSSI to relative highlight the importance of LLST, also reducing the degree of fragmentation of the final result. The number of  $k$  is tested from 1 to 19 for all 6 years by  $k$ -means to calculate the corresponding data variances. The optimal  $k$  is chosen as 7 where the clustering result can interpret over 99% of the data variance while more clusters are undesired, as shown in Figure 8. An odd value for breaks is also advantageous as it affords a central class for interpretation reference.



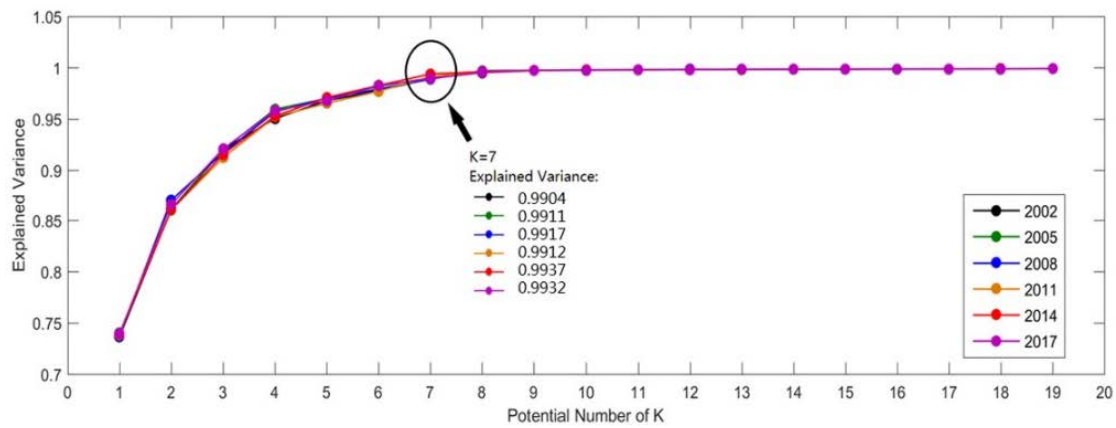


Figure 8. The determination of the optimal cluster number  $k$  for spatial clustering.

According to the spatial classification results of the 6 years, there is an expansion and dispersion tendency of high LST patterns from 2002 to 2017 as shown in Figure 9. Specifically, classes from 1 to 7 are sorted according to the mean LLST of each cluster, where 7 represent the highest mean value. As much as we try, the data of each year is still not normalized enough for direct comparison. Fortunately, the specific tendency can be checked through the spatial patterns of the 6 years generated through classification. The expansion and dispersion tendency is partially generated by urbanization and industrialization while referring to contemporaneous Landsat images. The main development activities are the Yangluo Economic Development Zone on the northeast corner, the East Lake High-Tech Development Zone in the lower part, and the other conventional expansion of the built up area.

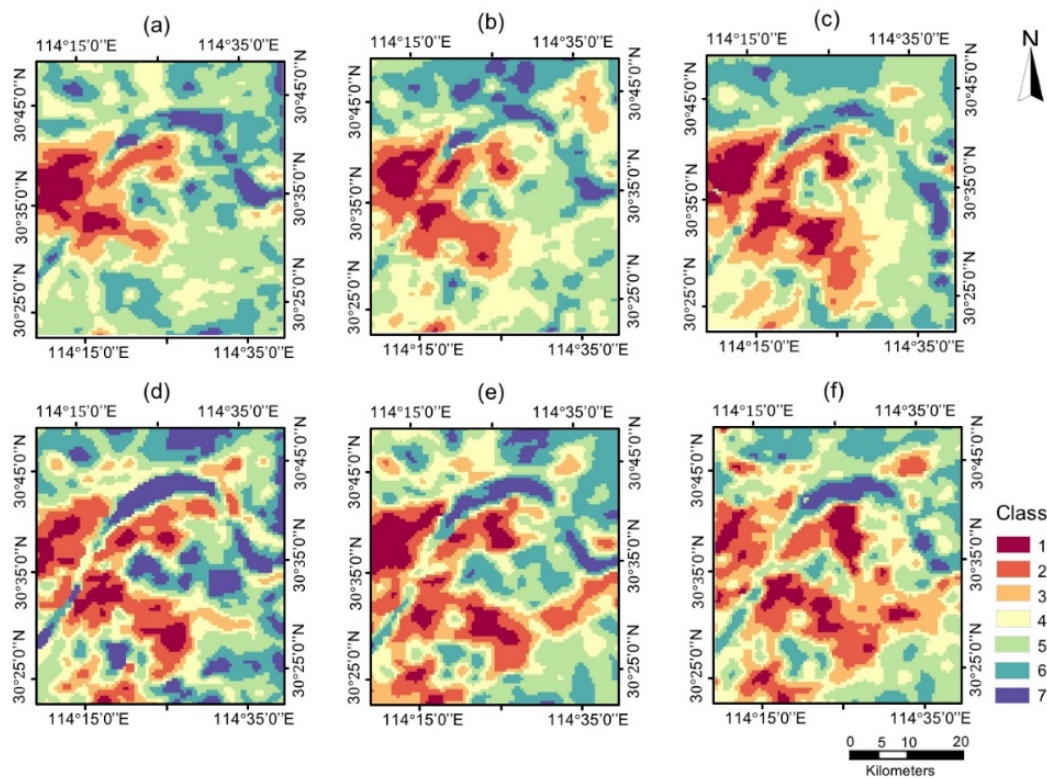
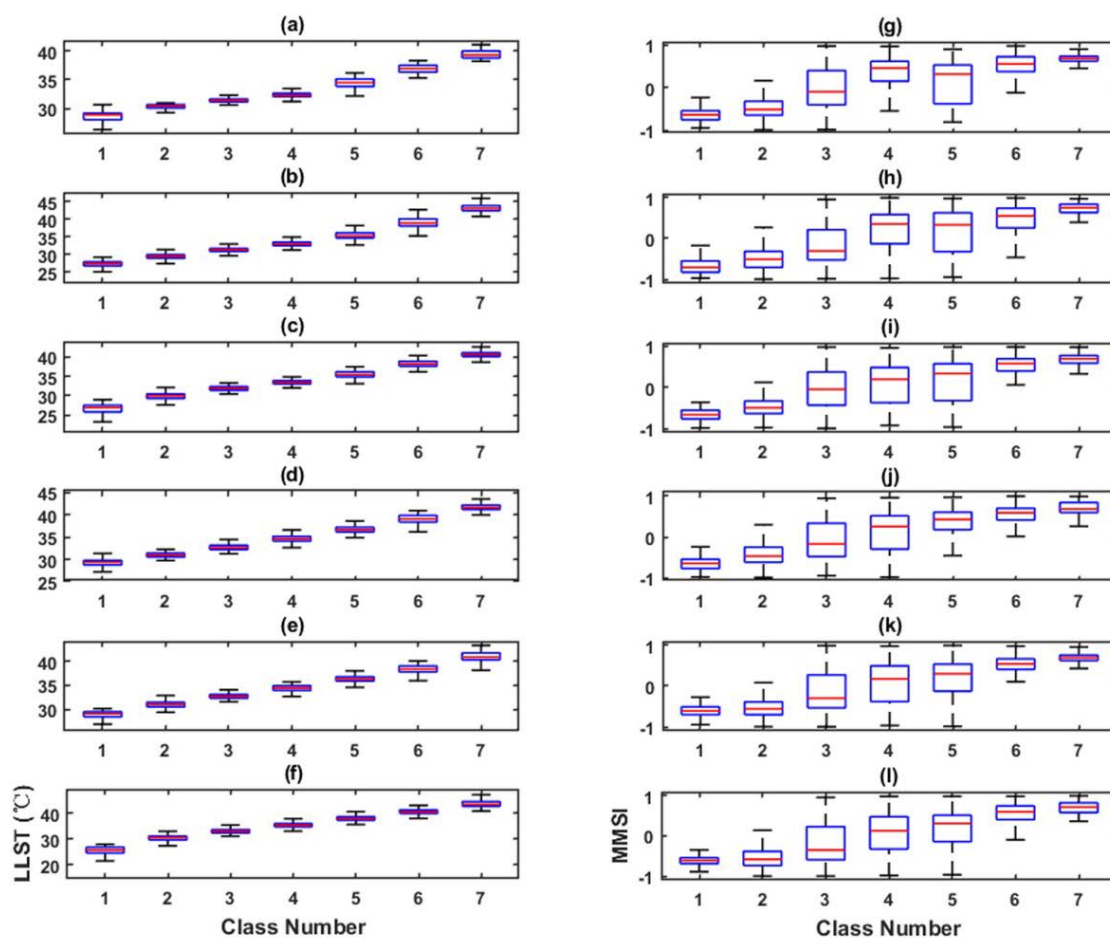


Figure 9. The spatial classification results of the 6 years. (a) 2002; (b) 2005; (c) 2008; (d) 2011; (e) 2014; (f) 2017.

LLST and MSSSI distribution among all the classes are further inspected through boxplot analysis (Figure 10). LLST rises monotonically from class 1 to class 7, whereas the rising tendency of MSSSI is blocked at class 5 for 2002 and 2005. The discordance occurs as MSSSI presents inconsistent variation trends with LLST. Specifically, class 5 represents the areas with higher LLST but lower MSSSI than class 4 for the two years. Contemporaneous Landsat images are referred to in order to identify the representative LULC type of each cluster. Generally, class 1 represents most of the water area such as the Yangtze River and a small part of vegetation area away from the main built up land. Class 2 signifies the water area nearer built up land and most of the vegetation area. Class 3 mainly contains the suburbs mixed with sparse buildings and some vegetation. Class 4 represents areas with relative higher building intensity than Class 3. Class 5 covers almost the urban development zone of each year. Class 6 and Class 7 represent the main developed area of each period with Class 7 revealing a severe thermal phenomenon at industrial districts and high building intensity area.

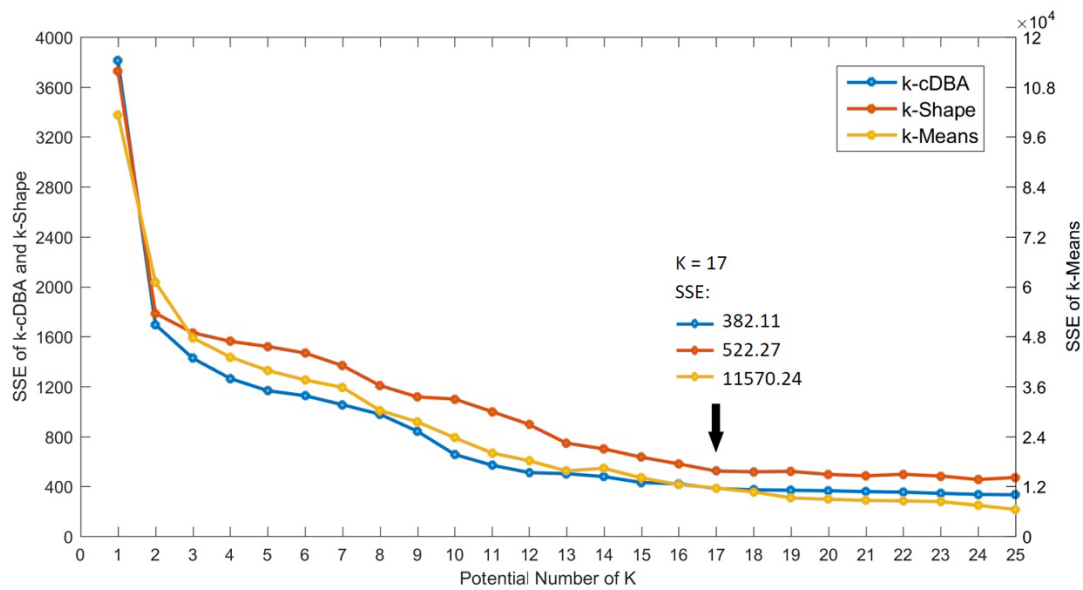


**Figure 10.** The boxplots of the spatial classification results. LLST of (a) 2002, (b) 2005, (c) 2008, (d) 2011, (e) 2014, (f) 2017; MSSSI of (g) 2002, (h) 2005, (i) 2008, (j) 2011, (k) 2014, (l) 2017.

#### 4.4. The Spatio-Temporal Pattern

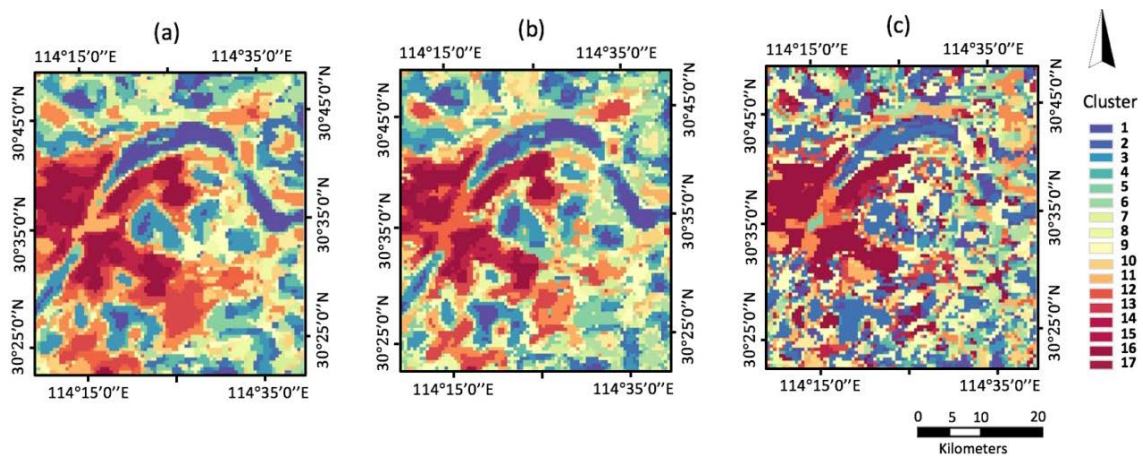
$k$ -means,  $k$ -cDBA, and  $k$ -shape are conducted to cluster the geo-referenced time series of LLSTs and MSSIs. The cluster number of  $k$  is tested from 1 to 25 to calculate the corresponding SSE. The optimal  $k$  is chosen as 17 where SSE becomes relative stable while more clusters leads to less significance but great interpretation difficulty for all algorithms, as shown in Figure 11.





**Figure 11.** The determination of the optimal cluster number  $k$  for time series clustering. The left ordinate measures the Sum of Squared Error (SSE) of  $k$ -cDBA and  $k$ -shape; the right ordinate measures the SSE of  $k$ -means.

$k$ -cDBA is found to be the optimal algorithm for our study. Specifically, clusters from 1 to 17 are sorted according to the mean LLST of each cluster, where 17 represents the highest mean value. The clustering results are shown in Figure 12. 400 pixels (around 5% of the total samples) are randomly selected to generate ground labels. The evaluation result of the 3 algorithms is represented in Table 4. All 10 external indexes support that  $k$ -cDBA as the best algorithm for our study, followed by  $k$ -means and  $k$ -shape.  $k$ -cDBA is advantageous for its shape-based and short term suited similarity measure, as well as the effective centroid computation method.  $k$ -means is once again testified as a robust and effective algorithm for clustering.  $k$ -shape behaves less satisfactory on our short term time series. Its advantages in scaling, translation, and shift invariance makes it more suitable for long term time series where the data structure is more important than the detailed changes in time points. Specifically, the clustering results of  $k$ -cDBA and  $k$ -means are overall similar, whereas that of  $k$ -shape is too fragmented for pattern analysis. The ground labels reveal that  $k$ -cDBA provides more detailed and precise information for developed and developing area in suburbs compared to  $k$ -means.

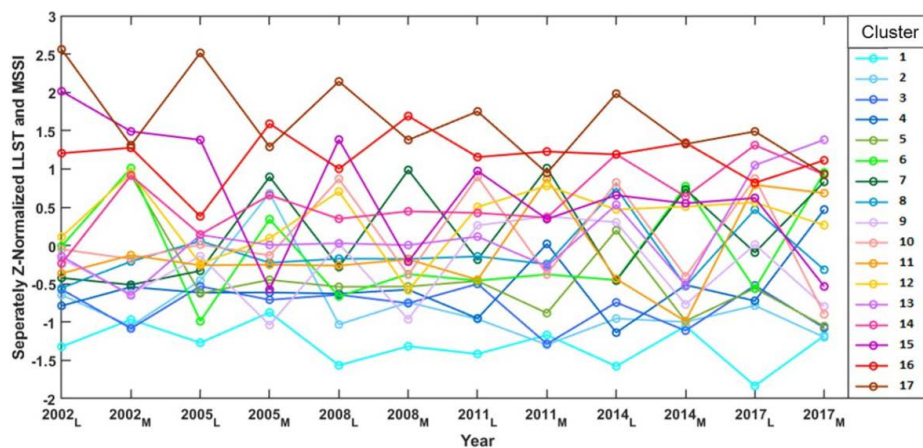


**Figure 12.** The time series clustering result of the algorithms. (a)  $k$ -means; (b)  $k$ -cDBA; (c)  $k$ -shape.

**Table 4.** The evaluation result of *k*-means, *k*-cDBA and *k*-shape, illustrated by 10 external indexes.

	<i>k</i> -Means	<i>k</i> -cDBA	<i>k</i> -Shape
Accuracy	0.6775	0.7725	0.6150
RI	0.9355	0.9387	0.9067
ARI	0.4680	0.4800	0.3896
Purity	0.6075	0.6475	0.5530
Jaccard Score	0.3355	0.3447	0.2423
F-measure	0.5735	0.6085	0.4623
FM	0.5025	0.5127	0.4523
CSM	0.5676	0.6174	0.5314
NMI	0.5991	0.6186	0.4808
Entropy	0.3865	0.3646	0.5058

The basic information of the 17 time series clusters generated by *k*-cDBA is further investigated. The time series centroids of all clusters are presented in Figure 13, where LLSTs and MSSIs are separately Z-normalized and then bound together, and finally sorted according to the year. Basic statistical indicators such as maximum, minimum, mean values and Standard Deviations (Stds) are represented in Table 5.

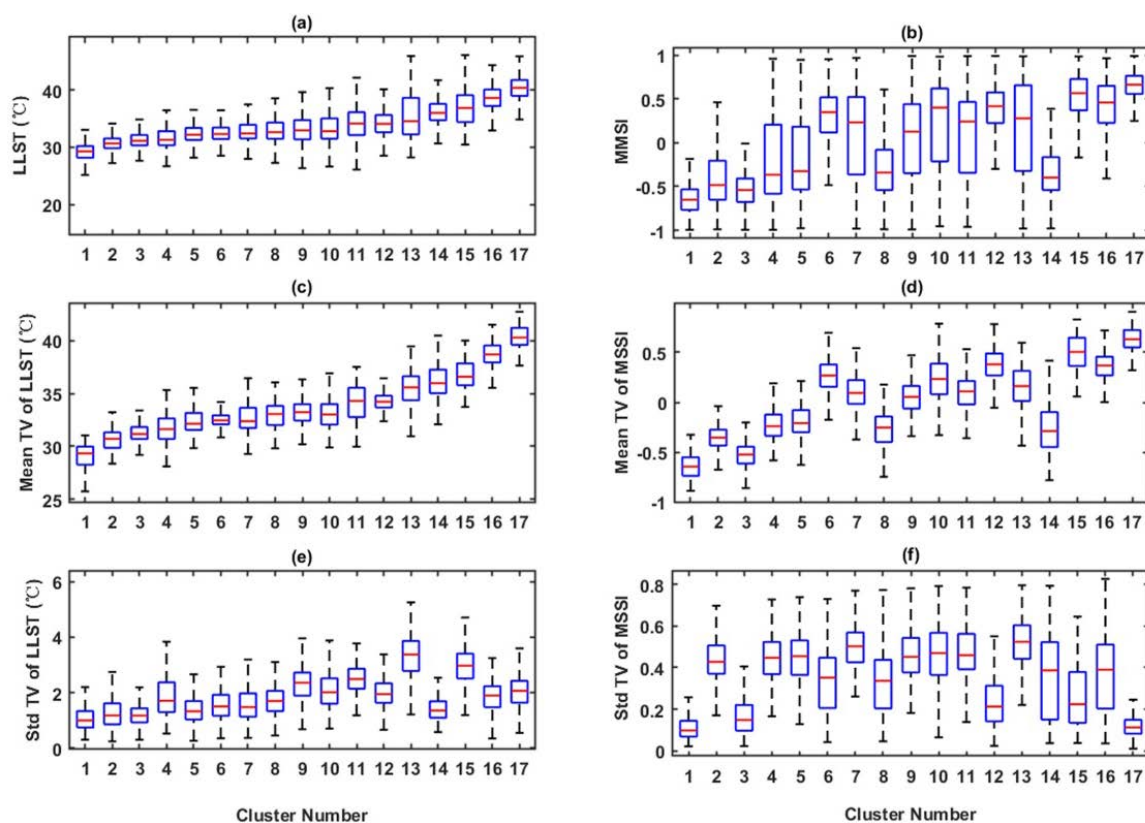


**Figure 13.** The time series centroids of all 17 clusters. 2002<sub>L</sub> and 2002<sub>M</sub> represent the LLST and MSSi of the year 2002, and so on.

**Table 5.** The statistical summary of *k*-cDBA.

Cluster Number	Maximum		Minimum		Mean		Standard Deviation	
	LLST	MSSi	LLST	MSSi	LLST	MSSi	LLST	MSSi
1	16.61	−0.99	33.95	0.44	29.05	−0.64	1.69	0.17
2	16.23	−0.99	42.63	0.97	30.72	−0.35	1.83	0.43
3	21.51	−0.99	36.33	0.33	31.22	−0.53	1.48	0.21
4	25.27	−0.99	41.13	0.96	31.75	−0.22	2.29	0.45
5	21.93	−0.98	41.69	0.95	32.43	−0.19	1.92	0.46
6	26.05	−0.98	41.03	0.96	32.60	0.26	1.80	0.38
7	19.07	−0.98	42.39	0.97	32.85	0.10	2.26	0.50
8	26.31	−0.99	39.95	0.96	32.95	−0.27	2.06	0.38
9	26.09	−0.99	41.95	0.99	33.21	0.05	2.47	0.47
10	25.23	−0.95	44.53	0.98	33.40	0.24	2.74	0.49
11	23.89	−0.96	42.91	0.99	34.14	0.10	3.02	0.48
12	28.51	−0.75	42.11	0.99	34.27	0.38	2.05	0.29
13	28.23	−0.98	45.89	0.99	35.42	0.16	3.53	0.53
14	28.89	−0.98	43.43	0.98	36.17	−0.26	2.19	0.44
15	30.45	−0.93	46.57	0.99	36.79	0.50	3.07	0.32
16	32.09	−0.95	46.07	0.97	38.65	0.36	2.12	0.41
17	34.39	−0.93	47.59	0.99	40.33	0.64	2.25	0.21

The structural information of the result generated by *k-c*DBA is further studied through boxplot analysis. There is a prominent rising tendency of LLST from cluster 1 to 17 while MSSSI exhibits a periodic up and down trend; as illustrated in Figure 14a,b. Take Clusters 14–16 for instance; the LLSTs ascend from 14 to 15, then 16; whereas the MSSSIs ascend from 14 to 16, then 15. It is because the decreasing trends of the LLSTs surrounding Cluster 16 and 14 are less significant, especially for Cluster 14. Besides, it is obvious that Clusters 8 and 9 preserve similar LLSTs but distinct MSSSIs. This demonstrates a desirable time series clustering result where pixels with uniform LLSTs but diversified MSSSIs are separated into different clusters. Boxplot analysis of the temporal dynamic is separately highlighted as the interpretation of time series clustering result is difficult especially when the data is high-dimensional [30]. Specifically, the means (Figure 14c,d) and standard deviations (Std) (Figure 14e,f) of the LLST and MSSSI time series of each pixel are analyzed through the boxplot. The results reveal that the temporal variations of the mean LLST and MSSSI exhibit similar patterns with the overall LLST and MSSSI data. However, the Std tendencies of LLST and MSSSI are relatively irregular and inconsistent with those of the overall LLST and MSSSI data. The heterogeneous temporal dynamics of the clusters can be further explored by consulting to the actual LULC changes over the 6 years.

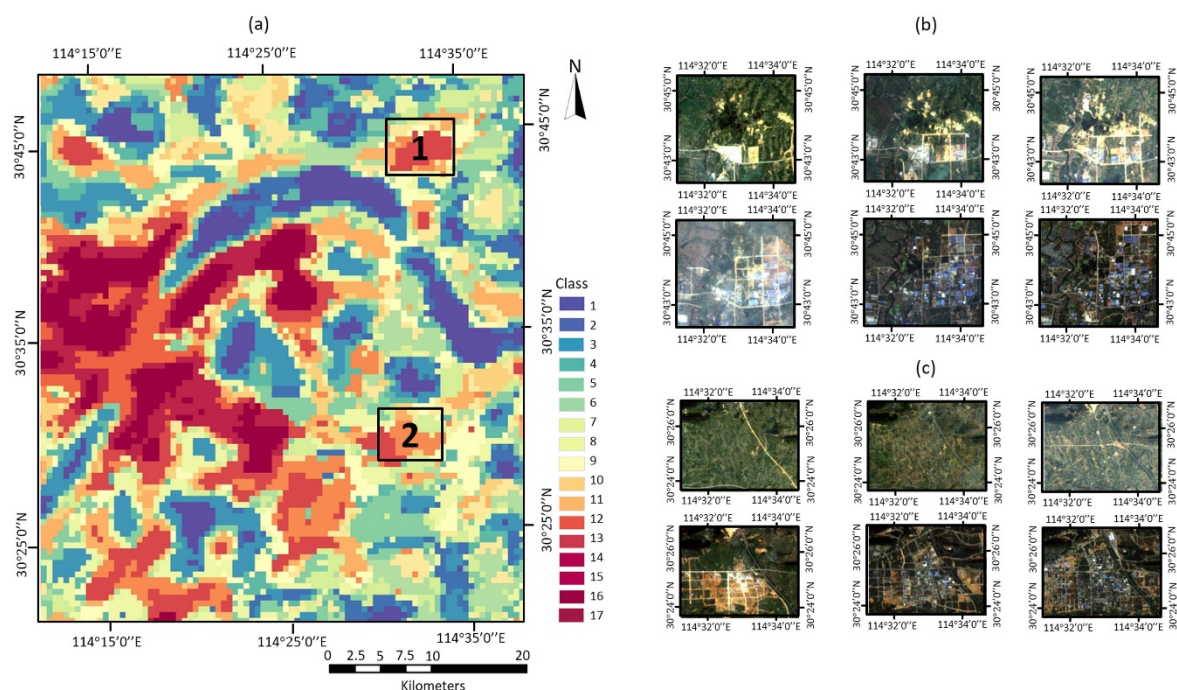


**Figure 14.** The boxplots of the time series clustering results. The whole boxplot of (a) LLSTs; (b) MSSSIs; The temporal variance (TV) boxplot of (c) the mean LLSTs of each time series; (d) the mean MSSSIs of each time series; (e) the standard deviations (Std) of LLSTs of each time series; (f) the Std of MSSSIs of each time series.

Those clusters preserving large Std values for LLST or MSSSI are worthy of great emphasis. Among them, Clusters 15 and 13 are taken as examples to analyze how LLST and MSSSI of the two clusters vary along with urbanization by taking the corresponding LULC trajectories from 2002 to 2017 with contemporaneous Landsat images as references. Specifically, Cluster 15 possesses the second highest Std of LLST among the 17 clusters whereas the variance of MSSSI is relatively weak (Figure 14e,f). The first box in Figure 15a is further taken as an example of Cluster 15. The LULC



trajectory represented in Figure 15b reveals that the area is firstly a vegetation area mixed with bare land, then gradually varies into built up land from 2002 to 2007, which results in a large variance of LLST. However, the LLST difference between the region with its surroundings of all 6 years is relative stable as there is an obvious but relative lagging development tendency of the surroundings, which makes the MSSI variance stay relatively small. Besides, Cluster 13 preserves the most remarkable variances of both LLST and MSSI. The 2nd box in Figure 15a is further taken as an example for Cluster 13. The region mainly indicates the newly-built Development Zones of the city which is situated far from the urban center. The LULC trajectory represented in Figure 15c suggests that it experienced significant construction activities from almost completely vegetation areas to built-up or developing areas. The LLST ascended evidently for the replacement of artificial surface elements for natural land cover. Most of the areas maintained as natural land cover in the earlier period which leads to low MSSI value in the first place, the MSSI then went up in the late period with the advancing of nonsynchronous construction between the region and its surroundings, making the temporal variance of MSSI distinct.



**Figure 15.** The Land Use and Land Cover (LULC) trajectory analysis of Clusters 15 and 13 by taking two areas as examples. (a) The time series clustering result of  $k$ -cDBA where the first box represents the example of Cluster 15 while the second represents Cluster 13; the LULC trajectory demonstration of box (b) 1 and (c) 2.

## 5. Discussion

This study focuses on the spatio-temporal pattern analysis of LST through time series clustering. Compared to previous studies [5–8,18], the presented research takes the temporal non-stationarity into consideration for the pattern recognition of LST. In fact, the covering of the temporal dimension increases the optimal cluster number from 7 to 17 as illustrated in Sections 4.3 and 4.4. Besides, although spatial patterns of the 6 years displayed in Figure 8 reveal an overall expansion and dispersion tendency of high LST area, the tendency is not necessarily accurate as the data is not completely normalized for direct comparison. To take one step back, the anisotropic tendency makes it difficult to precisely interpret the temporal variance of the spatial patterns. In contrast, time series clustering effectively classifies those trend lines into 17 heterogeneous clusters according to when major variance occurs and how much is the variance, as illustrated in Figures 13–15. More specifically, LLST and MSSI are

coupled for the pattern recognition of LST, hence, the pixels within one cluster preserve not only homogeneous LLST pattern but also parallel correlation with the surroundings. The incorporation of these two parameters follows the general science that cities are complex integrated systems other than a simple combination of isolated regions. The LST variance of one area is closely related to its surroundings, hence, the spatio-temporal pattern recognition of it should not violate the correlation.

This study facilitates the collaboration of urban thermal research and urban planning by transforming knowledge about LST patterns into the research scope and scale of urban planning. The incorporation of the urban and environmental systems is critical as urban systems may threaten the sustainability of the natural environment. The challenge is asserted to be whether science can deliver the related knowledge into domains like urban planning and management [63]. This study promotes the knowledge delivery on two levels. (1) the investigation of LST other than UHI breaks away from the constraint of “urban-rural” dichotomy into intra-city thermal pattern analysis. (2) The 17 time series clusters of homogeneous LST patterns further correspond to the zoning custom of urban planning and design, based on which customized strategies and measures can be implemented according to their specific spatio-temporal patterns.

Furthermore, this study complements the zoning scheme in urban planning in the sense of sustainability. Although sustainability has been defined as the simultaneous sustainability of society, economy, and environment [64], conventional zoning still places the emphasis on the functions of society and the economy and divides the land use types accordingly into residential, industrial, or commercial zones [53]. Some achievements in the science community have contributed to the compensation of such bias. Local Climate Zone (LCZ) [4,53,65] has been promoted to “classify urban areas into zones with homogeneous LST behaviors based on the climate sensitive indicators extracted from land surface factors” [53]. According to the framework proposed by Oke in 1982 [66], the recognition and description of the climate phenomenon should be firstly explored before turning it into the mechanism level. This study, hence, complements the content of LCZ by performing the zoning of LST patterns on the phenomenon level by incorporating time dimension with space dimension, and also LST with MSSSI for clustering. The clusters zoned by such LST patterns provide significant knowledge for the domain of urban planning to better understand the hierarchical environmental functions within the city.

However, there are two shortages worth discussing. (1) This study applies a 3-year interval for the pattern recognition of LST from 2002 to 2017 based on the underlying assumption that the impact of temporal scale on the final results is negligible, so that more emphasis could be placed on the selection of parameters and time series algorithms. However, the selection of temporal intervals may alter the final results of the time series clustering [67]. Hence, further research is demanded to investigate the temporal scale problem in all temporal analyses, including time series clustering and spatio-temporal modeling and forecasting. (2) Although this study employs 3 effective raw-based time series clustering algorithms (those work directly with the raw data) [22] to identify the optimal algorithm for our study, the first-rank  $k$ -cDBA reveals its deficiency in classifying the small part of the Yangtze River connecting the two downtown sides. Specifically,  $k$ -cDBA identifies the region into cluster 14 compared to that of cluster 12 for  $k$ -means, as shown in Figure 12. This is, to a great extent, resulted by the error in the original LST data. However, more attempts for algorithm selection are still demanded to improve the clustering accuracy. The exploration and comparison of some newly emerging model-based algorithms (indirectly with models built on the raw data) [22], especially those targeted for multidimensional data is thus, worthwhile. The ground label generation also resulted in new uncertainty although great emphasis has been attached.

Time series clustering can further be utilized as a pre-processing step for the forecast and analysis of LST [68–70]. The clustering of historical time series transforms the complicated datasets into a much smaller number of categories [30], thus, simplifying the traditional procedure of generating quality input data for later forecasting [70]. Hence, forecasting can be further operated on this basis through spatio-temporal modeling so that the impact of a planning program on the urban thermal environment



can be evaluated. To better fulfill this work, LST images with considerable resolutions, together with LULC or even more detailed planning indexes are thus required.

## 6. Conclusions

This study introduces time series clustering to characterize the spatio-temporal pattern of LST based on its latent pattern and morphology. *k*-means and BP-Net are firstly applied to extract the spatial pattern of the 6 years. All 6 years are divided into 7 classes. The spatial classes provide a basic understanding of the temporal variances over the years, which reveals a remarkable expansion and dispersion tendency of LST from 2002 to 2017. *k*-means, *k*-cDBA, and *k*-shape are then operated to generate time series clusters. Multiple external evaluation indexes indicate *k*-cDBA as the optimal algorithm for our study. The study area is divided into 17 geographic regions with heterogeneous temporal dynamics of LST patterns. Customized mitigation strategies from the view of urban planning can be allocated to these clusters according to their specific LLST and MSSSI patterns. Besides, further spatio-temporal modeling can be operated on the basis of these clusters to characterize the spatio-temporal dynamics of LST.

As investigated in the discussion section, this study advances the knowledge in urban thermal research and urban planning on three levels: (1) it promotes the collaboration between urban thermal research and urban planning as the time series clusters accommodate to the zoning custom in the planning domain, thus facilitating the transformation of knowledge about LST patterns into the research scope and scale of urban planning; (2) it complements the zoning in the urban planning domain based on environmental functions on the phenomena level, hence, contributing to the integrated development of urban sustainability; (3) it covers the non-stationarity in the temporal dimension to investigate how LST varies along with urbanization. However, the problem of temporal scale should not be neglected. Besides, more attempts on the selection of the optimal time series algorithm are demanded to promote the robustness of our research.

**Acknowledgments:** This research is supported by the National Natural Science Foundation of China (No. 41331175 and 51378399). We also thank the four anonymous reviewers for their insightful comments that have been very helpful in improving this paper.

**Author Contributions:** Huimin Liu proposed the research concept, designed the experiments, conducted data analysis, and composed the manuscript. Qingming Zhan designed the framework of this research and participated in manuscript editing. Chen Yang participated in the data preparation and analysis. Jiong Wang reviewed and edited the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

ARD	Automatic Relevance Determinant
ARI	Adjusted Rand Index
BP-Net	Back-Propagation Neural Networks
CC	Correlation Coefficient
cDTW	The Constrained Dynamic Time Warping
CSM	Cluster Similarity Measure
DBA	the Dynamic Time Warping Baryleft Averaging
ED	Euclidean Distance
FM	Folkes and Mallow index
GP	Gaussian Process
<i>k</i> -cDBA	Time Series Clustering Algorithm with <i>c</i> -DTW as the Distance Measure, and DBA for Centroid Computation
LCZ	Local Climate Zone
LST	Land Surface Temperature
LLST	Latent Pattern of LST
LULC	Land Use and Land Cover
MODIS	MODerate-resolution Imaging Spectroradiometer

MTGP	Multi-Task Gaussian Process Modeling
MSSI	Multi-Scale Shape Index
NMI	Normalized Mutual Information
RI	Rand Index
SI	The Shape Index
SSE	Sum of Squared Error
Std	Standard Deviation
TIR	Thermal Infrared

## References

1. Weng, Q. Thermal infrared remote sensing for urban climate and environmental studies: Methods, applications, and trends. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 335–344. [[CrossRef](#)]
2. Weng, Q.; Fu, P. Modeling annual parameters of clear-sky land surface temperature variations and evaluating the impact of cloud cover using time series of landsat tir data. *Remote Sens. Environ.* **2014**, *140*, 267–278. [[CrossRef](#)]
3. Fu, P.; Weng, Q. A time series analysis of urbanization induced land use and land cover change and its impact on land surface temperature with Landsat imagery. *Remote Sens. Environ.* **2016**, *175*, 205–214. [[CrossRef](#)]
4. Stewart, I.D.; Oke, T.R. Local climate zones for urban temperature studies. *Bull. Am. Meteorol. Soc.* **2012**, *93*, 1879–1900. [[CrossRef](#)]
5. Weng, Q.; Lu, D.; Schubring, J. Estimation of land surface temperature–vegetation abundance relationship for urban heat island studies. *Remote Sens. Environ.* **2004**, *89*, 467–483. [[CrossRef](#)]
6. Chen, X.; Zhao, H.; Li, P.; Yin, Z. Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes. *Remote Sens. Environ.* **2006**, *104*, 133–146. [[CrossRef](#)]
7. Song, J.; Du, S.; Feng, X.; Guo, L. The relationships between landscape compositions and land surface temperature: Quantifying their resolution sensitivity with spatial regression models. *Landsc. Urban Plan.* **2014**, *123*, 145–157. [[CrossRef](#)]
8. Wang, J.; Zhan, Q.; Guo, H.; Jin, Z. Characterizing the spatial dynamics of land surface temperature–impervious surface fraction relationship. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *45*, 55–65. [[CrossRef](#)]
9. Andrienko, G.; Andrienko, N.; Demšar, U.; Dransch, D.; Dykes, J.; Fabrikant, S.; Jern, M.; Kraak, M.; Schumann, H.; Tominski, C. Space, time and visual analytics. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 1577–1600. [[CrossRef](#)]
10. Mohan, M.; Kandya, A. Impact of urbanization and land-use/land-cover change on diurnal temperature range: A case study of tropical urban airshed of india using remote sensing data. *Sci. Total Environ.* **2015**, *506–507*, 453–465. [[CrossRef](#)] [[PubMed](#)]
11. Weng, Q.; Fu, P. Modeling diurnal land temperature cycles over Los Angeles using downscaled goes imagery. *ISPRS J. Photogramm. Remote Sens.* **2014**, *97*, 78–88. [[CrossRef](#)]
12. Fu, P.; Weng, Q. Temporal dynamics of land surface temperature from landsat TIR time series images. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 2175–2179. [[CrossRef](#)]
13. Sismanidis, P.; Keramitsoglou, I.; Kiranoudis, C.T. Identifying and characterizing the diurnal evolution of urban land surface temperature patterns. In Proceedings of the Urban Remote Sensing Event (JURSE 2017), Dubai, UAE, 6–8 March 2017.
14. Haashemi, S.; Weng, Q.; Darvishi, A.; Alavipanah, S.K. Seasonal variations of the surface urban heat island in a semi-arid city. *Remote Sens.* **2016**, *8*, 352. [[CrossRef](#)]
15. Sismanidis, P.; Bechtel, B.; Keramitsoglou, I.; Kiranoudis, C.T. Mapping the Spatiotemporal Dynamics of Europe’s Land Surface Temperatures. *IEEE Geosci. Remote Sens. Lett.* **2017**, *99*, 1–5. [[CrossRef](#)]
16. Amiri, R.; Weng, Q.; Alimohammadi, A.; Alavipanah, S.K. Spatial-temporal dynamics of land surface temperature in relation to fractional vegetation cover and land use/cover in the Tabriz urban area, Iran. *Remote Sens. Environ.* **2009**, *113*, 2606–2617. [[CrossRef](#)]
17. Xiong, Y.; Huang, S.; Chen, F.; Ye, H.; Wang, C.; Zhu, C. The impacts of rapid urbanization on the thermal environment: A remote sensing study of guangzhou, south china. *Remote Sens.* **2012**, *4*, 2033–2056. [[CrossRef](#)]
18. Wang, S.; Ma, Q.; Ding, H.; Liang, H. Detection of urban expansion and land surface temperature change using multi-temporal landsat images. *Resour. Conserv. Recycl.* **2016**. [[CrossRef](#)]

19. Thomson, D.J. The seasons, global temperature, and precession. *Science* **1995**, *268*, 59–68. [[CrossRef](#)] [[PubMed](#)]
20. Tran, D.X.; Pla, F.; Latorre-Carmona, P.; Myint, S.W.; Caetano, M.; Kieu, H.V. Characterizing the relationship between land use land cover change and land surface temperature. *ISPRS J. Photogramm. Remote Sens.* **2017**, *124*, 119–132. [[CrossRef](#)]
21. Kisilevich, S.; Mansmann, F.; Nanni, M.; Rinzivillo, S. Spatio-temporal clustering. In *Data Mining and Knowledge Discovery Handbook*; Springer: New York, NY, USA, 2009; pp. 855–874.
22. Liao, T.W. Clustering of time series data—A survey. *Pattern Recognit.* **2005**, *38*, 1857–1874. [[CrossRef](#)]
23. Paparrizos, J.; Gravano, L. k-Shape: Efficient and Accurate Clustering of Time Series. In Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data (SIGMOD 2015), Melbourne, Australia, 31 May–4 June 2015; pp. 1855–1870.
24. Aghabozorgi, S.; Shirkhorshidi, A.S.; Wah, T.Y. Time-series clustering—A decade review. *Inf. Syst.* **2015**, *53*, 16–38. [[CrossRef](#)]
25. Jain, A.K. Data Clustering: 50 Years beyond K-means. *Pattern Recognit. Lett.* **2010**, *31*, 651–666. [[CrossRef](#)]
26. Izakian, H.; Pedrycz, W.; Jamal, I. Fuzzy clustering of time series data using dynamic time warping distance. *Eng. Appl. Artif. Intell.* **2015**, *39*, 235–244. [[CrossRef](#)]
27. Thomas, B.; Cavanaugh, J.E. State: Pace discrimination and clustering of atmospheric time series data based on kullback information measures. *Environmetrics* **2007**, *19*, 103–121. [[CrossRef](#)]
28. Ji, M.; Xie, F.; Ping, Y. A dynamic fuzzy cluster algorithm for time series. *Abstr. Appl. Anal.* **2013**, *2013*, 1–7. [[CrossRef](#)]
29. Gorji Sefidmazgi, M.; Sayemuzzaman, M.; Homaifar, A.; Jha, M.K.; Liess, S. Trend analysis using non-stationary time series clustering based on the finite element method. *Nonlinear Process. Geophys.* **2014**, *21*, 605–615. [[CrossRef](#)]
30. Hallac, D.; Vare, S.; Boyd, S.; Leskovec, J. Toeplitz inverse covariance-based clustering of multivariate time series data. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, 13–17 August 2017; pp. 215–223.
31. Wang, J.; Zhan, Q.; Guo, H. The Morphology, Dynamics and Potential Hotspots of Land Surface Temperature at a Local Scale in Urban Areas. *Remote Sens.* **2016**, *8*, 18. [[CrossRef](#)]
32. Schwarz, N.; Schlink, U.; Franck, U.; Großmann, K. Relationship of land surface and air temperatures and its implications for quantifying urban heat island indicators—An application for the city of leipzig (Germany). *Ecol. Indic.* **2012**, *18*, 693–704. [[CrossRef](#)]
33. Goodchild, M.F. A spatial analytical perspective on geographical information systems. *Int. J. Geogr. Inf. Syst.* **1987**, *1*, 327–334. [[CrossRef](#)]
34. Goodchild, M.F. Integrating GIS and remote sensing for vegetation analysis and modeling: Methodological issues. *J. Veg. Sci.* **1994**, *5*, 615–626. [[CrossRef](#)]
35. Streutker, D.R. A remote sensing study of the urban heat island of houston, Texas. *Int. J. Remote Sens.* **2002**, *23*, 2595–2608. [[CrossRef](#)]
36. Streutker, D.R. Satellite-measured growth of the urban heat island of houston, Texas. *Remote Sens. Environ.* **2003**, *85*, 282–289. [[CrossRef](#)]
37. Hung, T.; Uchihama, D.; Ochi, S.; Yasuoka, Y. Assessment with satellite data of the Urban Heat Island effects in Asian mega cities. *Int. J. Appl. Earth Obs.* **2006**, *8*, 34–48. [[CrossRef](#)]
38. Keramitsoglou, I.; Kiranoudis, C.T.; Ceriola, G.; Weng, Q.; Rajasekar, U. Identification and analysis of urban surface temperature patterns in Greater Athens, Greece, using MODIS imagery. *Remote Sens. Environ.* **2011**, *115*, 3080–3090. [[CrossRef](#)]
39. Rajasekar, U.; Weng, Q. Urban heat island monitoring and analysis using a non-parametric model: A case study of Indianapolis. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 86–96. [[CrossRef](#)]
40. Crosson, W.L.; Al-Hamdan, M.Z.; Hemmings, S.N.J.; Wade, G.M. A daily merged modis aqua–terra land surface temperature data set for the conterminous United States. *Remote Sens. Environ.* **2012**, *119*, 315–324. [[CrossRef](#)]
41. Jin, M. Interpolation of surface radiative temperature measured from polar orbiting satellites to a diurnal cycle: 2. Cloudy-pixel treatment. *J. Geophys. Res.* **2000**, *105*, 4061–4076. [[CrossRef](#)]

42. Shen, H.; Huang, L.; Zhang, L.; Wu, P.; Zeng, C. Long-term and fine-scale satellite monitoring of the urban heat island effect by the fusion of multi-temporal and multi-sensor remote sensed data: A 26-year case study of the city of wuhan in china. *Remote Sens. Environ.* **2016**, *172*, 109–125. [[CrossRef](#)]
43. Wan, Z. *Collection-5 Modis Land Surface Temperature Products Users' Guide*. Institute for Computational Earth System Science (ICESS); University of California, Santa Barbara: Santa Barbara, CA, USA, 2007.
44. Wan, Z.; Dozier, J. A generalized split-window algorithm for retrieving land-surface temperature from space. *IEEE Trans. Geosci. Remote Sens.* **1996**, *34*, 892–905.
45. Wan, Z. New refinements and validation of the MODIS land-surface temperature/emissivity products. *Remote Sens. Environ.* **2008**, *112*, 59–74. [[CrossRef](#)]
46. Wan, Z. New refinements and validation of the collection-6 modis land-surface temperature/emissivity product. *Remote Sens. Environ.* **2014**, *140*, 36–45. [[CrossRef](#)]
47. Bonilla, E.V.; Chai, K.M.; Williams, C. Multi-task gaussian process prediction. In *Advances in Neural Information Processing Systems*; Donnelley & Sons: San Francisco, CA, USA, 2008.
48. Ackerman, S.A.; Strabala, K.I.; Menzel, W.P.; Frey, R.A.; Moeller, C.C.; Gumley, L.E. Discriminating clear sky from clouds with MODIS. *J. Geophys. Res.* **1998**, *103*, 32141–32157. [[CrossRef](#)]
49. Bonde, U.; Badrinarayanan, V.; Cipolla, R. *Multi Scale Shape Index for 3D Object Recognition*; Springer: Berlin/Heidelberg, Germany, 2013.
50. Koenderink, J.J.; van Doorn, A.J. Surface shape and curvature scales. *Image Vis. Comput.* **1992**, *10*, 557–564. [[CrossRef](#)]
51. Lindeberg, T. Feature detection with automatic scale selection. *Int. J. Comput. Vis.* **1998**, *30*, 79–116. [[CrossRef](#)]
52. Josien, K.; Liao, T.W. Simultaneous grouping of parts and machines with an integrated fuzzy clustering method. *Fuzzy Sets Syst.* **2003**, *126*, 1–21. [[CrossRef](#)]
53. Wang, J.; Ouyang, W. Attenuating the surface Urban Heat Island within the Local Thermal Zones through land surface modification. *J. Environ. Manag.* **2017**, *187*, 239–252. [[CrossRef](#)] [[PubMed](#)]
54. Sakoe, H.; Chiba, S. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Trans. Acoust. Speech Signal* **1978**, *26*, 43–49. [[CrossRef](#)]
55. Petitjean, F.; Ketterlin, A.; Gancarski, P. A global averaging method for dynamic time warping, with applications to clustering. *Pattern Recognit.* **2011**, *44*, 678–693. [[CrossRef](#)]
56. Keogh, E.; Kasetty, S. On the need for time series data mining benchmarks: A survey and empirical demonstration. *Data Min. Knowl. Discov.* **2003**, *7*, 349–371. [[CrossRef](#)]
57. Ding, H.; Trajcevski, G.; Scheuermann, P.; Wang, X.; Keogh, E. Querying and mining of time series data: Experimental comparison of representations and distance measures. In Proceedings of the 34th International Conference on Very Large Databases (VLDB), Auckland, New Zealand, 24–30 August 2008; pp. 1542–1552.
58. Wang, X.; Mueen, A.; Ding, H.; Trajcevski, G.; Scheuermann, P.; Keogh, E. Experimental comparison of representation methods and distance measures for time series data. *Data Min. Knowl. Discov.* **2013**, *26*, 275–309. [[CrossRef](#)]
59. Faloutsos, C.; Ranganathan, M.; Manolopoulos, Y. Fast subsequence matching in time-series databases. In Proceedings of the 1994 ACM SIGMOD International Conference on Management of Data (SIGMOD 1994), Minneapolis, Minnesota, 24–27 May 1994; pp. 419–429.
60. Ulanova, L.; Begum, N.; Keogh, E. Scalable Clustering of Time Series with U-Shapelets. In Proceedings of the 2015 SIAM International Conference on Data Mining, Vancouver, British Columbia, BC, Canada, 30 April–2 May 2015.
61. Vlachos, M.; Lin, J.; Keogh, E.; Gunopulos, D. A Wavelet-Based Anytime Algorithm for K-Means Clustering of Time Series. In Proceedings of the Workshop on Clustering High Dimensionality Data and Its Applications, San Francisco, CA, USA, 1–3 May 2003; pp. 23–30.
62. Weng, Q. A remote sensing-GIS evaluation of urban expansion and its impact on surface temperature in the Zhujiang Delta, China. *Int. J. Remote Sens.* **2001**, *22*, 1999–2014. [[CrossRef](#)]
63. Seto, K.C.; Golden, J.S.; Alberti, M. Sustainability in an urbanizing planet. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 8935–8938. [[CrossRef](#)] [[PubMed](#)]
64. Wu, J. Landscape sustainability science: Ecosystem services and human well-being in changing landscapes. *Landsc. Ecol.* **2013**, *28*, 999–1023. [[CrossRef](#)]

65. Stewart, I.; Oke, T. Newly developed “thermal climate zones” for defining and measuring urban heat island magnitude in the canopy layer. In Proceedings of the Eighth Symposium on Urban Environment, Phoenix, AZ, USA, 11–15 January 2009.
66. Oke, T.R. The energetic basis of the urban heat island. *Q. J. R. Meteorol. Soc.* **1982**, *108*, 1–24. [[CrossRef](#)]
67. Fischer, M.M.; Nijkamp, P. *Handbook of Regional Science*; Springer: Berlin/Heidelberg, Germany, 2014.
68. Wang, N.Y.; Chen, S.M. Temperature prediction and taifex forecasting based on automatic clustering techniques and two-factors high-order fuzzy time series. *Expert Syst. Appl.* **2009**, *36*, 2143–2154. [[CrossRef](#)]
69. Askari, S.; Montazerin, N. A high-order multi-variable fuzzy time series forecasting algorithm based on fuzzy clustering. *Expert Syst. Appl.* **2015**, *42*, 2121–2135. [[CrossRef](#)]
70. Ghofrani, M.; Carson, D.; Ghayekhloo, M. Hybrid clustering-time series-bayesian neural network short-term load forecasting method. In Proceedings of the North American Power Symposium (NAPS 2016), Denver, CO, USA, 18–20 September 2016.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).