Urban heat island estimation from improved selection of urban and rural stations by DTW algorithm

Yonghong Hu (yonghonghu@ceode.ac.cn)
Chinese Academy of Sciences https://orcid.org/0000-0002-1520-4386
Gensuo Jia
Chinese Academy of Sciences
Jinlong Ai
Yiyang Vocational and Technical College
Yong Zhang
China Meteorological Administration
Meiting Hou
China Meteorological Administration
Yapeng Li
Space Star Technology Co.

Research Article

Keywords: urban heat island, air temperature sequences, station selection, urban and rural stations

Posted Date: March 8th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-164196/v1

Version of Record: A version of this preprint was published at Theoretical and Applied Climatology on August 13th, 2021. See the published version at https://doi.org/10.1007/s00704-021-03749-z.

Urban heat island estimation from improved selection of urban and rural stations by DTW algorithm

Yonghong Hu[1*], Gensuo Jia[2], Jinlong Ai[3], Yong Zhang[4*], Meiting Hou[5], Yapeng Li[6]

1 Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China

2 Key Laboratory of Regional Climate-Environment for East Asia, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China

3 Department of Biology and Information Engineering, Yiyang Vocational and Technical College, Yiyang 413049, China

4 National Satellite Meteorological Center, China Meteorological Administration, Beijing 100081, China

5 China Meteorological Administration Training Centre, China Meteorological Administration, Beijing 100081, China

6 Space Star Technology Co.,Ltd, Beijing 100095, China

Corresponding Author: Yonghong Hu (E-mail: huyh01@radi.ac.cn

Phone: +86 10 82178126)

Abstract: Typical urban and rural temperature records are essential for the estimation and comparison of urban heat island effects in different regions, and the key issues are how to identify the typical urban and rural stations. This study tried to analyze the similarity of air temperature sequences by using dynamic time warping algorithm (DTW) to improve the selection of typical stations. We examined the similarity of temperature sequences of 20 stations in Beijing and validated by remote sensing, and the results indicated that DTW algorithm could identify the difference of temperature sequence, and clearly divide them into different groups according to their probability distribution information. The analysis for station pairs with high similarity could provide appropriate classification for typical urban stations (FT, SY, HD, TZ, CY, CP, MTG, BJ, SJS, DX, FS) and typical rural stations (ZT, SDZ, XYL) in Beijing. We also found that some traditional rural stations can't represent temperature variation in rural surface because of their surrounding environments highly modified by urbanization process in last decades, and they may underestimate the urban climate effect by 1.24℃. DTW algorithm is simple in analysis and application for temperature sequences, and has good potentials in improving urban heat island estimation in regional or global scale by selecting more appropriate temperature records.

Key words: urban heat island, air temperature sequences, station selection, urban and rural stations

1 Introduction

Urban heat island (UHI) is quantified by the temperature difference between urban and rural area with temperature records collected from meteorological stations or satellite platforms (Camilloni and Barros 1997; Stewart and Oke 2012; Voogt and Oke 2003). A pair of stations in urban and rural region are usually used to delineate the situation of surface temperature variation and estimate the magnitude of canopy urban heat island, and the accuracy of estimated UHI intensity highly depended on whether urban/rural station could represent "real" temperature of urban and rural surface (Camilloni and Barros 1997; Stewart and Oke 2012; Voogt and Oke 2003).

UHI estimations were also different with each other due to urban/rural sites identification from different data sources, such as population, night light images and land use/cover datasets (He et al. 2007; Mohsin and Gough 2012; Parker 2010), which may cause difficulties for UHI comparison across different regions and more universal results needed by urban climate research (Oke 2006). Single pair of stations for UHI estimation may be quite reasonable for those cities with relative stable city limits and rural environment, which didn't experience rapid urban expansion. However, urban development induced by city economy or population increase usually resulted in urban expansion and urban morphology change, further modified land properties of urban surroundings (Hu et al. 2015; Jia et al. 2015; Normile 2008). Uncertainties of local climate estimations would be introduced by inhomogeneous air temperature records by such urbanization process (Stewart and Oke 2012; Wu and Yang 2013; Yan et al. 2010).

It is important to identify the bias of UHI estimation induced by the difference of representative sites selection (Mohsin and Gough 2012; Stewart and Oke 2012). As more and more field meteorological instruments equipped at urban and its surroundings, averaged temperature records from multiple stations in typical land cover have already been used to quantify UHI (He et al. 2007; Kim and Baik 2005; Tan et al. 2010). The temperature from each meteorological station could only represent its limited footprint of surface energy variation (He et al. 2013). Therefore, local climate zone was proposed to identify more homogeneous surface temperature record in urban area for comparing with other rural sites to derive more reasonable UHI intensities and historical climate analysis (Stewart and Oke 2012). More physical explanation of UHI intensities could be derived from the comparison between the common surface and exposure characteristics of the selected field sites in specific cities. However, it is not convenient to compare UHI effects across different regions based on UHI intensities derived differently. It is still a challenge to derive the more representative UHI results by using air temperature data from limited number of observation sites.

Temperature time series provide the possibility of insights for climate trends

variation, especially the evolution of urban climate. Land surface temperature from satellite images also provide detailed spatial characterization of temperature variation in difference spatial scales, which has also been widely used in surface UHIs estimation (Voogt and Oke 2003; Weng et al. 2004). They are highly correlated to air temperature variation and can estimate air temperature for those regions with sparse meteorological observation (Vancutsem et al. 2010; Zhu et al. 2013). Therefore, satellite-based land surface temperature (LST) is a good reference for examining surface thermal conditions of local climate zone, further evaluating whether appropriate observation sites included in urban climate analysis. Air temperature and LST records could be used to quantify UHI happened in urban canopy and urban surface, respectively (Hu et al. 2019). Averaged temperature are usually used to quantify UHI intensity to eliminate the possible uncertainties induced by surface heterogeneity of meteorological stations (Jin 2012). Local climate zones have defined the appropriate region for identifying UHI intensity to compare with each other in different cities (Stewart and Oke 2012). In fact, there are similar fluctuations of temperature records in local climate zones, and they are helpful in identifying UHI effect with appropriate observation records. Rapid urbanization has resulted in many meteorological stations suffering from the invasion of man-made infrastructures, then inhomogeneous signals will be introduced to temperature records (Yan et al. 2010). Subsequently, urban climate effect will be underestimated.

Dynamic Time Warping (DTW) is an algorithm for measuring similarity between two temporal sequences which may vary in speed, and it has potential in distinguishing the temporal similarity of temperature series from local climate zones. Therefore, this study tried to use DTW algorithm to improve the accuracy of UHI estimation by selecting typical urban and rural station combinations of 20 stations in Beijing. The algorithm was also used to verify the inhomogeneous signals among air temperature time series for further identifying which one is the best combination of observation sites in estimating UHI intensity. Then, we analyzed the possible improvement in UHI estimation by using the combination analysis from specific multiple-group temperature records.

2 Methods

2.1 Dynamic time warping algorithm

As its advance in automatic recognition and partial shape matching application for sequences of different lengths, DTW algorithm has been widely applied to analyze the consistency of temporal sequences (Keogh and Pazzani 1999; Wang et al. 2013). We can construct an $n \times m$ matrix for temperature sequences of A and B with length n and m respectively:

$$A = a_1, a_2, a_3, ..., a_i, ..., a_n \quad (1)$$

$$B = b_1, b_2, b_3, ..., b_j, ..., b_m \quad (2)$$

where the (i_{th}, j_{th}) element of this matrix contains the distance $d(i,j)$ between the two points A_i and B_j. In order to derive the optional distance, a warping path W,

$$\mathrm{W} = \varpi 1, \varpi 2 ... \varpi \mathrm{k} \qquad \max(n, m) \le k \le n + m - 1 \quad (3)$$

, can be conducted for sequences A and B. In order to find the best match or alignment between these two sequences, we need to find a path in distance matrix which minimizes the total distance between them. The cumulative distance, $\gamma(i, j)$, is the minimum of the sum of the distances between the individual elements on the path divided by the sum of the weighting function, which can be calculated as followed:

$$\gamma(i, j) = d(A_i, B_j) + \min\{\gamma(i-1, j-1), \gamma(i-1, j), \gamma(i, j-1)\} \quad (4)$$

Finally, the cumulative distance for each point could be considered as a factor to quantify the consistency between sequence A and B, which also could be used to examine the difference between temperature time series. The cumulative distance derived from DTW algorithm are used to compare the similarity of temperature sequences (STS), further to identify temperature sequences from which station pairs are better for estimating UHI intensities. Temperature sequences with smaller distance from DTW algorithm have better consistency in temperature variation. Therefore, the statistical characterization for STS is used here to distinguish station combinations, and it provides an efficient way in understanding the relationship of different climate record series. Here, DTW algorithm is programmed based on Python language to

conduct the similarity analysis of temperature sequences.

2.2 Meteorological stations selection for estimating UHI intensity

This study collected air temperature sequences from 1987-2016 of 20 meteorological stations in Beijing, which were released by China Meteorological Administrations (Figure 1). Beijing has experienced rapid urban expansion during last decades with its urbanization rate greater than 85% in 2019. Urbanization has already greatly altered the local climate. A more extensive observation network was established in Beijing to monitor urban environment and urban climate. The surroundings of those stations invaded by urban land will cause the rise of air temperature records in addition to regional climate warming (He et al. 2013), especially in rural stations. Therefore, it is hard to select the suitable urban or rural stations with stable local climate zone due to the inhomogeneous information from station relocation or instrument replacement (Yan et al. 2010). It is appropriate to evaluate the applicability of DTW algorithm across urban-rural transect in Beijing because of the abundance of observation data from meteorological stations of different kinds.

We divided 20 stations into two groups urban and rural stations according to their location and land use information from National land use/cover datasets of China in 2015 (Hu et al. 2015) (Table 1). Urban and rural stations have been usually defined by city population and their location. Remote sensing products, land use or night light information, was considered as the optimized way to identify station types. However, there are still difficulties or uncertainties in examining the representativeness of temporal change of temperature sequences by using land use information from the limited area of station surroundings, such as 1km×1km. Here, the similarity of of air temperature sequences derived from DTW algorithm is helpful to validate station types. Theoretically, stable urban or rural stations has the similar air temperature fluctuations or warming signals induced by climate background (He et al. 2013), which may cause lower DTW distances of a pair of stations. We calculated all the DTW distances for 20 stations to create matrix to examine their similarity of temperature series. Then, the cumulative probability distribution analysis for those

179 distance information was used to analyze their agglomeration. The stations with
180 higher agglomeration and lower DTW distance was examined and classified as typical
181 urban and rural sites, which was considered as the most appropriate sites to examine
182 UHI intensity.

183 In order to further validate the UHI results from selected sites, the composite
184 land surface temperature products (MYD11A2 and MOD11A2) from 2003-2016 were
185 selected to estimate UHI intensity. Surface temperature from 4 times observation by
186 Terra/Aqua-MODIS (local time 10:30 AM and 22:30 PM from Terra, 13:30 PM and
187 01:30 AM from Aqua) illustrates surface radiation energy variation influenced by the
188 daily insolation, which is correlated to daily mean air temperature observed in
189 meteorological station. UHI sequences could be also derived from the land surface
190 temperature difference between urban and rural surface. We examine the similarity of
191 UHI sequences from satellite-based land surface temperature and air temperature.
192 Then, we evaluate their difference of UHI intensity to further analyze which station
193 combinations are better for urban climate effect estimation.

194 **3 Results and discussion**

195 3.1 The similarity analysis of air temperature sequences

196 According to the station location and land use information from remote sensing
197 images, we divided the 20 stations into two groups: urban stations (PG, CP, SY, MTG,
198 SJS, HD, CY, TZ, BJ, FT, DX, FS) and rural stations (THK, SDZ, FYD, YQ, HR, MY,
199 ZT, XYL) (Table 1). The 20 stations exhibit clear intra-group differences in local
200 environment, such as altitude and surrounding land cover type. Some stations were
201 relocated due to the rapid urbanization, such as about 2 times relocation happened in
202 SY station, MTG station, SJS station, HD station, TZ station and FS station (Table 1).
203 Therefore, we need to carefully identify whether they are suitable for the evaluation of
204 urban climate effects. The distance calculated by DTW algorithm are quite different
205 with the values ranged from 95.6-1560.1 according to the similarity analysis of air
206 temperature sequences of 20 stations from 1987-2016 in Beijing (Figure 2). As a
207 lower distance appeared in temporal sequences of higher similarity, STS among the
208 same type of stations should be higher than that of different kinds of stations.

However, 32% of station pairs among rural stations has lower STS with the DTW distance greater than 800, while no urban station pairs could reach to 800, especially larger difference appeared in those rural stations with higher elevation, such as FYD station (Table1 and Figure 2). This phenomenon indicated that rural stations has larger differences with each other, and we should pay more attention to the selection of typical rural station for urban climate analysis.

According to the cumulative probability distribution analysis for STS results, lower values of the distance (less than 280) are concentrated at the left end of the horizontal axis, accounting for about 33% of all STS results (Figure 3), which indicates that local environment difference of each station would greatly alter temperature variation. We further analyzed the detailed pattern about STS results influenced by urbanization processes in last decades (Hu et al. 2015). Table 2 provide the detailed pattern about STS results among urban and rural stations with DTW distance less than 280 to validate whether it is suitable for typical stations identification by spatial coverage derived from remote sensing images. As expected, some traditional rural stations cannot capture the real rural temperature trend because of their high similarity to urban stations, such as DTW distance by 167.3 and 216.7 for MY station versus PG station and HR station versus FS station, respectively. By contrast, traditional urban stations (BJ, CY, HD et al.,) still have high similarity with each other, even some of them experienced relocation during the study period. This phenomenon proved that stations in urban area have relative stable environment of local climate zone with large area impervious surface, while rapid urbanization mainly modified natural surroundings of rural stations. We also found that some rural stations at the remote area of the municipality can't be considered as typical rural station, such as the lower similarity of FYD station with all other stations with the averaged distance by 1353.9±171.9. FYD station is located at the top of Foyeding mountain with an altitude of 1224.7m, and its temperature variation more closed to air temperature trend from the radiosonde profile at this level.

3.2 Identification of typical urban and rural stations by DTW algorithm

Meteorological stations may be influenced by urbanization with stations'

surroundings invaded by city infrastructures, such as roads, factories and residential part of city, especially in developing countries(Jia et al. 2015). The uncertainties of temperature variation are usually from two aspects: 1) warming signals induced by urbanization process(Jia et al. 2015), and 2) the fluctuation or unhomogeneity of air temperature sequences resulted by station relocation(Yan et al. 2010). As its ability in analyzing the temporal similarity of data sequences, DTW algorithm is also useful for detecting unhomogeneous signals of air temperature sequences, further selecting typical stations in urban climate effect evaluation. According to the location, the numbers of station relocation, the altitude and land cover of site surroundings, we selected two stable sites, FT station and SDZ stations, as the referenced stations for typical urban and rural station identification and validation. First, we examined the similarity of FT/SDZ stations with other stations to identify possible unhomogeneous conditions of these temperature sequences (Figure 4). Except for the station combination of FT station versus PG station, higher similarity can be found among urban stations with the distance less than 200. PG station shouldn't be considered as a typical urban station with the STS values by 404.3, meanwhile it is also relative far away from city limit of Beijing (Figure 1). SDZ station is a typical rural station that hasn't undergone the relocation, and its temperature change has higher similarity with other rural stations, such as DTW distance by 264 and 180 for XYL station and ZT station, respectively. The STS analysis also suggests that FYD station, MY station and HR station aren't appropriate to represent air temperature change in rural surface. According to the environmental evolution of urban station and rural station, we can conclude that urban stations have usually maintained its relative stable environment, while the surroundings of rural stations always experienced the dynamic change with the increased urbanization signals, especially those stations located at urban fringe. Urban land teleconnection analysis also suggests the traditional classification based on discrete categorize of a rural-urban dichotomy couldn't meet the requirements of the continuum analysis for urban economy and sustainability(Yan et al. 2010). Therefore, the place-based conception for urban and rural station identification can't fulfil the task for typical urban and rural station selection either. According to these

STS results from FT station and SDZ station, we can further divide 20 stations into three groups: typical urban stations (FT, SY, HD, TZ, CY, CP, MTG, BJ, SJS, DX, FS), typical rural stations (ZT, SDZ, XYL), other stations (THK, YQ, PG, FYD, HR, MY). Here, other stations means that their air temperature change couldn't represent typical characterization of local climate of urban or rural surface.

In order to validate our classification results about typical urban and rural stations derived from FT station and SDZ station, we further analyze the STS between each station and our classification group, then examine the percentage of each station similar to our classification group using the distance threshold 280 (Figure 5). As expected, the percentage analysis showed the same patterns with our station classification results, such as the percentage of FT, SY, HD, TZ, CY, CP, MTG, BJ, SJS, DX, FS station similar to typical urban station reach to 81% and the percentage of ZT, SDZ, XYL station similar to typical urban station greater than 65%. Meanwhile, YQ, FYD, THK and MY station aren't similar to neither urban station nor rural station, while only 36.4% results from HR station and 9.1% results from PG station are similar to urban station. Therefore, our classification for typical urban and rural station based on DTW algorithm is reasonable.

3.3 Evaluation of UHI derived from selected stations

3.3.1 Comparison of UHI from new station groups and traditional station groups

We compared UHI derived from traditional station groups and new station groups from DTW algorithm to identify the difference of urban climate effect induced by station selection (Figure 6). Here, the typical mountain station, FYD station, was also used to examine UHI variation altered by its altitude. Generally, the comparison of UHI derived from different station combinations in Beijing indicated that traditional station groups will underestimate UHI effects. Averaged UHI effects derived from station groups by DTW algorithm reach 2.31±0.51℃, while they decrease to 1.07±0.57℃ by using traditional station groups. This phenomenon is probably caused by warming signals in rural stations induced by urbanization process, and the environment investigation by remote sensing images indicates that urban land has already accounted for about 50% and 38% of the surroundings of MY station and

HR station. UHI effects derived from urban stations versus mountain station reach to 7.13±0.79℃, higher than UHI effects than other two combinations, mainly because air temperature greatly decreased along rising altitude. UHI effects have been intensified among three station combinations from 1987-2016, and it increase by 2.13℃/decade and 0.95℃/decade by new station groups and traditional station groups, respectively.

3.3.2 Comparison of UHI from air temperature and satellite-based land surface temperature

UHI is usually classified as surface urban heat island (SUHI), canopy urban heat island (CUHI) and boundary urban heat island (BUHI) according to their characterization in different layers of urban atmosphere (Hu et al. 2019; Yuan and Bauer 2007). SUHI can be derived from land surface temperature records, while CUHI is calculated by the difference between urban and rural temperature records. Annual SUHI and CUHI records from different mega-cities have similar trends and close magnitude (Hu et al. 2019). Meanwhile, surface temperature from large area of urban and rural surface are adopted in quantifying SUHI, which maintain more information than single stations. Figure 7 shows the relationship between air temperature and satellite-based land surface temperature from 2003-2016, and the result indicates that they have good consistency in temperature variation at urban and rural stations except for air temperature lower than land surface temperature about 5℃. Land surface temperature variation mainly represent skin temperature, more controlled by the change of solar radiation. Land surface temperature records derived from remote sensing images usually capture the averaged temperature conditions of land surface with more spatial details.

The SUHI from 2003-2016 was quantified by land surface temperature difference between urban and rural surface (Hu et al. 2019), then we examined the similarity between SUHI and CUHI sequences, and found that the similarity between SUHI sequences and CUHI from new station groups reach 81.94, which is less than that from SUHI versus CUHI from traditional station groups (113.76) and SUHI versus CUHI from mountain station (832.77). These results indicated that SUHI

sequences has better similarity with CUHI calculated from new station groups than that derived from traditional station group. The difference between averaged SUHI and CUHI effects from 2003-2016 also shows that the minimum deviation (0.3℃) was found between SUHI and CUHI from new station groups, while this deviation increase to 1.02℃ for SUHI versus CUHI from traditional station groups, and the largest difference happened between SUHI and CUHI from mountain station by 4.96℃. Those above results suggest that the temporal change and overall magnitude of UHI from new station groups identified by DTW algorithm more approach the SUHI results, which prove that new station group is helpful for UHI estimation. DTW algorithm may have special advantages in selecting the typical station groups for comparing UHI results across regions, especially for rapid urbanization regions. In these regions, urban fringe are considered as a transition zone from urban to rural area with highly interaction between urban expansion and rural activities, and it already became the fastest changing regions with their rapid land surface transformation. In fact, it is hard to exactly distinguish the border of suburban, urban fringe and rural area according to the urban-rural continuum characterization. Remote sensing images may provide exact spatial information. However, we may not sure whether meteorological stations have already experienced the transformation of station type because of the possible warming signals induced by urbanization process from near surface urban land (Hu et al. 2019). Thus, we think it is better for typical urban and rural station selection by considering spatial coverage information from remote sensing and temporal change information by DTW algorithm.

4 Conclusion

The dynamic time warping algorithm (DTW) was investigated in this study to improve the selection of urban and rural stations for UHI estimation. The similarity of temperature sequences (STS) was derived from air temperature sequences from 1987-2016 of 20 meteorological stations in Beijing using DTW algorithm. Then, we analyzed the station combinations with high similarity of temperature sequences using the cumulative probability distribution analysis, further divided those station into three groups to identify typical urban stations (FT, SY, HD, TZ, CY, CP, MTG, BJ,

SJS, DX, FS) and typical rural stations (ZT, SDZ, XYL). According to the STS analysis and validation by remote sensing images, new station groups for UHI estimation is reasonable. Meanwhile, this method is also helpful in detecting inhomogeneous signals of temperature sequences of traditional rural stations, and we found that PG station, MY station and HR station is not suitable for representing temperature conditions in rural surface any more because they have been influenced by urbanization. We also found that mountain stations, such as FYD station, have big differences with rural station and urban station, and they aren't appropriate to analyze urban climate effect.

Our study analyzed the bias in the selection of typical urban and rural stations, and the results proved that dynamic urbanization process also caused potential uncertainties for urban identification, especially for those station located at urban fringe experienced rapid urbanization. Therefore, DTW algorithm analysis for air temperature sequences may provide another simple way to investigate them to understand their representativeness. Another advantages of this method is less requirement for background information of meteorological stations. Therefore, it is more convenient for researchers to understand the stations situation in a new region, even in regional or global scale. Certainly, it also has some limitations or uncertainties when less samples of meteorological stations adopted in this algorithm, and we should know more background information for better understanding under this situation.

Conflict of interest

The authors declared that they have no conflicts of interest to this work.

Authors' contributions

Conceptualization, Yonghong Hu and Gensuo Jia; Investigation, Jinlong Ai; Writing and original draft, Yonghong Hu; Writing-reviewing and editing, Yonghong Hu and Yong Zhang; Methodology, Meiting Hou and Yapeng Li. All the authors have read and agreed to the published version of the manuscript.

Funding

This study is supported by CAS Strategic Priority Research Program (XDA19030401), the National Natural Science Foundation of China (41861124005) and Natural Science Foundation for Young Scientists of Hunan Province (2020JJ5557).

References:

Camilloni, I., and Barros, V. (1997). On the urban heat island effect dependence on temperature trends. Climatic Change, 37, 665-681.

He, J.F., Liu, J.Y., Zhuang, D.F., Zhang, W., and Liu, M.L. (2007). Assessing the effect of land use/land cover change on the change of urban heat island intensity. Theoretical and Applied Climatology, 90, 217-226.

He, Y., Jia, G., Hu, Y., and Zhou, Z. (2013). Detecting urban warming signals in climate records. Advances in Atmospheric Sciences, 30, 1143-1153.

Hu, Y., Hou, M., Jia, G., Zhao, C., Zhen, X., and Xu, Y. (2019). Comparison of surface and canopy urban heat islands within megacities of eastern China. ISPRS Journal of Photogrammetry and Remote Sensing, 156, 160-168.

Hu, Y., Jia, G., Hou, M., Zhang, X., Zheng, F., and Liu, Y. (2015). The cumulative effects of urban expansion on land surface temperatures in metropolitan JingjinTang, China. Journal of Geophysical Research: Atmospheres, 120, 9932-9943.

Hu, Y., Jia, G., Pohl, C., Zhang, X., and van Genderen, J. (2015). Assessing Surface Albedo Change and its Induced Radiation Budget under Rapid Urbanization with Landsat and GLASS Data. Theoretical and Applied Climatology, 123,711-722.

Jia, G., Xu, R., Hu, Y., and He, Y. (2015). Multi-scale remote sensing estimates of urban fractions and road widths for regional models. Climatic Change, 3, 543-554.

Jin, M.S. (2012). Developing an Index to Measure Urban Heat Island Effect Using Satellite Land Skin Temperature and Land Cover Observations. Journal of Climate, 25, 6193-6201.

419 Keogh, E.J., and Pazzani, M.J. (1999). Scaling up Dynamic Time Warping to Massive
420 Datasets. In, European Conference on Principles of Data Mining and Knowledge
421 Discovery (pp. 1-11).
422 Kim, Y.H., and Baik, J.J. (2005). Spatial and temporal structure of the urban heat
423 island in Seoul. Journal of Applied Meteorology, 44, 591-605.
424 Mohsin, T., and Gough, W.A. (2012). Characterization and estimation of urban heat
425 island at Toronto: impact of the choice of rural sites. Theoretical and Applied
426 Climatology, 108, 105-117.
427 Normile, D. (2008). China's living laboratory in urbanization. Science, 319, 740-743.
428 Oke, T.R. (2006). Towards better scientific communication in urban climate.
429 Theoretical and Applied Climatology, 84, 179-190.
430 Parker, D.E. (2010). Urban heat island effects on estimates of observed climate
431 change. Wiley Interdisciplinary Reviews-Climate Change, 1, 123-133.
432 Ren, G.Y., Chu, Z.Y., Chen, Z.H., and Ren, Y.Y. (2007). Implications of temporal
433 change in urban heat island intensity observed at Beijing and Wuhan stations.
434 GEOPHYSICAL RESEARCH LETTERS, 34, L05711.
435 Ren, G.Y., Zhou, Y.Q., Chu, Z.Y., Zhou, J.X., Zhang, A.Y., Guo, J., and Liu, X.F.
436 (2008). Urbanization effects on observed surface air temperature trends in north
437 China. Journal of Climate, 21, 1333-1348.
438 Seto, K.C., Reenberg, A., Boone, C.G., Fragkias, M., Haase, D., Langanke, T.,
439 Marcotullio, P., Munroe, D.K., Olah, B., and Simon, D. (2012). Urban land
440 teleconnections and sustainability. Proceedings of the National Academy of
441 Sciences, 109, 7687-7692.
442 Stewart, I.D. (2011). A systematic review and scientific critique of methodology in
443 modern urban heat island literature. International Journal of Climatology, 31, 200,
444 217.
445 Stewart, I.D., and Oke, T.R. (2012). Local climate zones for urban temperature studies.
446 Bulletin of the American Meteorological Society, 93, 1879-1900.
447 Tan, J., Zheng, Y., Tang, X., Guo, C., Li, L., Song, G., Zhen, X., Yuan, D., Kalkstein,
448 A.J., and Li, F. (2010). The urban heat island and its impact on heat waves and

449 human health in Shanghai. International Journal of Biometeorology, 54, 75-84.
450 Vancutsem, C., Ceccato, P., Dinku, T., and Connor, S.J. (2010). Evaluation of MODIS
451 land surface temperature data to estimate air temperature in different ecosystems
452 over Africa. Remote Sensing of Environment, 114, 449-465.
453 Voogt, J.A., and Oke, T.R. (2003). Thermal remote sensing of urban climates. Remote
454 Sensing of Environment, 86, 370-384.
455 Wang, X., Mueen, A., Ding, H., Trajcevski, G., Scheuermann, P., and Keogh, E.
456 (2013). Experimental comparison of representation methods and distance
457 measures for time series data. Data Mining and Knowledge Discovery, 26,
458 275-309.
459 Weng, Q.H., Lu, D.S., and Schubring, J. (2004). Estimation of land surface
460 temperature-vegetation abundance relationship for urban heat island studies.
461 Remote Sensing of Environment, 89, 467-483.
462 Wu, K., and Yang, X. (2013). Urbanization and heterogeneous surface warming in
463 eastern China. Chinese Science Bulletin, 58, 1363-1373.
464 Yan, Z., Li, Z., Li, Q., and Jones, P. (2010). Effects of site change and urbanisation in
465 the Beijing temperature series 1977-2006. International Journal of Climatology,
466 30, 1226-1234.
467 Yuan, F., and Bauer, M.E. (2007). Comparison of impervious surface area and
468 normalized difference vegetation index as indicators of surface urban heat island
469 effects in Landsat imagery. Remote Sensing of Environment, 106, 375-386.
470 Zhang, D.L., Shou, Y.X., and Dickerson, R.R. (2009). Upstream urbanization
471 exacerbates urban heat island effects. Geophysical Research Letters, 36, L24401.
472 Zhu, W., Lű, A., and Jia, S. (2013). Estimation of daily maximum and minimum air
473 temperature using MODIS land surface temperature products. Remote Sensing of
474 Environment, 130, 62-73.
475
476
477
478
479

Table list:

Figure list:

Figures

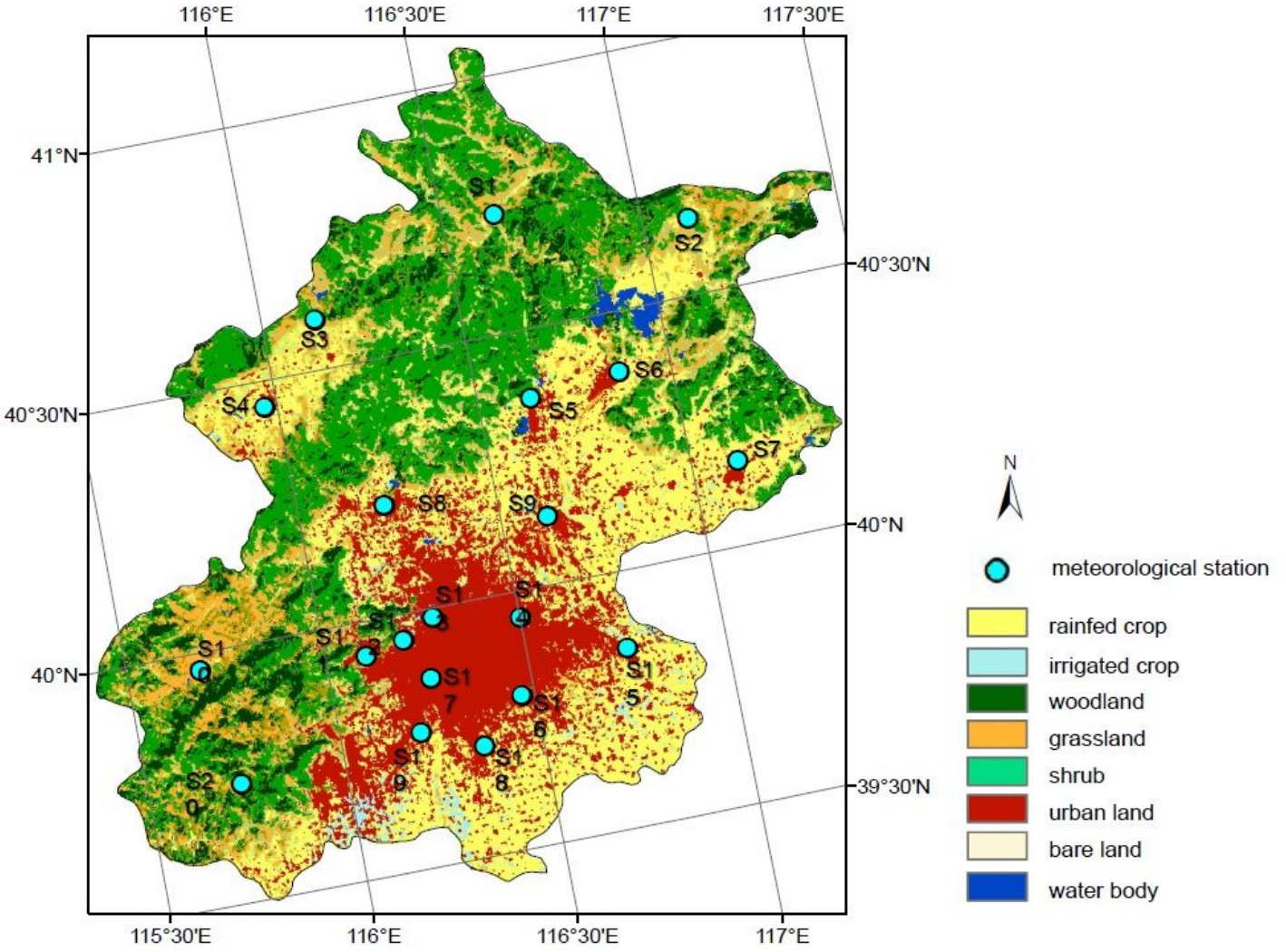

Figure 1

The location and surrouding land use type of meteorological stations in Beijing. Details about these meteologcial stations can refer to Table 1. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

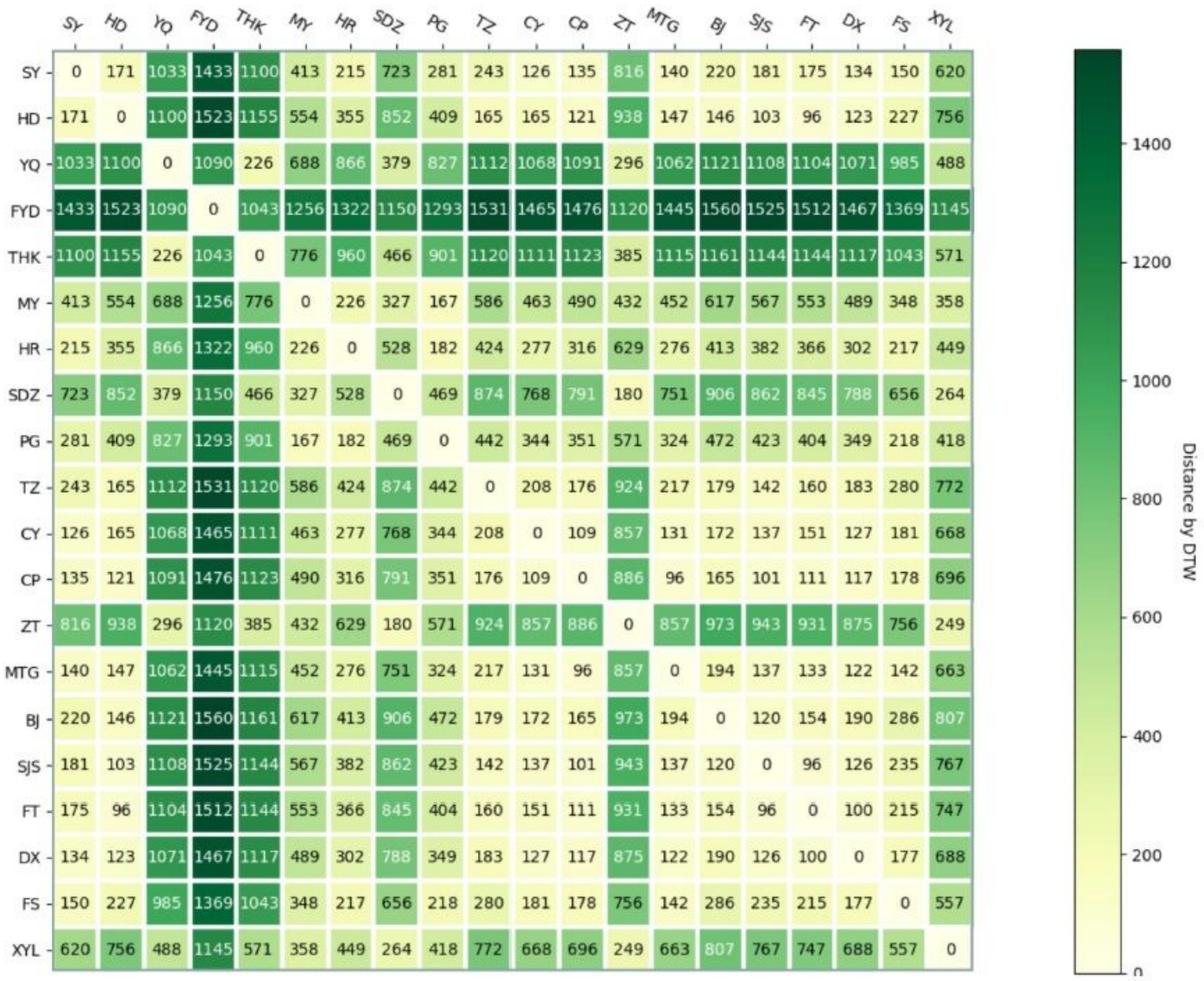

Figure 2

The distances derived from DTW algorithm for air temperature sequences from 1987-2016 of every pair of station in Beijng.

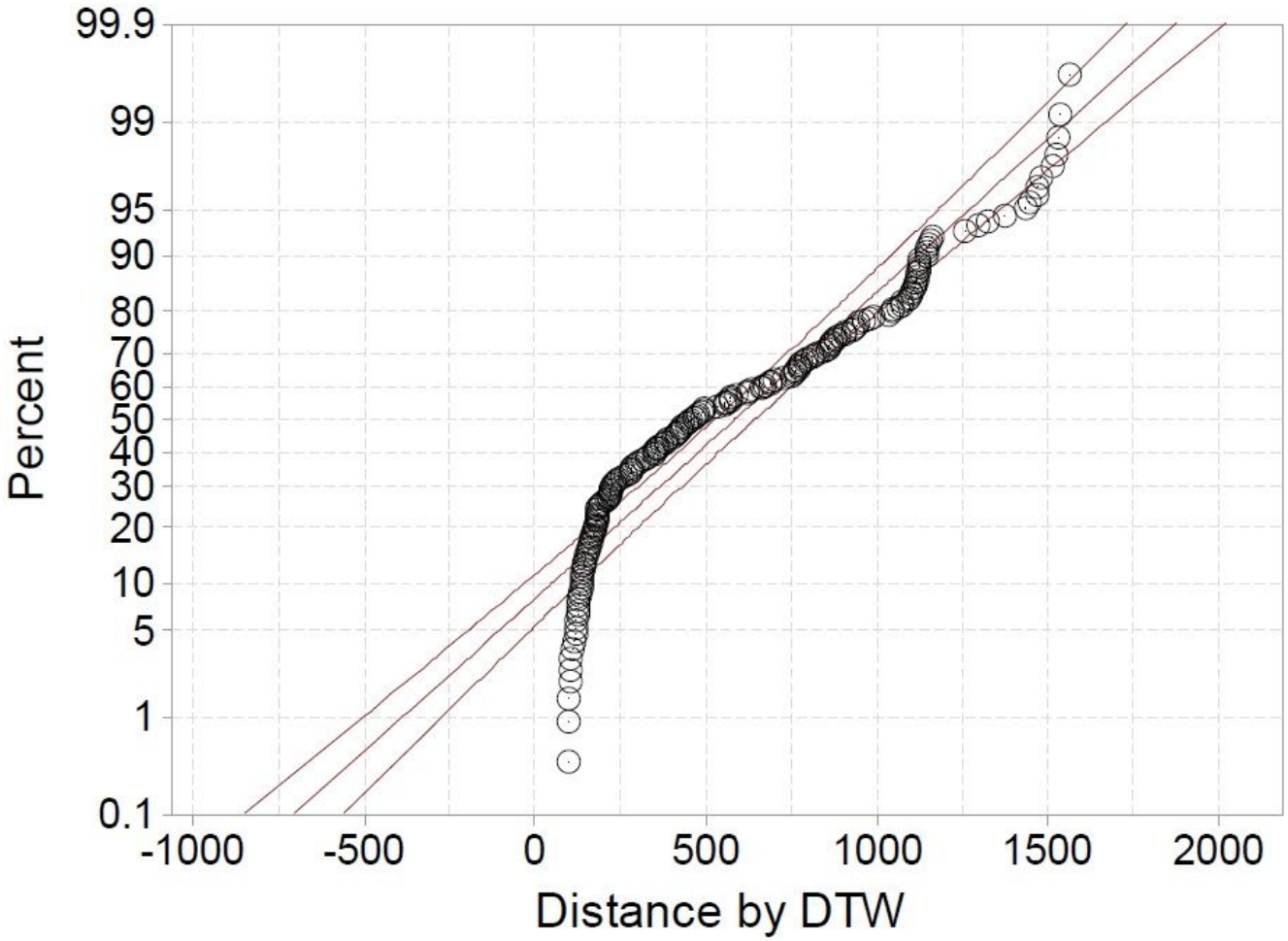

Figure 3

The cumulative probability distribution for the distance values in analyzing the similarity of air temperature sequences.

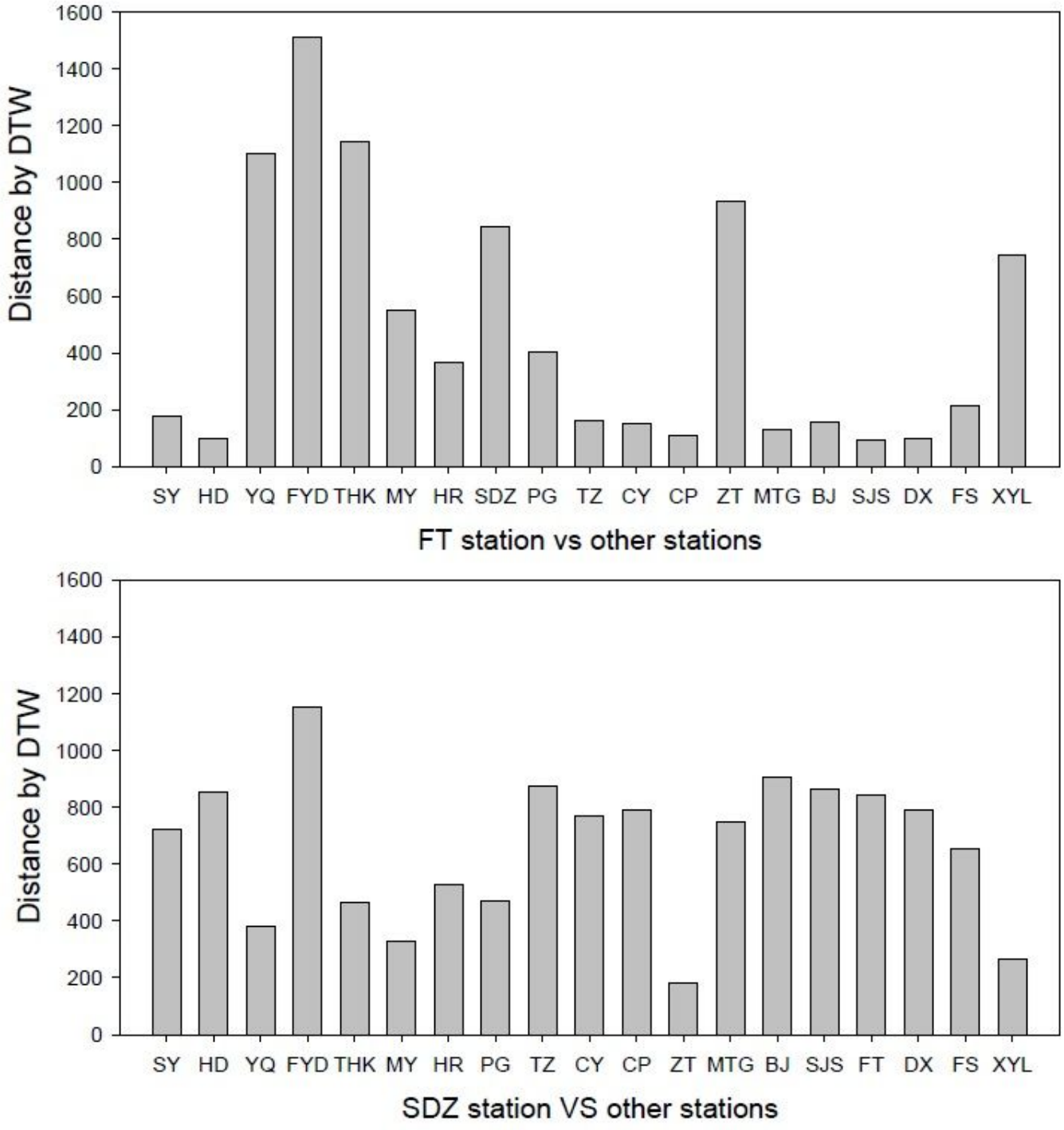

Figure 4

The DTW distance of air temperature sequences from 1987-2016 of the selected typical urban/rural station with other stations in Beijing.

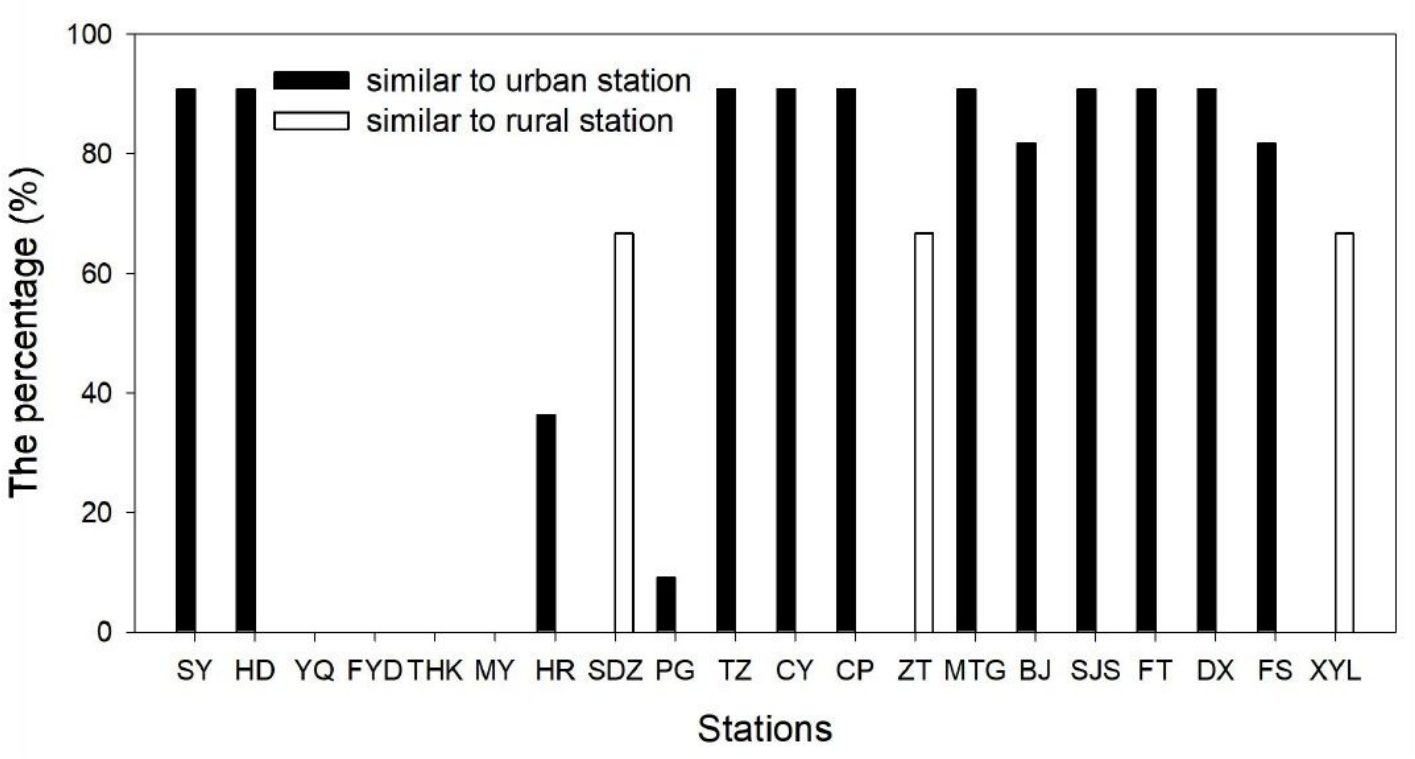

Figure 5

The percentage of each station similar to typical urban/rural stations derived from referenced station.

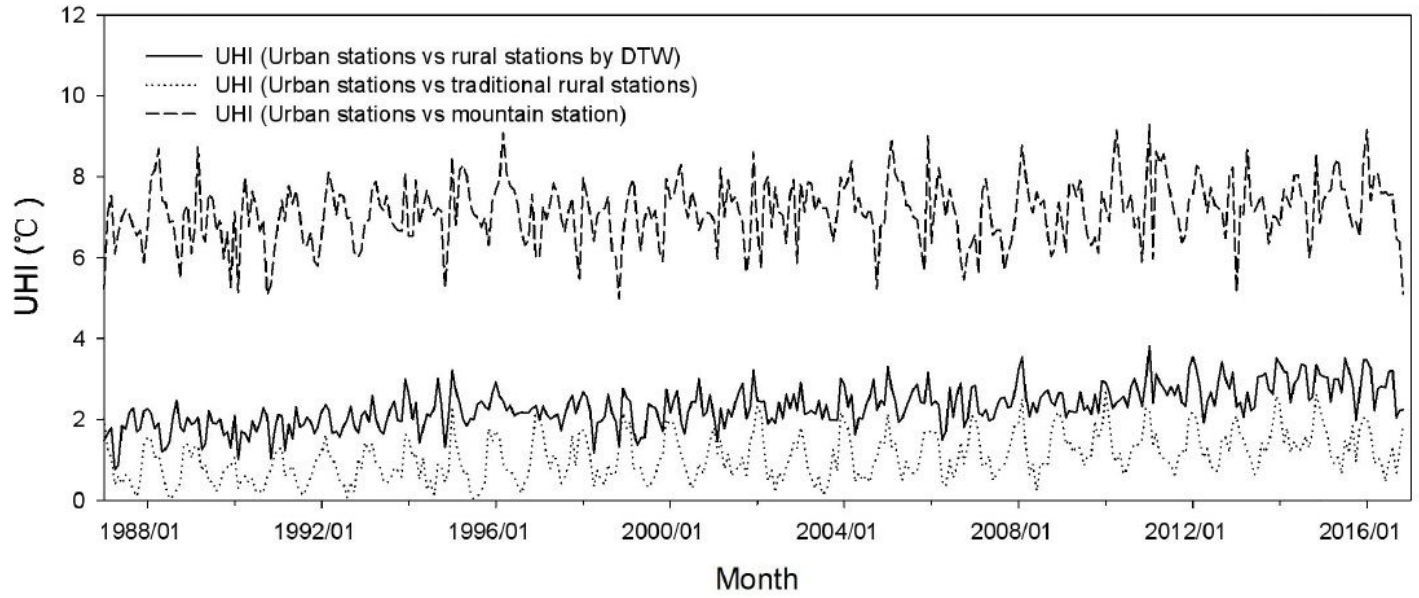

Figure 6

The comparison of UHI estimation from different station combinations in Beijing, including urban stations vs rural stations by DTW, urban stations vs traditional rural stations, and urban stations vs

mountain station.

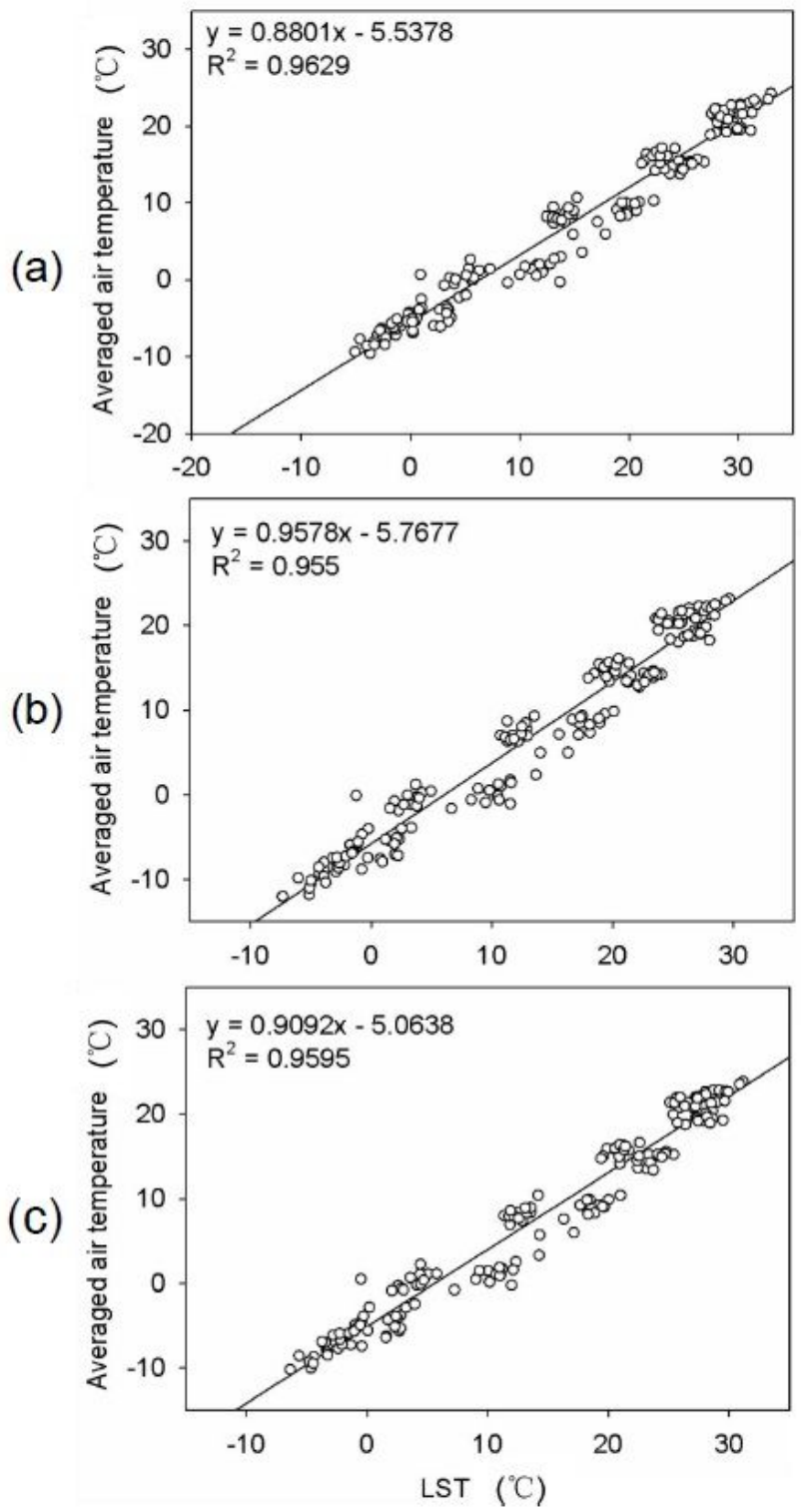

Figure 7

The relationship between land surface temperature (LST) and air temperature in typical urban stations (a), typical rural stations (b) and other stations (c) in Beijing from 2003- 2016.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table.pdf