ICIC Express Letters
Part B: Applications
Volume **10**, Number **10**, October **2019**

ICIC International ©2019 ISSN 2185-2766
pp. 929–936

DEVELOPING A TIME SERIES CLUSTERING METHOD FOR URBAN AIR POLLUTION IN CHINA

Degang Wang[1] and Wenyan Song[2,*]

[1]School of Control Science and Engineering
Dalian University of Technology
No. 2, Linggong Road, Dalian 116024, P. R. China
wangdg@dlut.edu.cn

[2]School of Economics
Dongbei University of Finance and Economics
No. 217, Jianshan Street, Dalian 116025, P. R. China
*Corresponding author: songwydufe@dufe.edu.cn

Received May 2019; accepted August 2019

Abstract. *Air pollution is an austere and widespread phenomenon around the world, especially in some developing countries. The study of emission classification for cities is helpful to control the pollution. Considering that air pollution is time variable and persistent, time series clustering technique is suitable to deal with this problem. A fuzzy clustering method for time series is proposed and used to discuss Chinese urban air pollution in this paper. Firstly, the weighted dynamic time warping distance is adopted as measurement between time series in fuzzy clustering process. Then with the purpose of satisfactory classifying quality, a new kind of fitness function is designed for parameter training in a modified particle swarm optimization algorithm. Simulation results show that this method is effective for some common data sets of time series. Finally, the proposed approach is applied to empirical research of air pollution time series in China, and some different dynamic characteristics for pollutant discharges have been discovered.*
Keywords: Air pollution, Time series, Fuzzy clustering, Dynamic time warping

1. **Introduction.** Air pollution is a major environmental risk to the entire world. Air pollution causes 1.4 million deaths from stroke, 2.4 million deaths due to heart disease, and 1.8 million deaths due to lung disease and cancer every year [1]. Air pollutants mainly include particulate matter (PM), ozone (O_3), nitrogen dioxide (NO_2) and sulphur dioxide (SO_2) [2,3]. Particulate matters are inhalable particles composed of ammonia, black carbon, nitrates, sodium chloride, sulphate, mineral dust and water. Particles with a diameter of less than 10 microns (PM10), including fine particles less than 2.5 microns (PM2.5), can penetrate people's lungs and enter bloodstream, so they can cause harm to the respiratory and cardiovascular systems predominantly [4,5].

Plenty of academic groups, environmental protection organizations and government agencies focused on air pollution problem over the past decades increasingly. Among existing studies, atmospheric pollutants forecasting [6-10] and air quality assessment [11,12] are the most concentrated aspects. Based on forecasting and assessment results, some targeted air pollution remedy methods have been made. In recent years, along with the development of real-time pollution monitoring techniques, the changing process of air pollution receives more and more attention. Accordingly, some time series classification or clustering approaches have been used to deal with the classified governance for pollutant emission problems.

DOI: 10.24507/icicelb.10.10.929

In fact, a lot of researches have been conducted on time series clustering, and different time series clustering algorithms have been applied in different domains. See [13-15] for reviews of relevant approaches. Among all methods applied to clustering time series, shape-based approaches for whole time series clustering are the mainly used ones, in which the dynamic time warping (DTW) distance is the most popular way to measure difference or similarity between time series, especially between time series with different length. In [16], DTW distance is used to design a hybrid fuzzy clustering method based on Fuzzy C-Means (FCM) and Fuzzy C-Medoids (FCMdd) methods. In [17], a weighted DTW (WDTW) method for time series classification is proposed, which is a combination of the original DTW and the derivative DTW (DDTW). Existing researches have shown that WDTW can improve clustering performances of many time series data sets [18,19]. Inspired by these, we develop a fuzzy clustering method based on the WDTW distance to deal with multivariate air pollution time series for cities of China in this study.

Considering that on the one hand many researchers have found that FCMdd is sensitive to initialization and may generate results in local optima, and on the other hand the parameter of the WDTW distance has a significant impact on the similarity measures for time series, in this paper FCMdd method is used to obtain clustering centers in the first phase, and then taking these corresponding centers as initialized particles for the particle swarm optimization (PSO) algorithm [20], a modified fitness function based on the WDTW distance is designed to determine the weight parameter and the cluster prototypes in the subsequent iteration process.

The remaining sections of this paper are organized as follows. In Section 2, the basic principles of FCM and FCMdd methods and the WDTW distance function are introduced. Section 3 describes the proposed optimization algorithm of time series clustering. In Section 4, experimental results of some datasets and the urban air pollution time series in China are presented. Some conclusions and future works are summarized in Section 5.

2. Preliminaries.

2.1. Fuzzy clustering techniques. Fuzzy C-Means and Fuzzy C-Medoids methods are two typical fuzzy clustering algorithms, which are widely used at present. They can both give membership degree of each sample to each cluster. Taking time series as data set to be classified, FCM and FCMdd clustering processes can be realized as follows.

Let $X = \{x_1, \ldots, x_N\}$ be a set of N time series, in which all x_k $(k = 1, \ldots, N)$ have the same length of sequence and the same dimensionality at every time point. The results of clustering for X correspond to a set of c cluster centers $V = \{v_1, \ldots, v_c\}$ and a partition matrix $U = [u_{ik}]_{c\times N}$, where $u_{ik} \in [0, 1]$, $\sum_{i=1}^{c} u_{ik} = 1$, $k = 1, \ldots, N$, and $0 < \sum_{k=1}^{N} u_{ik} < N$, $i = 1, \ldots, c$. The expression of fuzzy clustering objective function is shown as below:

$$J(U, V) = \sum_{i=1}^{c}\sum_{k=1}^{N} u_{ik}^{m} d^2(v_i, x_k), \tag{1}$$

where $d(\cdot, \cdot)$ is a distance function and m $(m > 1)$ is a fuzzification coefficient. Iterative algorithm starting with randomly initialized partition is always used to minimize $J(U, V)$.

Accordingly, the clustering prototypes and partition matrix can be calculated by

$$v_i = \sum_{k=1}^{N} u_{ik}^{m} x_k \Bigg/ \sum_{k=1}^{N} u_{ik}^{m}, \quad u_{ik} = 1 \Bigg/ \sum_{j=1}^{c} \left(\frac{d(v_i, x_k)}{d(v_j, x_k)}\right)^{2/(m-1)}. \tag{2}$$

While for FCMdd, the prototypes are selected within the time series dataset X. After initializing c cluster centers randomly, the new centers v_i $(i = 1, \ldots, c)$ in the next iterative step are determined in the following form:

$$v_i = x_l, \quad l = \arg\min_{1\le j\le N}\left\{\sum_{k=1}^{N} u_{ik}^m d^2(x_j, x_k)\right\}. \tag{3}$$

2.2. **Weighted dynamic time warping.** Dynamic time warping is an efficient distance measure for time series clustering and classification. The minimum path between time series can be decided by allowing a non-linear mapping of one sequence to another.

Assume two time series: $x = \{x(1), \ldots, x(m)\}$ and $y = \{y(1), \ldots, y(n)\}$. Construct an $m \times n$ distance matrix $D = [d_{ij}]_{m\times n}$, where $d_{ij} = d(x(i), y(j))$ and $d(\cdot,\cdot)$ is a distance function. Then we create a warping path $W = \{w_1, \ldots, w_K\}$, which should subject to three conditions:

(i) boundary conditions: $w_1 = d(x(1), y(1))$ and $w_K = d(x(m), y(n))$;

(ii) continuity conditions: for $w_l = d(x(i), y(j))$ and $w_{l+1} = d(x(i'), y(j'))$, $i' - i \le 1$ and $j' - j \le 1$;

(iii) monotonicity conditions: $i' - i \ge 0$ and $j' - j \ge 0$.

Then the path which minimizes the warping cost corresponds to the value of the DTW distance:

$$DTW(x, y) = \min_{W}\left\{\sum_{k=1}^{K} w_k\right\}.$$

In practice, the DTW distance can be calculated through a cumulative distance matrix Γ, in which $\Gamma(i, j) = d(x(i), y(j)) + \min\{\Gamma(i-1, j), \Gamma(i, j-1), \Gamma(i-1, j-1)\}$ and $\Gamma(0, 0) = 0$. Hence, $DTW(x, y) = \Gamma(m, n)$.

Considering the dynamic variation property of time series, derivative dynamic time warping and weighted version of DTW are proposed consecutively [17].

Define the first order difference sequence x' of time series x as $x'(i) = x(i+1) - x(i)$, $i = 1, \ldots, m-1$, and y' of time series y as $y'(j) = y(j+1) - y(j)$, $j = 1, \ldots, n-1$. Then the DDTW distance of these two time series x and y is the DTW distance between x' and y', i.e.,

$$DDTW(x, y) = DTW(x', y') \tag{4}$$

For DTW and DDTW, a weighted DTW (WDTW) distance can be defined as

$$WDTW(x, y) = (1-\alpha)\cdot DTW(x, y) + \alpha \cdot DDTW(x, y) \tag{5}$$

where $\alpha \in [0, 1]$ is a parameter.

3. **Proposed Clustering Algorithm Based on WDTW.** In this section, we design a modified PSO algorithm to solve fuzzy clustering of multivariate time series based on the weighted DTW distance function. For time series dataset $X = \{x_1, \ldots, x_N\}$, we use the WDTW distance in FCMdd method to create initial clustering centers as population of particles, and meanwhile generate uniformly distributed random numbers as initial particles for the weight parameter α in WDTW.

Hence, every particle of swarm can be denoted by $\left(X^i, \alpha_i, V_i^X, V_i^\alpha, P_{best_i}^X, P_{best_i}^\alpha\right)$, where X^i, α_i, V_i^X, V_i^α are the positions and velocities, and $P_{best_i}^X$, $P_{best_i}^\alpha$ are the personal best positions in the i-th particle. In particular, $X^i = (x_1^i, \ldots, x_c^i)$, $V_i^X = \left(v_{i1}^X, \ldots, v_{ic}^X\right)$, $P_{best_i}^X = \left(pbest_{i1}^X, \ldots, pbest_{ic}^X\right)$, where c is the number of clusters, and $NumP$ is the number of particles, $i = 1, \ldots, NumP$. The best positions of all the particles in current step are $G_{best}^X = \left(gbest_1^X, \ldots, gbest_c^X\right)$ and G_{best}^α.

Then X^i, α_i, V_i^X, V_i^α are updated by the following formulae:

$$\begin{gathered} v_{ij}^X(Iter+1) \\ = v_{ij}^X(Iter) + C_1 rand_1\left(pbest_{ij}^X(Iter) - x_j^i(Iter)\right) + C_2 rand_2\left(gbest_j^X(Iter) - x_j^i(Iter)\right), \\ x_j^i(Iter+1) = x_j^i(Iter) + v_{ij}^X(Iter+1), \quad j = 1, 2, \ldots, c, \\ \alpha_i(Iter+1) = \alpha_i(Iter) + V_i^\alpha(Iter+1), \end{gathered}$$

$$V_i^{\alpha}(Iter+1)$$
$$= V_i^{\alpha}(Iter) + C_1 rand_1(P_{best_i}^{\alpha}(Iter) - \alpha_i(Iter)) + C_2 rand_2(G_{best}^{\alpha}(Iter) - \alpha_i(Iter)),$$

where $Iter$ and $Iter+1$ are the current and the next iteration numbers, C_1 and C_2 are the acceleration parameters, and $rand_1$ and $rand_2$ are random numbers in the interval $[0,1]$. In reality, the velocities are always set to some search space bounds by presetting $v_{\max}^X$ and $v_{\max}^{\alpha}$, i.e., $\left|v_{ij}^X\right| < v_{\max}^X$ and $\left|V_{\max}^{\alpha}\right| < v_{\max}^{\alpha}$.

In each iteration, a fitness function defined as follows is used to select the personal best and global best particles:

$$\begin{aligned} Fitness_{Internal} &= \frac{1}{N-c}\sum_{k=1}^{N}\sum_{j=1}^{c} d\left(c_j^i, x_k \middle| \alpha_i\right) \\ &= \frac{1}{N-c}\sum_{k=1}^{N}\sum_{j=1}^{c} \left((1-\alpha_i)d_1\left(c_j^i, x_k\right) + \alpha_i d_2\left(c_j^i, x_k\right)\right), \end{aligned} \tag{6}$$

where c_j^i represents the current particle such as x_j^i, $d(\cdot,\cdot)$ is the weighted distance between time series with parameter α_i, and d_1, d_2 are distance functions.

As noted as some existing research, if the length of time series and the size of the dataset are large, it will always cost much computation time of clustering algorithms when the DTW distance is used. To solve this issue, several hierarchical clustering techniques and multistage clustering processes are proposed. In this study, we consider a two stage method to overcome this problem. First, FCMdd method based on the WDTW is used to generate the initial particles for PSO. Then in the course of iterative optimization, a weighted Euclidean distance is adopted in the fitness function (6), where the distances between original series and differential series are mixed together to form a convex combination.

The implementation process of the proposed fuzzy clustering method for time series can be summarized as follows.

Step 1: Input training data set $X = \{x_1, \ldots, x_N\}$. Assign the number of clusters c, parameter m and the number of iterations in the FCMdd process.

Step 2: Initialize clustering centers and the weight parameter α^0 randomly, then use the WDTW as distance function in (1) to (3) to cluster data into c categories by the FCMdd method, and meanwhile record the primary optimal prototypes X^0.

Step 3: Set maximum iterations of PSO algorithm. Take X^0 and α^0 as the initial cluster centers and the weight value in WDTW. Set the number of particles $NumP$, acceleration parameters C_1, C_2, and upper limit values $v_{\max}^X$ and $v_{\max}^{\alpha}$.

Step 4: Based on fitness function Equation (6), use PSO to determine the eventual cluster centers and the weight parameter.

Step 5: Calculate the partition matrix and get the class label for each training series.

In the realization course of the above method, several respects could be investigated further. For instance, in view of that the value of α determines the clustering validation directly, in order to select appropriate α efficiently, an enumeration method can be used to obtain the subinterval for it firstly, and then the eventual value of α can be trained by the proposed PSO algorithm upon the corresponding sub-range. In addition, the fitness function (6) can be replaced by other internal measures for clustering, such as Calinski-Harabasz index and Modified Hubert Γ statistic. Moreover, if we have class labels for objects to be classified, the fitness function can be designed according to some external measure indexes, like purity, and entropy [19]. Besides, when the dimension of time series is high, the computational complexity of clustering will increase. Therefore, some descending dimension method like the principal component analysis (PCA), could be considered to pretreatment of data.

4. **Empirical Analysis.** In this section, some datasets from the UCR Time Series Classification Archive (http://www.cs.ucr.edu/∼eamonn/time_series_data/) are considered to demonstrate the effectiveness of the proposed algorithm. Then monthly air pollution time series of some cities in China (http://www.cnemc.cn/) are clustered and analyzed.

4.1. **UCR datasets.** Since the UCR time series data are labeled time series, some external measurements can be used to verify quality of clustering results. To compare with the literature, two indexes are calculated in the following experiments [19].

$$\text{Purity Index: } \mathit{Purity} = \frac{1}{N}\sum_{i=1}^{c}\max_{j} n_{ij} \tag{7}$$

$$\text{Entropy Index: } \mathit{Entropy} = -\sum_{j=1}^{c}\frac{n_j}{N}\sum_{i=1}^{c}\frac{n_{ij}}{n_j}\log_2\frac{n_{ij}}{n_j} \tag{8}$$

where N is the total amount of the dataset, n_j is the number of elements in the j-th cluster, and n_{ij} is the number of elements of the i-th class in the j-th cluster. It is worth mentioning that for purity and negative entropy indicators, a higher value means a better clustering performance.

Seven datasets are selected, and their main characteristics are summarized in Table 1. The purity and entropy indexes of the proposed method are reported in Table 2. For both FCM and FCMdd techniques, the fuzzification coefficient m is set to 2 and the maximum number of iterations is set to 50. It can be seen from Table 2 that compared with existing parametric dynamic time warping clustering technique, using fuzzy clustering means and PSO algorithm we can get more satisfactory results. Besides, by weighted dynamic time warping distance, the change of time series is fully taken into account, and this method can be used in both binary and multi-class classification problems.

Table 1. Description of characteristics for some UCR datasets

Dataset	Length of time series	Number of time series	Number of classes
ItalyPowerDemand	24	1096	2
TwoLeadECG	82	1162	2
MoteStrain	84	1272	2
CBF	128	930	3
DistalPhalanxOutlineAgeGroup	80	539	3
MedicalImages	99	1141	10
Swedish Leaf	128	1125	15

Table 2. Results of UCR datasets

Dataset	Purity				Entropy*(−1)			
	DD_DTW [19]	DTW	DDTW	WDTW	DD_DTW [19]	DTW	DDTW	WDTW
ItalyPowerDemand	0.578	0.6451	0.6515	**0.6788**	−0.93	−0.6469	−0.6451	**−0.6276**
TwoLeadECG	0.501	0.6179	0.6437	**0.6954**	−1	−0.6648	−0.6395	**−0.6091**
MoteStrain	0.539	0.7869	0.7225	**0.8247**	−0.995	−0.5074	−0.5872	**−0.4516**
CBF	0.673	0.7183	0.6075	**0.8065**	−0.666	−0.6691	−0.8415	**−0.5514**
DistalPhalanxOutlineAgeGroup	0.776	0.7792	0.7792	**0.7829**	−0.799	−0.5637	−0.5634	**−0.5593**
MedicalImages	0.546	0.6152	0.6179	**0.6468**	−1.614	−1.2004	−1.157	**−1.1357**
Swedish Leaf	0.22	0.2667	0.3182	**0.3236**	−3.123	−1.6248	−1.496	**−1.4592**

4.2. **Air pollution time series.** In this paper, monthly urban air quality time series of 74 cities are collected to be clustered. The length of each time series is 43, which include PM2.5, PM10, SO_2, CO, NO_2 and O_3 emissions loaded from December 2013 to June 2017. In addition, the corresponding air quality index (AQI) time series are gathered together. The AQI values are provided according to the Technical Regulation on Ambient Air Quality Index [21], which are used to measure the overall quality of the air. The basic statistics for all the data are listed in Table 3.

In terms of PM2.5, if the mean discharge is considered, then some cities would be classified into the same level. Taking Chengdu and Guiyang as examples, the average emissions of PM2.5 for these two cities are $67.23\mu g/m^3$ and $40.38\mu g/m^3$ respectively. Reference to classification criteria both of them could be categorized into the second level. However, if these two PM2.5 emission time series are compared carefully, different

TABLE 3. Statistical analysis for the urban air pollution data

	PM2.5	PM10	SO_2	CO	NO_2	O_3	AQI
	$(\mu g/m^3)$	$(\mu g/m^3)$	$(\mu g/m^3)$	(mg/m^3)	$(\mu g/m^3)$	$(\mu g/m^3)$	\
Min	9.5	19.5	2.9	0.39	7.2	11.5	25
Max	276.3	390.3	203.6	4.65	100.9	191.7	301
Median	48.8	84.15	18.6	0.95	38.5	88.45	85
Mean	56.59	95.66	25.93	1.09	40.1	90.36	91.06

TABLE 4. Classification results of 74 cities

	Type 1	Type 2
AQI Time Series	Shanghai, Chongqing, Guangzhou, Shenzhen, Hangzhou, Qinhuangdao, Zhangjiakou, Chengde, Hohhot, Dalian, Changchun, Harbin, Nanjing, Wuxi, Changzhou, Suzhou, Nantong, Lianyungang, Haian, Yancheng, Yangzhou, Zhenjiang, Taizhou, Suqian, Ningbo, Wenzhou, Jiaxing, Huzhou, Jinhua, Quzhou, Zhoushan, Taizhou, Lishui, Shaoxing, Fuzhou, Xiamen, Nanchang, Qingdao, Changsha, Zhuhai, Foshan, Jiangmen, Zhaoqing, Huizhou, Dongguan, Zhongshan, Nanning, Haikou, Guiyang, Kunming, Lhasa, Lanzhou, Xining, Yinchuan	Beijing, Baoding, Tianjin, Shijiazhuang, Tangshan, Handan, Xingtai, Cangzhou, Langfang, Hengshui, Taiyuan, Shenyang, Xuzhou, Hefei, Jinan, Zhengzhou, Wuhan, Chengdu, Xi'an, Urumqi
Multivariate Time Series	Shanghai, Guangzhou, Shenzhen, Hangzhou, Nanjing, Wuxi, Changzhou, Suzhou, Nantong, Lianyungang, Haian, Yancheng, Yangzhou, Zhenjiang, Taizhou, Suqian, Ningbo, Wenzhou, Jiaxing, Huzhou, Jinhua, Quzhou, Zhoushan, Taizhou, Lishui, Shaoxing, Hefei, Fuzhou, Xiamen, Nanchang, Qingdao, Wuhan, Changsha, Zhuhai, Foshan, Jiangmen, Zhaoqing, Huizhou, Dongguan, Zhongshan, Nanning, Haikou, Guiyang, Kunming, Lhasa, Lanzhou, Urumqi	Beijing, Baoding, Tianjin, Chongqing, Shijiazhuang, Tangshan, Qinhuangdao, Zhangjiakou, Handan, Xingtai, Chengde, Cangzhou, Langfang, Hengshui, Taiyuan, Hohhot, Shenyang, Dalian, Changchun, Harbin, Xuzhou, Jinan, Zhengzhou, Chengdu, Xi'an, Xining, Yinchuan

trajectories can be discovered. Accordingly, by proposed time series clustering algorithm Chengdu could be classified to the third class and Guiyang could be classified to the second class. In fact, along with deepening of research it has become increasingly clear that the pollution sources and patterns are indeed different between these two regions.

Besides, according to the Technical Regulation on Ambient Air Quality Index, based on mean values of the AQI indexes we cluster 74 cities into 4 levels. Only Baoding is classified into level 4 and Haikou is classified into level 1, and all the other cities are classified into level 2 and level 3. Therefrom, by the designed WDTW algorithm we cluster all the cities into 2 clusters for analysis. The classification results based on both AQI time series and original six kinds of emissions vector time series are shown in Table 4. And it can be seen from it that the partitions are not exactly the same, which could be helpful for policy makers, when they take comprehensive consideration of both the overall conditions and the change situations of various pollution emissions to make pertinent treatment measures.

5. **Conclusions.** In this paper, a time series clustering method based on weighted DTW is proposed. By using fuzzy clustering and PSO algorithms, the purity index and the entropy index of classification results can be effectively improved. Therefrom, urban air pollution time series are graded and analyzed, and the dynamic changes of these pollutants are found accordingly, which can be further utilized to generate more effective air pollution control strategy. In future, some more efficient distances of time series will be investigated, and the structure and interaction of different emissions will be considered deeply.

Acknowledgment. This work is partially supported by the National Natural Science Foundation of China (71571035, 61773088). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] *http://www.who.int/airpollution.*

[2] Health aspects of air pollution with particulate matter, ozone and nitrogen dioxide, Report on a WHO working group, *World Health Organization*, vol.6, p.30, 2003.

[3] D. M. Stieb, S. Judek and R. T. Burnett, Meta-analysis of time-series studies of air pollution and mortality: Effects of gases and particles and the influence of cause of death, age, and season, *Air Repair*, vol.52, no.4, pp.470-484, 2002.

[4] R. M. Harrison and J. Yin, Particulate matter in the atmosphere: which particle properties are important for its effects on health?, *Science of the Total Environment*, vol.249, no.1, pp.85-101, 2000.

[5] J. O. Anderson, J. G. Thundiyil and A. Stolbach, Clearing the air: A review of the effects of particulate matter air pollution on human health, *Journal of Medical Toxicology*, vol.8, no.2, pp.166-175, 2012.

[6] A. Kurt, B. Gulbagci, F. Karaca and O. Alagha, An online air pollution forecasting system using neural networks, *Environment International*, vol.34, no.5, pp.592-598, 2008.

[7] I. Kyriakidis, K. D. Karatzas and G. Papadourakis, Using preprocessing techniques in air quality forecasting with artificial neural networks, *Springer Berlin Heidelberg*, pp.357-372, 2009.

[8] Y. Zhang, M. Bocquet, V. Mallet, C. Seigneur and A. Baklanov, Real-time air quality forecasting, part I: History, techniques, and current status, *Atmospheric Environment*, vol.60, pp.632-655, 2012.

[9] Y. Zhang, M. Bocquet, V. Mallet, C. Seigneur and A. Baklanov, Real-time air quality forecasting, part II: State of the science, current research needs, and future prospects, *Atmospheric Environment*, vol.60, pp.656-676, 2012.

[10] D. Domańska and M. Wojtylak, Explorative forecasting of air pollution, *Atmospheric Environment*, vol.92, pp.19-30, 2014.

[11] F. V. Yee, P. Simon, K. J. Jiří and L. T. Chew, A review on air emissions assessment: Transportation, *Journal of Cleaner Production*, vol.194, no.1, pp.673-684, 2018.

[12] Y. Li, D. Guan, S. Tao, X. Wang and K. He, A review of air pollution impact on subjective well-being: Survey versus visual psychophysics, *Journal of Cleaner Production*, vol.184, pp.959-968, 2018.

[13] W. T. Liao, Clustering of time series data: A survey, *Pattern Recognition*, vol.38, no.11, pp.1857-1874, 2005.

[14] A. Saeed, S. S. Ali and W. T. Ying, Time-series clustering – A decade review, *Information Systems*, vol.53, pp.16-38, 2015.

[15] T. Rajesh, Y. S. Devi and K. V. Rao, Hybrid clustering algorithm for time series data – A literature survey, *International Conference on Big Data Analytics & Computation Intelligence*, pp.343-347, 2017.

[16] I. Hesam, P. Witold and J. Iqbal, Fuzzy clustering of time series data using dynamic time warping distance, *Engineering Applications of Artificial Intelligence*, vol.39, pp.235-244, 2015.

[17] J. Young-Seon, J. K. Myong and A. O. Olufemi, Weighted dynamic time warping for time series classification, *Pattern Recognition*, vol.44, no.9, pp.2231-2240, 2011.

[18] G. Tomasz and Ł. Maciej, Multivariate time series classification with parametric derivative dynamic time warping, *Expert Systems with Applications*, vol.42, no.5, pp.2305-2312, 2015.

[19] Ł. Maciej, Hierarchical clustering of time series data with parametric derivative dynamic time warping, *Expert Systems with Applications*, vol.62, pp.116-130, 2016.

[20] J. Kennedy and R. Eberhart, Particle swarm optimization, *Proc. of IEEE International Conference on Neural Networks*, vol.4, pp.1942-1948, 1995.

[21] State Environmental Protection Agency of China, *Technical Regulations on Ambient Air Quality Index (in Chinese)*, HJ633-2012, Beijing, 2012.