University of Reading

Network optimization for enhanced resilience of urban heat island measurements

Article

Accepted Version

Honjo, T., Yamato, H., Mikami, T. and Grimmond, C. S. B. ORCID: https://orcid.org/0000-0002-3166-9415 (2015) Network optimization for enhanced resilience of urban heat island measurements. Sustainable Cities and Society, 19. pp. 319-330. ISSN 2210-6707 doi: 10.1016/j.scs.2015.02.004 Available at https://centaur.reading.ac.uk/39688/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.
Published version at: http://dx.doi.org/10.1016/j.scs.2015.02.004
To link to this article DOI: http://dx.doi.org/10.1016/j.scs.2015.02.004

Publisher: Elsevier

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

University of Reading

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Network optimization for enhanced resilience of urban heat island measurements

Tsuyoshi Honjo1, Hiroaki Yamato2, Takehiko Mikami3, C.S.B. Grimmond4

1Department of Environmental Science and Landscape Architecture, Faculty of Horticulture, Chiba University, Matsudo, Chiba, 271-8510, Japan, 2Chuo University, 1-3-27 Kasuga, Bunkyo-ku, Tokyo, 112-8551, Japan, 3Faculty of Liberal Arts, Teikyo University, Hachioji, Tokyo, 192-0395, Japan, 4Department of Meteorology, University of Reading, Earley Gate, PO Box 243, Reading RG6 6BB, UK. Tsuyoshi Honjo,honjo@faculty.chiba-u.jp

Abstract
The urban heat island is a well-known phenomenon that impacts a wide variety of city operations. With greater availability of cheap meteorological sensors, it is possible to measure spatial patterns of urban atmospheric characteristics with greater resolution. To develop robust and resilient networks, recognizing sensors may malfunction, it is important to know when measurement points are providing additional information and also the minimum number of sensors needed to provide spatial information for particular applications. Here we consider the example of temperature data, and the urban heat island, through analysis of a network of sensors in the Tokyo metropolitan area (Extended METROS). The effect of reducing observation points from an existing meteorological measurement network is considered, using random sampling and sampling with clustering. The results indicated the sampling with hierarchical clustering can yield similar temperature patterns with up to a 30% reduction in measurement sites in Tokyo. The methods presented have broader utility in evaluating the robustness and resilience of existing urban temperature networks and in how networks can be enhanced by new mobile and open data sources.

Keywords: meteorological measurement network, random sampling, hierarchical clustering, IDW (Inverse Distance Weighting), urban heat island

1. Introduction

The urban heat island is a well-known phenomenon (Stewart, 2011) that impacts a wide variety of city operations (e.g. energy usage, need for snow clearance). With the drive to develop both "Smart Cities" and the "Internet of Things" (Perera, 2014), and the availability of relatively cheap meteorological sensors, it is now desirable and possible to observe the spatial pattern of atmospheric variables with much more detail than in the past. Many cities now have multiple sensors installed (see the examples in Table 1). Such networks allow the spatial dynamics of air temperature to be considered.

Table 1. Examples of air temperature networks that have been installed in cities. Table ordered by year first installed. Note that some cities have had multiple networks installed, these are listed separately.

City	*Number of sensors*	*Spatial extent*	*Period of operation*	*Project*	*Reference*
Tokyo, Japan	120	2187	2002-2005	METROS	Mikami et al., 2003
Taipei, Taiwan	60	271.79	2003-present	TWIN	Chang et al, 2010
Washington. DC, USA	16	177	2003-present	DCNet	Hicks et al., 2012
Helsinki, Finland	102	150	2005-present	Helsinki Testbed	Koskinen et al., 2011
Cambridge, MA, USA	25	18.47	2006-2010	CitySense	Murty et al., 2008
Tokyo, Japan	200	7000	2006-present	Extended METROS	Yamato et al., 2009
Oklahoma City, USA	40	1440	2007-2010	OKCNET	Basara et al., 2010
Hong Kong	105	1104	2007-present	Co-WIN	Hung and Wo, 2012
London, UK	91	1577	2009-present	OPAL	Davies et al., 2011
Worldwide	N/A	N/A	2013-present	Weather Signal	Overeem et al, 2013
Shanghai, China	200	6340.5	N/A-present	SUIMON	Tan et al., 2014

In the design and operation of a meteorological measurement network for spatial characterization of the urban heat island, many questions arise. In the planning stage, key questions include what is the optimal number of measurement points and what is their optimal distribution? Should measurement points be distributed evenly or unevenly? Are there important features which must be measured? Once measurements are initiated, if instruments malfunction or are withdrawn, questions need to be answered as to whether sensors should be substituted at these points or is this unnecessary? How significant are missing data to the overall performance of the network? A priori it might be expected that a larger number of measurement points is better, but what is the trade-off with respect to resources, such as labor to service and sustain the network?

In this study, we analyse an existing meteorological measurement network around the Tokyo metropolitan area - Extended METROS (Yamato et al., 2009) to consider optimal network design for urban heat island measurements to address some of these questions.

2. Material and Methods

2.1 Data sets

The measurement points of Extended METROS are shown in Fig.1. The area is one of the most populated in the world, with more than 30 million people. The area analyzed extends from 35.3° to 36.3°N latitude (about 90.4 km) and from 139.2° to 140.2°E longitude (about 111.3 km) and includes Tokyo Bay (Fig 1, note 1 station is beyond the specified area). To the east of the study area is the Pacific Ocean. Elevation varies from mean sea level to 300 m. The density of the stations is approximately 1 station per

50 km2 or a mean separation of 7 km. However, given no measurement points are located in 30% of the area (Tokyo Bay: 9%, mountain area and unpopulated areas 20 %), the mean separation of sensors is actually about 6 km.

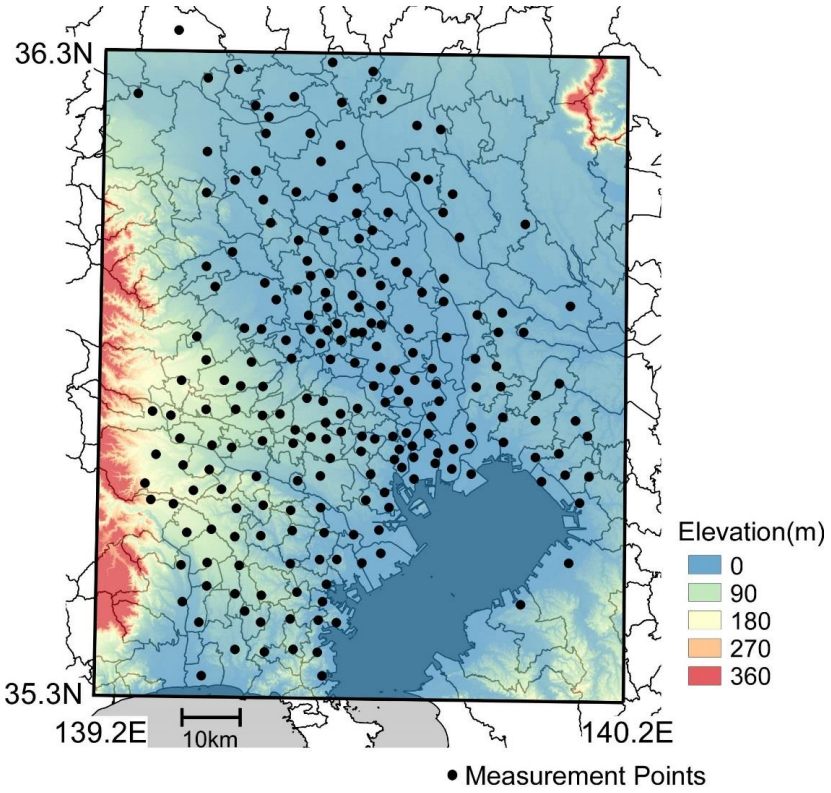


Fig. 1 The location of measuring points of the Extended METROS network in the greater Tokyo area. Lines within the land indicates borders of local governments, i. e., ward, city, and town.

Thermometers (T&D, TR5106 thermistor sensor with RTR-52a data logger) are housed in unaspirated Stevenson screens at 1.5 m above ground level sited in elementary schools. The individual network, which is operated/coordinated by METROS Research Project, was set up by 12 independent groups. No individual has been to every site, as will be the case for many large networks that are being established (formally or ad hoc networks of opportunity). The temperature measurements, collected since 2006, consist of samples taken every 10 min from approximately 200 sites. The actual number at any given time varies because of instrument related issues (malfunctions etc.).
For the analysis presented here, examples from the hottest and coldest periods are selected to consider the implications of network distribution and density. The data were sampled at 4:00, 10:00, 16:00, 22:00 (local time, note Japan does not use summer time) on 15 and 16 August 2007 (8 summer cases) and 13 and 14 February 2008 (8 winter cases) as explained in Table 2. The mean and standard deviation of the temperature and locations of the heat island of the cases are also reported. The numbers of measurement points were 183 in the summer cases and 200 in the winter cases. Histograms of the original data are shown in Fig. 2.

Table 2. The mean and standard deviation of the air temperature for the 183 (summer) and 200 (winter) sensors for the 16 cases analysed.

Case No.	Time and Day		Temp Mean (°C)	Temp SD (°C)	Heat Island (HI) location
1	4:00	15th Aug.	27.22	0.92	HI at the city center
2	10:00	15th Aug.	33.60	1.19	transition
3	16:00	15th Aug.	35.28	1.69	HI at inland
4	22:00	15th Aug.	30.36	1.14	transition
5	4:00	16th Aug.	28.64	0.98	HI at the city center
6	10:00	16th Aug.	35.06	1.14	transition
7	16:00	16th Aug.	34.29	2.33	HI at inland
8	22:00	16th Aug.	30.63	1.36	transition
9	4:00	13th Feb.	1.19	0.74	HI at the coast
10	10:00	13th Feb.	3.39	1.05	no HI
11	16:00	13th Feb.	4.65	1.10	no HI
12	22:00	13th Feb.	0.67	1.29	HI at the coast
13	4:00	14th Feb.	-1.02	1.77	HI at the coast
14	10:00	14th Feb.	6.36	0.98	no HI
15	16:00	14th Feb.	8.81	0.79	no HI
16	22:00	14h Feb.	2.64	1.18	HI at the coast

2.2 Random sampling and analysis of mean temperature

Random sampling was undertaken to analyse the effect of reducing the number of measurement points. Each station (183 in summer; 200 in winter) was numbered and for each case a selection of stations was sampled for a pre-determined percentage of the data 1000 times. The percentage of data analysed varied from10% to 90%, in 10% increments. Sampling 10% of the data resulted in a 90% reduction of the data.

Air temperature data Ti,c measured at site i of then sites (summer=183, winter=200) at a particular time and date c (c =1 to 16, Table 2) were analysed. For the random sampling, a set of random order ri,j was used, where i is from 1 to n (as specified above), which belong to the j random sampling set (j= 1 to 1000). Same sets of random order were used for these cases and random sampling from 10% to 90% considering real situation of networks. When the different sets of random order were used in pre-calculation, characteristics were same with the usage of same sets. The selected data Tsk,l,c from the original data can be expressed with rk,l (k=1 to m, l=1 to 1000) as follows,

$$T_{S_{k,l,c}} = T_{r_{k,l,c}} \quad (1)$$

$$m = n \cdot s \quad (2)$$

where s is a sampling ratio from 0.1 to 0.9 with 0.1 interval, and m is the number of sampled points. Thus when n=183 and s=0.2, m=37. If s is 1, the original data Ti,c all are used instead of Tsk,l,c.. From the sampled data, mean temperature is calculated:

$$T_{a_{l,c}} = \sum_{k=1}^{m} T_{S_{k,l,c}}/m \quad (3)$$

In the analysis of mean temperature with random sampling, 1000 cases of mean temperature are calculated in each sampling ratio (x 9) and each case (x 16).

2.3 Inverse distance weighting (IDW) method with random sampling

To understand the spatial influence of reducing the measurement points, interpolation is necessary to obtain a contour map or 2-D image of the distribution. Frequently used method, especially in soil science and hydrology but also applied to meteorological data, are kriging and inverse distance weighting (IDW) algorithm.

Kriging and IDW have been compared in many studies. Kriging shows better results in some studies (e.g. Zimmerman et al., 1999; Chaplot et al., 2006), while IDW better results in others (e.g. Weber and Englund, 1992; Nalder and Wein, 1998). In other studies, the differences were insignificant (Jarvis and Stuart, 2001; Kravchenko, 2003; Mueller et al., 2004). In this study, a simple interpolation method, IDW was used.

With IDW the interpolated temperature Tl,c(x) was calculated from sampled data Tsk,l,c by:

$$\mathrm{T}_{l,c}(x) = \frac{\sum_{k=1}^{m} w_k(x) T s_{k,l,c}}{\sum_{k=1}^{m} w_k(x)} \quad (4)$$

where the weighting function

$$W_k(x) = \frac{1}{d(x,x_k)^p} \quad (5)$$

is related to x, the coordinate vector of an interpolated point, xk is a measurement point, d(x,xk) denotes the distance from x to xk, m is the number of sampled points, and the positive number p=2 is used in this study to reflect the local temperature variation. To obtain an interpolated image of the original data, Ti,c is used instead of Tsk,l,c. The same IDW method is used after sampling with hierarchical clustering in section 2.4.

Using the IDW method, spatial domains with 201 x 201= 40401 interpolated points were generated for analysis. These can also be visualized. To estimate the similarity between two spatial domains or interpolated images, normalized cross-correlation, Rncc (hereafter the correlation) and root-mean-square error (RMSE) were used:

$$R_{ncc} = \frac{\sum_{ix=1}^{N}\sum_{iy=1}^{N} T_1(ix,iy) T_2(ix,iy)}{\sqrt{\sum_{ix=1}^{N}\sum_{iy=1}^{N} T_1(ix,iy)^2 \sum_{ix=1}^{N}\sum_{iy=1}^{N} T_2(ix,iy)^2}} \quad (6)$$

$$RMSE = \sqrt{\frac{\sum_{ix=1}^{N}\sum_{iy=1}^{N} (T_1(ix,iy) - T_2(ix,iy))^2}{N^2}} \quad (7)$$

whereT1 (ix,iy)and T2 (ix,iy)are the interpolated values of air temperature from the original and sampled data, respectively; ix and iy are coordinates of the interpolated image; and N is the number of pixels in the x-y dimension (N=201). For the analysis of the IDW method with random sampling, 1000 images were made and compared in each sampling ratio (x 9) and each case (x 16).

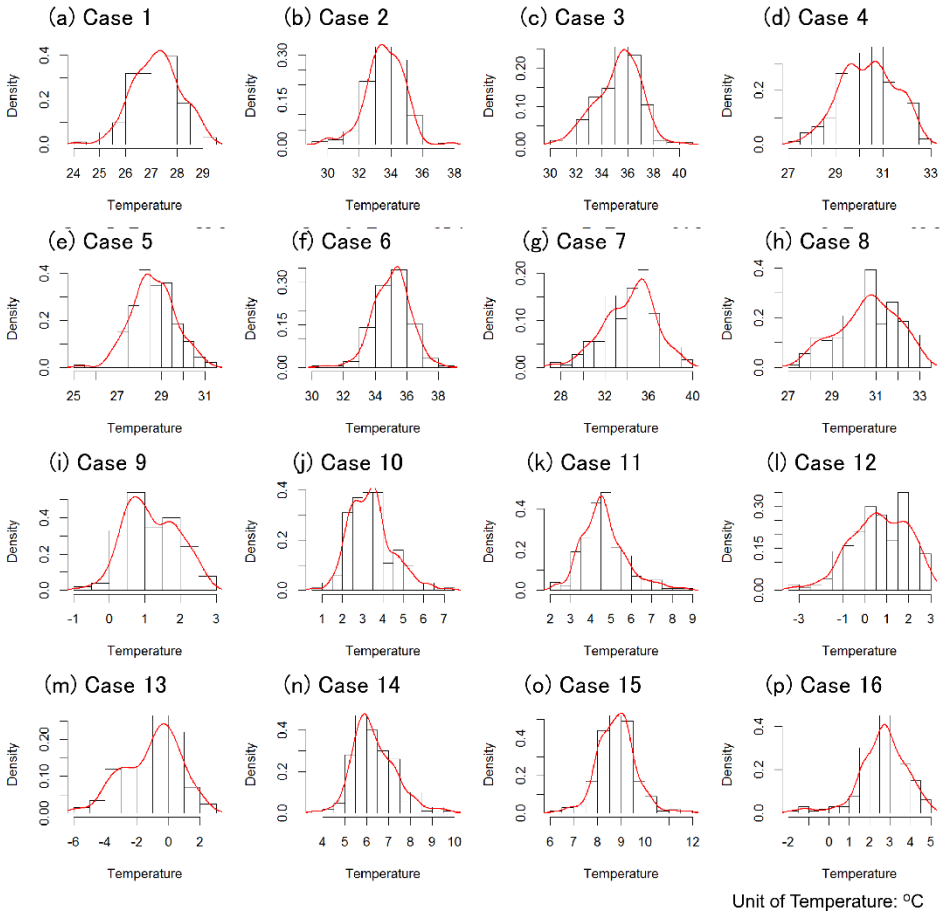


Fig. 2 Histograms of original temperature (◦C) distribution when 183 (summer, cases 1-8) and 200 (winter, cases 9-16) sensors are used. See Table 1 for details of each case.

2.4 IDW method with hierarchical clustering and sampling

In many studies, optimal sampling or selection of measurement points is undertaken using spatial simulated annealing (SSA) with

a criterion like mean kriging variance (Van Groenigen, 1999;Heuvelink et al., 2006;Brus and Heuvelink, 2007). Baume et al. (2011) indicated "greedy algorithms" perform better than SSA for optimal spatial measurement network design. These studies have primarily considered soil and water quality, not meteorological data which fluctuates more temporally.

Brus et al. (2006) suggested that simple methods with k-means clustering could be an attractive alternative to optimization by SSA with the kriging variance. Examples of designing spatial coverage with k-means clustering are provided by Walvoort (2010). One of the difficulties in k-means clustering is that the method does not always produce same sampling results because of problems of convergence to local minima and the influence of initial conditions.

An alternative approach is hierarchical clustering. Measurement points are classified into m categories over the measurement domain and then stations within each category of the network are selected (Fig. 3) - hereafter referred to as clustering. In this case, the longitude and latitude of the measurement points were used for the classification by the clustering. In this study, hierarchical clustering was used for sampling instead of k-means clustering and the results of classifications are always same.

After the clustering and calculation of the center position of each category, the nearest point to the center position of each category was chosen as the sampled point and the temperature of that point was used as the representative value of the category.

By using the sampled point with clustering, and the same procedures as section 2.3, 201 x 201 interpolated data were obtained with IDW. The correlation and RMSE were also calculated from the interpolated data. In the analysis of the IDW method with clustering, one image was made and compared in each sampling ratio (x 9) and each case (x 16).

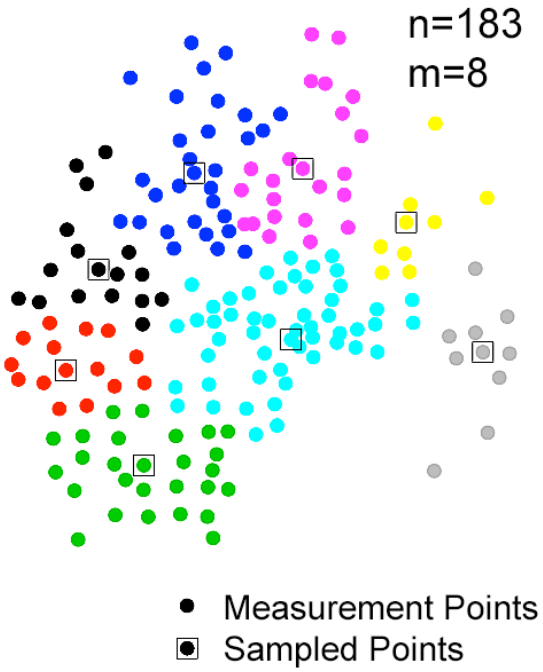


Fig. 3 Example of choosing m=8 points from n=183 measurement points by sampling with hierarchical clustering. First, n points are classified into m categories expressed as different colors. Second, center points of each category are selected as sampled points.

3. Results

Although data from only four days and four times of day were analysed (Table 2), a large number of different calculations have been performed to assess the spatial features of the air temperature at these times. In summer, the mean air temperature shows the expected pattern of being coolest at 04:00 local time (27 to 29°C) and warmest at 16:00 (34 to 35°C), with the least and greatest spatial variability, respectively, also being recorded at these times (Table 2). For both days, urban heat island maxima were seen both at the city center (04:00) and inland (16:00), with transitions occurring at the other two times (10:00, 22:00). As expected, the winter temperatures were much cooler, ranging from means of -1°C to 9°C for the eight sample times. In the winter, the spatial standard deviations were generally smaller and the UHI was evident at the coast.

The overall analysis considers the impact of sensor location and density. To ensure the results are robust, the 1000 randomly selected cases were analysed. This was repeated for each case (Table 2, c =1 to 16) and sampling ratio (s = 0.1 to 0.9) for both the random and clustered network patterns. For each, the mean air temperature was determined.

3.1 Random sampling and analysis on mean temperature

First we consider a network with randomly located sites (Fig. 1), so any sensor in the domain could be sampled in the analysis with equal probability. For each sample the mean temperature deviation, which is the difference in mean air temperature compared to the original mean temperature, was determined. In Fig. 4, the distributions of the mean temperature deviations are shown for Case 1. In other cases, similar distributions were obtained (not shown). The deviation indicates the error of estimating mean temperature with random sampling. In Fig. 4, all the data are symmetrically distributed and show a high density near zero deviation according to central limit theorem. The ranges are larger when the sample numbers are smaller. When the sampling ratio is 0.9, most of the deviations are within a +0.05°C range, whereas for the 0.7 sampling ratio, most of the deviations are within a +0.1°C range. Thus, a reduction in the network of 30% is acceptable if a 0.1°C deviation is allowable for the application at hand. In this example, such a reduction corresponds to a change in mean sensor spacing from 7.1 to 8.5 km (assuming the whole area is 10000 km^2 and there are 200 original measurement points).

For all 16 cases, the standard deviation of mean temperature across all sampling ratios (each with 1000 values) shows a clear decrease in size as the number of sensors increases (Fig. 5). The maximum standard deviation, when the sampling ratio is s= 0.1, varies from 0.21 to 0.51°C in summer, and 0.16 to 0.37°C in winter. The case with the largest standard deviation in both summer (Case 7, Fig. 5a) and winter (Case 13, Fig. 5b) is consistent with the time when there was the largest standard deviation observed in the original data (Table 2). When the standard deviation of mean temperature is divided by the original standard deviation as shown in Fig. 5c, the values of all cases are very similar. This shows the temperature distributions of histograms (Fig. 2) were similar although their spatial patterns were different.

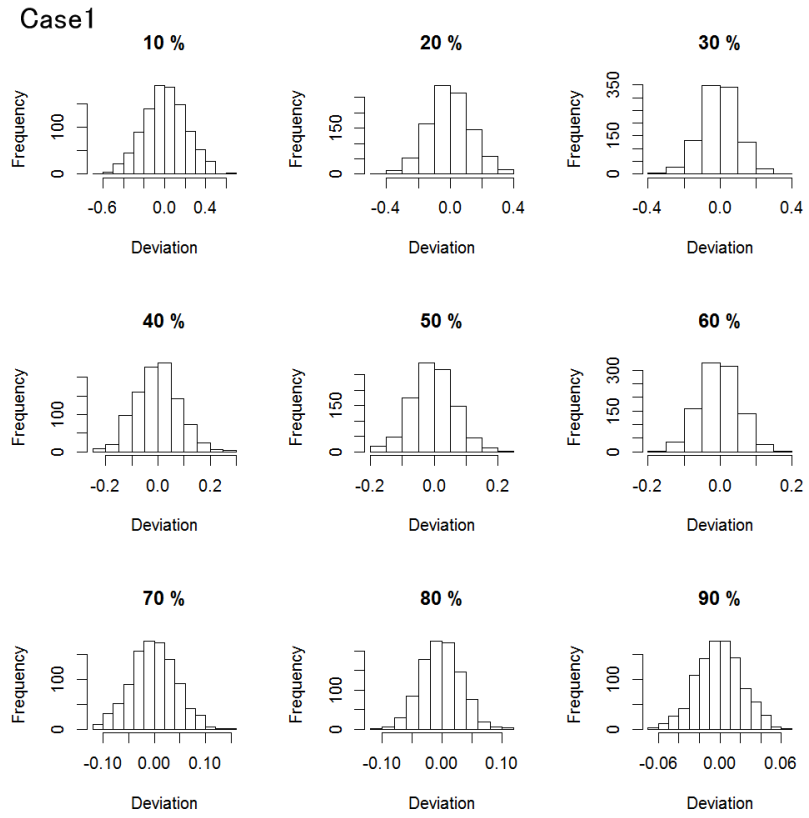


Fig. 4 Histograms of temperature deviation (sampled mean temperature - original mean temperature) (°C) in random sampling (Case 1, Table 1).Compare with Fig. 2a (case 1) when 100 % of the data are used to show the actual values. Frequency is the number of a particular temperature deviation with all subplots having 1000 samples. The number above each subplot indicates the percentage of sensors (from the original 183) randomly sampled to be analysed for statistics to be calculated.

The maximum change (decrease) of standard deviation between sampling ratios occurs consistently when the ratio increases from 0.1 and 0.2, which doubles the number of sensors analysed. If a certain standard deviation is specified as the desired accuracy of estimation of mean temperature, the appropriate sampling ratio or reduction of the data in each case from Fig. 5 can be selected.

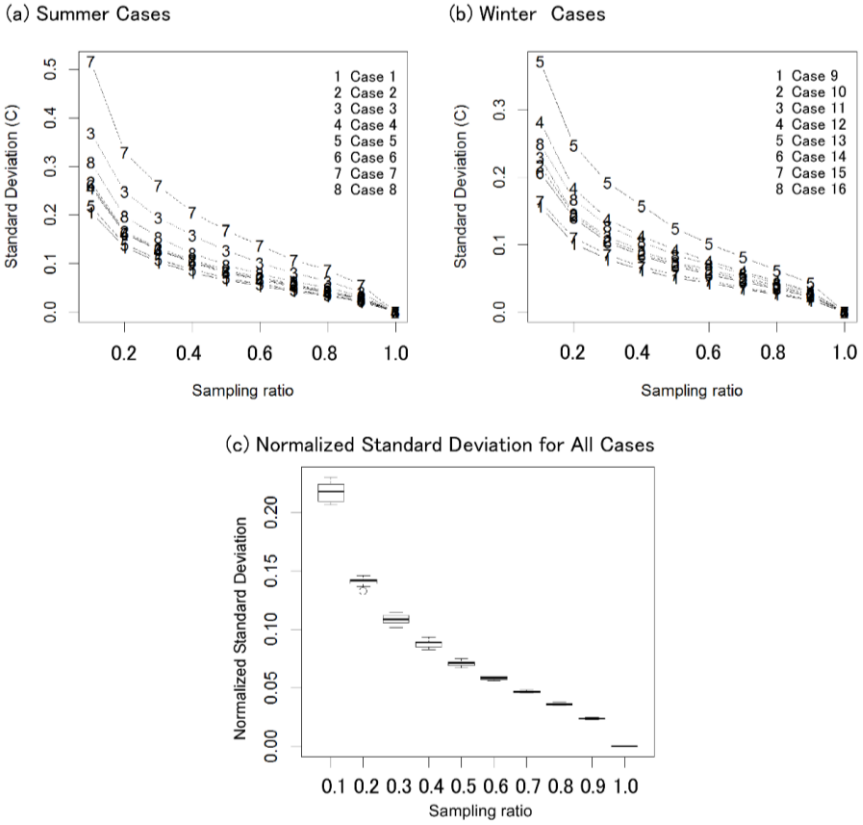


Fig. 5 Standard deviation (°C) of mean temperature (from the 1000 samples) and normalized standard deviation (the standard deviation of mean temperature divided by the original standard deviation) with changing proportion of sensors present. Numbers on lines refer to the case number. See Table 2 for more details.

3.2 IDW method with random sampling

The IDW interpolated images when all sensors are included (Fig. 6) show clearly the spatial distributions of the temperature. The location of the UHI maxima can be seen to change with time (Table 2, Fig 6). For each case (Table 2), sampling ratio and random sampling combination, a map of the air temperature pattern can be created.

With random sampling here, this creates 144,000 maps (16 cases x 9 sampling ratios x 1000 samples), thus it is not possible to show even close to 0.1 % of these. Moreover, there are many more possible combinations beyond this subset of the 1000 per c/s; with a population of n sensors with a sampling ratio s, there are

^{n}Csn combinations[1]:

$$^{n}C_{sn}=[n!]/\{[n-(sn)]!\,(sn)!\} \quad (8)$$

In addition to the images, the correlations (eq. 6) and the RMSEs (eq 7) were calculated.

Thus, it is possible to identify the best and worst correlations for a case and sampling ratio. In Fig.7, for Case 1 (Table 2) with a [1]sampling ratio of 0.2 (c/s= 1/0.2), despite the low sampling rate, there are images with high correlations (>0.9) which are visually very similar to the original image (compare to Fig. 6a- Case 1).

[1] Note with n=30 and s=0.1 there are 4060combinations

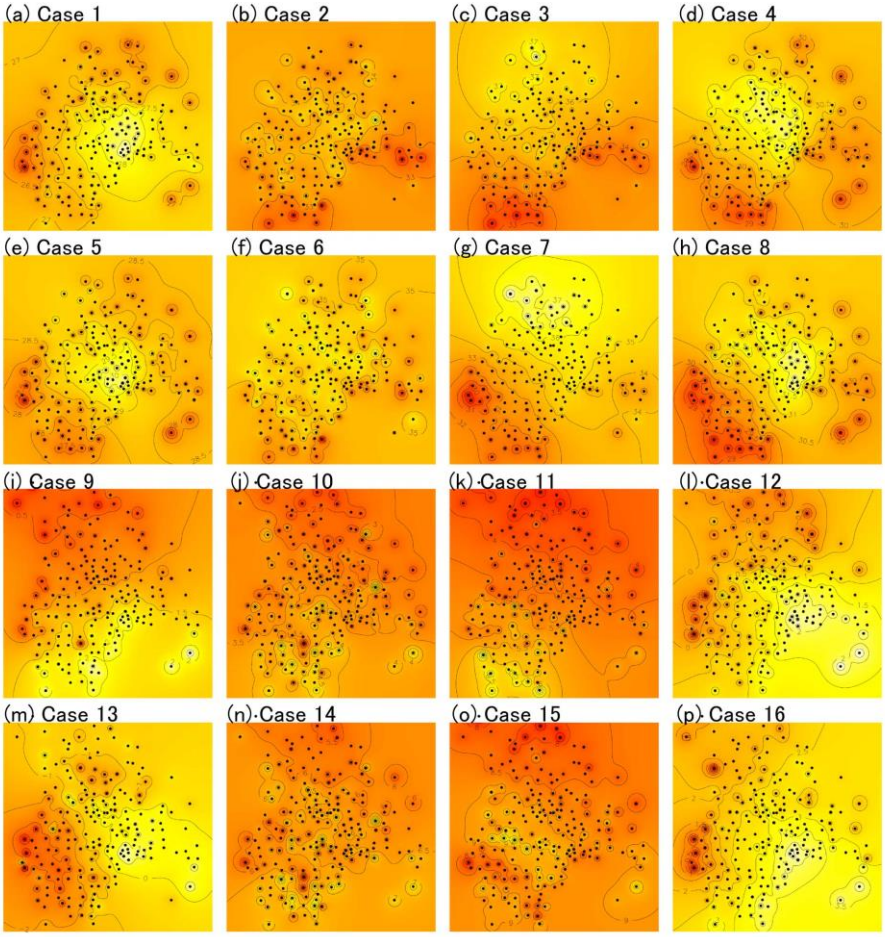


Fig. 6 Spatial pattern of air temperature based on inverse distance weighted (IDW) using all available sensors for the 16 cases (see Table 2 for more details of the cases). The shading from red to yellow indicates change of temperature from low to high. See Fig.1 for location of Tokyo Bay and sensor locations etc.

Through analysis of all the images with spatially equal probability sampling, those with the highest or best correlations compared to those with the poorest correlation or biased sampling can be identified (e.g. Fig. 7). To review all 1000 combination analysed for a c/s set histograms of the correlation and the RMSE statistics, along with the relation between the two are used (e.g. Fig. 8). This allows determination of whether the "best" or "worst" (e.g. Fig. 7) outcomes are unlikely or not. For c/s=1/0.2 (Fig 8a, b) and c/s=9/0.2 (Fig. 8d, e), we see that the correlation is positively skewed and the RMSE negatively skewed. The relation between the correlation and RMSE is strong, with little scatter above a clear lower limit (Fig. 8c, f). In all c/s sets similar distributions and relations are observed. Thus it is possible to make some estimate of what the poorest relation maybe independent of the specific combination of sites used. This allows a probable error estimate across the network to be determined.

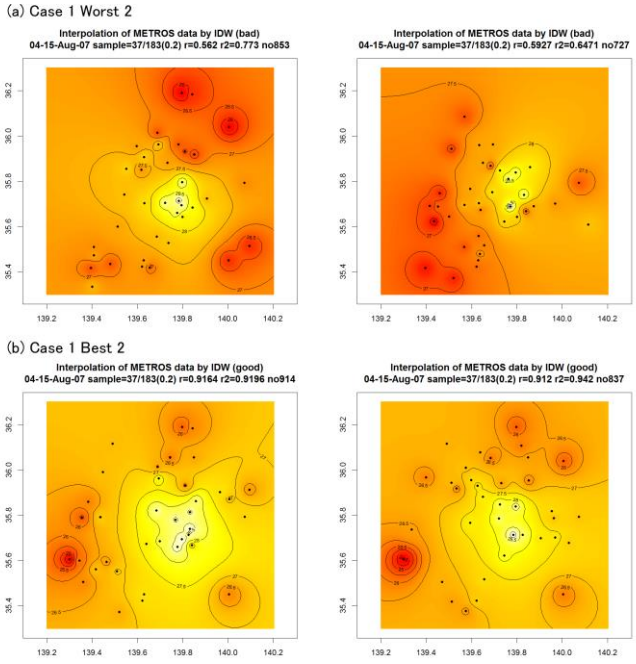


Fig. 7 As for Fig. 6 but for Case 1 with a sampling ratio of 0.2, with four of the possible 1000 randomly (homogenous) selected sensors points. The four are chosen based on the statistical analysis of their properties relative to Fig.6a pixels using equations 6 and 7. They represent the two best (highest correlation) and worst (lowest correlation).

Summarizing this for all the c/s sets, for all cases higher correlations and lower RMSEs were evident for higher sampling ratios (Fig. 9a, b). The four lowest correlations all occur at 10:00 (c=2, 6, 10 and 14, Fig 9a) when no clear temperature maximum area (heat island) existed (Table 2, Fig 6).
If a threshold of acceptable (good) similarity for the correlation of 0.9 is chosen, this would suggest a sampling ratio of 0.7 is the lower limit (Fig. 9a), which corresponds to a mean distance between sensors of 12.9 to 13.5 km as sufficient. Given there are no stations located in Tokyo Bay currently, the land based separation is effectively smaller (section 2.1). Even at the sampling ratio 0.2, the correlations were sometimes more than 0.9. The network can be considered to be very robust to the loss of measurement data in these cases. Given the nature of the sites and instruments used in this network, this suggests it is measuring local scale features well, but not necessarily micro-scale variability (vertical and horizontal) in and around each of the sites.

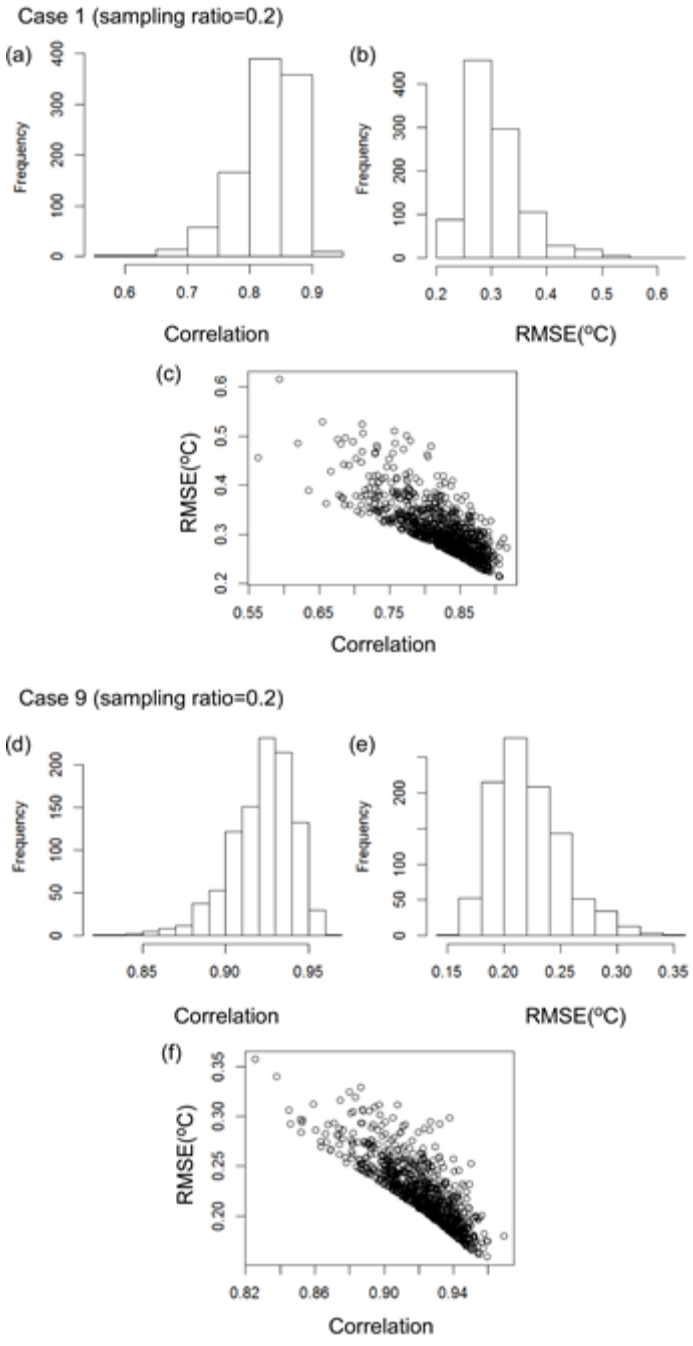


Fig. 8 Example of the statistics (correlation (-), RMSE (°C)) determined from analysis of the IDW spatial domains for two c/s (a-c) Case 1 sampling ratio=0.2 and (d-f) Case 9, sampling ratio: 0.2). Histograms of (a,c) correlation (eq. 6) and (b,d) RMSE (eq 7) plus the relation between the (c,f) correlation and RMSE.

Visual evaluation by human eye is complicated, but when the correlation is more than 0.9, the images are recognizably very similar visually. If the correlation is 0.8, the images show good visual similarity with the original images. The correlation can be used as a criterion of spatially good selection of the data.

In Fig. 9b, the largest mean RMSE across all the c/s is 0.8°C, which was associated with c/s= 7/0.1 for the summer afternoon situation (16:00). This was the case studied with the largest spatial standard deviation (Table 2) with only 18 sensors used to characterize the pattern. Across all cases, selecting the 0.7 (s) as above, results in a RMSE of about 0.2 °C or less.

3.3 IDW method with hierarchical clustering and sampling

In virtually all urban areas there are known variations in topography (e.g. changes in elevation, distance of sites to water bodies like lakes or sea, etc.) which will affect meteorological variables such as temperature. Knowledge of these should guide sensor network design and installation. In this study the presence of large water bodies (Tokyo Bay and the Pacific Ocean) and hills surrounding Tokyo (Fig. 1) are expected to impact the thermodynamics of the city's atmosphere, as well as constrain where the city has and can develop.

In the prior section, it was possible for all sensors randomly selected to be clustered together in a relatively small area. For example, with 37 sensors selected in c/s= 1/0.2, the poorest points (lowest correlation) in Fig. 7a or Fig. 8c likely reflected that situation. In the following analysis, the same methods are used as section in 3.2 except that the sampling is once with hierarchical clustering for each s and c. Selection of sensors was required to include geo-location of a specified density based on s as shown in Fig. 10. Thus a distributed pattern at the larger scale (related to s) is forced and this ensures that the stations are spread across the spatial domain. Sensors could still cluster in certain areas but maintain an overall coverage across the domain (Fig 10c, d).

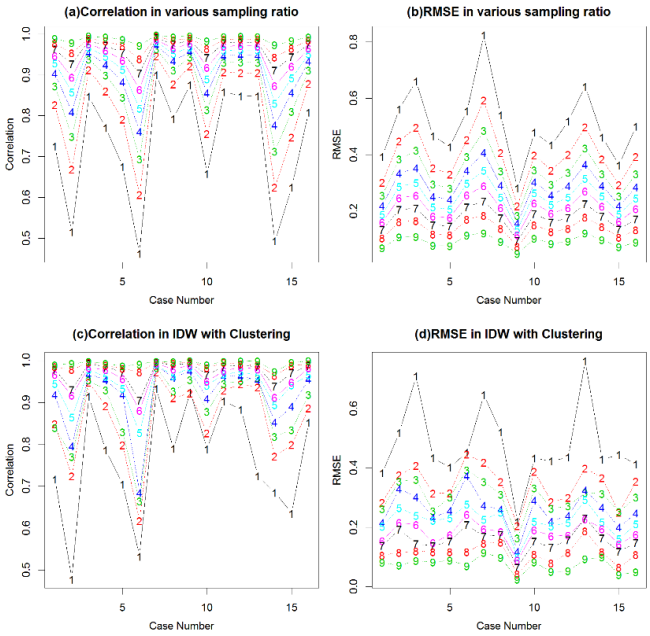


Fig. 9. The mean (a) correlation and (b) RMSE for each c/s based on the 1000 samples. The lines join a sampling ratio (1 is 0.1, 2 is 0.2 and 9 is 0.9) across the cases (Table 1). The results of IDW with hierarchical clustering are also shown in (c) correlation and (d) RMSE.

Fig.11 allows direct comparison between the IDW interpolated images and one example of clustered selection c/s=1/0.1, 1/0.2, 1/0.5 (and equivalently for cases 2, 9, 10). First, the placement of sensors selected for analysis can be compared between the original (Fig.6a, b, i, j or Fig. 11 row 4 c/s =1/1), randomly sampled (Fig.7 c/s= 1/0.2) and the sampling with clustering (Fig.11 row 2 c/s=1/0.2) schemes. With the hierarchical clustering, the sensors selected are forced to cover the

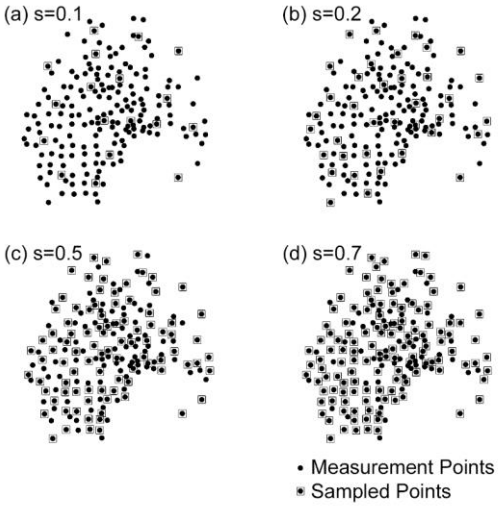
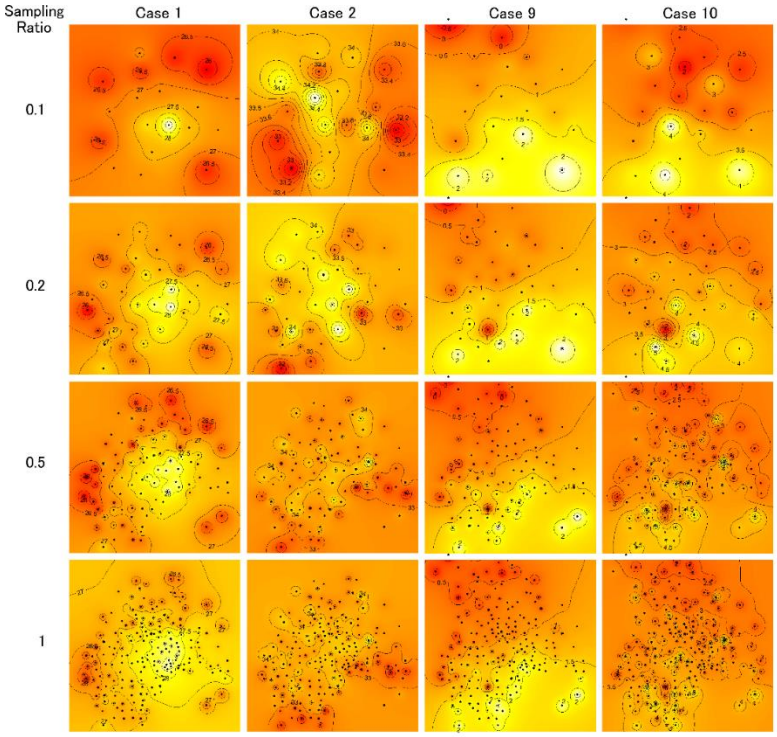


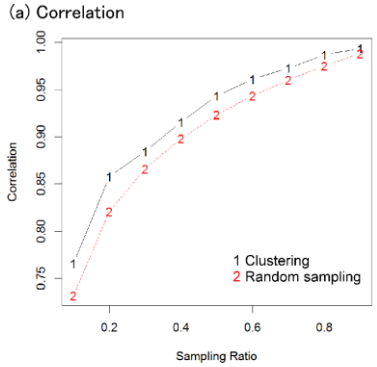
Fig. 11 As for Fig. 6 but for Cases 1, 2, 9 and 10 with sampling ratio of (row 1) 0.1, (row 2) 0.2, (row 3) 0.5 and (row 4) 1.0 (all/original) and images made by IDW with selected points by sampling with hierarchical clustering. For rows 1-3 of Case 1 and Case2, selected sensors points shown in Fig. 10 are used.

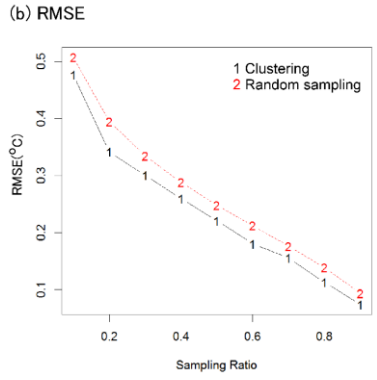
whole domain at each sampling ratio. Images of sampling ratio 0.5 and 1 are similar in all cases. Even at the sampling ratio 0.2, some images are similar to the original images.

Fig. 10 Sampled points from measurement points by sampling with hierarchical clustering at the sampling ratio, s of 0.1, 0.2, 0.5 and 0.7 in summer Cases (n=183) with the same method explained in Fig. 3.



Following the approach in section 3.2, after analysis of all the c/s combinations, the mean correlations and RMSEs are summarized in Fig. 9c, d. The characteristics are very similar to the randomly sampled results of previous section (Fig. 9a, b), with higher correlations and lower RMSEs with higher sampling ratios. The cases with high/low values were almost the same using both methods (Fig. 9). At the sampling ratio 0.7, the correlations are over 0.9 and RMSEs are less than 0.2°C for all cases. Again, this means very similar IDW images were obtained with 30 % data loss. Even at the sampling ratio 0.5, the correlations are over 0.8 and IDW images are very similar to the original one (Fig. 9, 11).





3.4 Implications for designing an urban air temperature network

From this analysis of the two sampling strategies (random and clustering), it is clear that the sampling with hierarchical clustering improves the network's capability to reproduce the full network results with fewer sensors (Fig. 12). The correlations are larger and RMSEs lower than with only random sampling. This indicates the hierarchical clustering is effective in guiding where sensors should be located.

Fig. 12 Comparison of the Correlation and the RMSE in random sampling and sampling with hierarchical clustering.

These results inform how we should approach air temperature network design in the future. Many meteorological networks evolve drawing on stations originally installed for a wide variety of purposes (e.g. case of London, Grimmond, 2013). Techniques, such as those presented here, allow us to evaluate the spatial data recorded by such networks, and importantly how they can best be enhanced. Also, given it is inevitable that instruments will fail, need to be replaced and site access lost, the sampling methods and analyses in this study that introduces some redundancy and resilience to the network, may be useful from the outset of network design in guiding better use of resources. Through ongoing use of the type of analyses presented here, QA/QC (quality assurance/quality control) processes also can be enhanced as this approach can also be used as a tool to identify sensors in need of attention. Here only one type of sampling with hierarchical clustering has been considered (based on location). However, it is highly likely that further gains can be made drawing on more detailed metadata (e.g. Muller et al., 2013a; Muller et al.,2013b) and siting characteristics (e.g. Stewart and Oke, 2012) . This will become even more important as a wider spectrum of sensors types are used (e.g. smart phones, Overeem et al., 2013) and mobile settings (Mahoney et al., 2010) which will yield temporally and spatially more dynamic data. This will enhance data for both near real-time (e.g. energy related decision making) or long-term climate (e.g. architectural and engineering design) users.

4. Conclusions

Extended METROS data gathered in Tokyo were used to analyze the effect of reducing observation points from an existing meteorological measurement network. For the analysis of estimation of mean temperature and interpolation with IDW, random sampling was undertaken. Generally, it was possible to obtain almost the same result with a 30% reduction in the number of sensors. IDW with hierarchical clustering was used to select points homogeneously.

The original network instrument siting in Tokyo, sensors in elementary schools across the city, capture local scale variations in temperature, but not detailed vertical and horizontal variability at each of the sites. Thus this analysis has allowed us to look at local scale features. From this analysis, a mean instrument spacing of the order 13 km (at the time of 30% reduction) allows the key features of Tokyo's temperature domain to be reproduced. As all cities have their own distinct geographical setting, which will impact the evolution of the urban heat island, we suggest that methods presented in this study like application of IDW with hierarchical clustering improves the ability to characterize spatial sampling. It allows either improved performance with the same number of sensors or sustains the same level of performance with fewer sensors (Fig. 12).

The techniques used here allow assessments of networks as new sensors are added or existing networks combined, to ensure robust and resilient networks not dependent on individual sensors. With the development of open data policies and crowdsourcing of meteorological data (e.g. Weather Observation Website, http://wow.metoffice.gov.uk/; Weather Signal, http://weathersignal.com/; AMeDAS, http://www.jma.go.jp/jma/indexe.html, etc.) such techniques will allow adaptive but rigorous use of available data.

References

Basara, J. B., Illston, B. G., Fiebrich, C. A., Browder, P. D., Morgan, C. R., McCombs, A., Bostic, J. P., McPherson R. A., Schroeder, A. J. & Crawford, K. C. (2010). The Oklahoma City Micronet. *Meteor. Appl., 18*, 252–261.

Baume, O. P., Gebhardt, A., Gebhardt, C., Heuvelink, G. B. M., &Pilz, J. (2011). Computers & Geosciences Network optimization algorithms and scenarios in the context of automatic mapping.*Computers and Geosciences*, *37*(3), 289–294.

Brus, D., Gruijter, J. De, &Groenigen, J. Van. (2006). Designing spatial coverage samples using the K-means clustering algorithm.*Developments in Soil Science*, *31*, 183–192.

Brus, D. J., &Heuvelink, G. B. M. (2007). Optimization of sample patterns for universal kriging of environmental variables. *Geoderma*, *138*(1-2), 86–95.

Chang, B., Wang, H.-Y., Peng, T-Y. & Hsu, Y-S. (2010). Development and evaluation of a city-wide wireless weather sensor network. *Educational Technology & Society 13*, 270–280.

Chaplot, V., Darboux, F., Bourennane, H., Leguédois, S., Silvera, N. & Phachomphon, K. (2006). Accuracy of interpolation techniques for the derivation of digital elevation models in relation to landform types and data density. *Geomorphology, 77*, 126–141.

Davies, L., Bell, J. N. B., Bone, J., Head, M., Hill, L., Howard, C., Hobbs, S. J., Jones, D. T., Power, S. A., Rose, N., Ryder, C., Seed, L., Stevens, G., Toumi, R., Voulvoulis, N., White, P. C. L. (2011). Open air laboratories (OPAL): a community-driven research programme. *Environmental Pollution*, 159(8–9): 2203–2210.

Grimmond, C. S. B. (2013). Observing London: Weather data needed for London to thrive. London Climate Change Partnership, 81pp http://climatelondon.org.uk/wp-content/uploads/2013/07/Observing-London-Final.pdf, Last accessed 10 August 2014.

Heuvelink, G., Brus, D., &Gruijter, J. de. (2007). Optimization of sample configurations for digital mapping of soil properties with universal kriging.*Developments in Soil Science*, *31*, 137–151.

Hicks, B. B., Callahan, W. J., Pendergrass, W. R. III, Dobosy, R. J., Novakovskaia, E. (2012). Urban turbulence in space and time. *Journal of Applied Meteorology and Climatology*, 51, 205–218.

Hung, T. K. & Wo, O. C. (2012). Development of a community weather information network (Co-WIN) in Hong Kong. *Weather, 67*, 48–50.

Jarvis, C. H. & Stuart, N., (2001).A Comparison among Strategies for Interpolating Maximum and Minimum Daily Air Temperatures. Part II: The Interaction between Number of Guiding Variables and the Type of Interpolation Method. *Journal of Applied Meteorology, 40*, 1075-1084.

Koskinen, J. T., Poutiainen, J., Schultz, D. M., Joffre, S., Koistinen, J., Saltikoff, E., Gregow, E., Turtiainen, H., Dabberdt, W. F., Damski, J., Eresmaa, N., Göke, S., Hyvärinen, O., Järvi, L., Karppinen, A., Kotro, J.,Kuitunen, T., Kukkonen, J., Kulmala, M., Moisseev, D., Nurmi, P., Pohjola, H., Pylkkö, P., Vesala, T.,Viisanen, Y. (2011). The Helsinki Testbed: a mesoscale measurement, research, andservice platform. *Bull Amer. Meteor. Soc., 92*, 325-342.

Kravchenko, A. N. (2003). Influence of Spatial Structure on Accuracy of Interpolation Methods. *Soil Science Society of America Journal, 67*, 1564–1571.

Mahoney, B., Drobot, S., Pisano, P., McKeever, B., and O'Sullivan, J. (2010) Vehicles as Mobile Weather Observation Systems. *Bull. Amer. Meteor. Soc., 91*, 1179–1182.

Mikami, T., Ando H., Morishima W., Izumi T., & Shioda T. (2003). A new urban heat island monitoring system in Tokyo. Proc. Fifth Int. Conf. on Urban Climate, Lodz, Poland, International Association for Urban Climate. http://nargeo.geo.uni.lodz.pl/~icuc5/text/O_3_5.pdf, Last accessed 10 August 2014.

Muller, C. L., Chapman,L., Grimmond, C.S.B., Young D. T., & Cai, X. M. (2013a). Toward a Standardized Metadata Protocol for Urban Meteorological Networks. *Bull. Amer. Meteor. Soc., 94*, 1161–1185.

Muller, C. L., Chapman L., Grimmond, C. S. B., Young D. T., & Cai, X. M. (2013b). Sensors and the city: a review of urban meteorological networks. *Int. J. Climatol., 33*, 1585–1600.

Mueller, T. G., Mijatovic, B., Sears, B. G., Pusuluri, N. & Stombaugh, T. S., (2004). Map quality for ordinary kriging and inverse distance weighted interpolation. *Soil Science Society of America Journal, 68*, 2042–2047.

Murty, R. N., Mainland, G., Rose, I., Chowdhury, A. R., Gosain, A., Bers, J. & Welsh, M. (2008). CitySense: an urban-scale wireless sensor network and testbed. In Proceedings of the 2008 IEEE International Conference on Technologies for Homeland Security, Waltham, MA. 583-588.

Nalder, I.A.&Wein, R.W. (1998). Spatial interpolation of climatic normals: test of a new method in the Canadian boreal forest. *Agriculture and Forest Meteorology, 92*, 211-225.

Overeem, A. Robinson, J. C. R. Leijnse, H., Steeneveld, G. J., Horn, B. K. P. and Uijlenhoet, R. (2013). Crowdsourcing urban air temperatures from smartphone battery temperatures. Geophysical Research Letters, 40, 4081–4085.

Perera, C., Zaslavsky, A., Christen, P. & Georgakopoulos, D. (2014). Sensing as a service model for smart cities supported by Internet of Things, *Trans. Emerging Tel. Tech., 25*, 81–93.

Stewart, I. D. (2011). A systematic review and scientific critique of methodology in modern urban heat island literature, *Int. J. Climatol., 31*, 200 – 217.

Stewart I. D. & Oke, T. R. (2012). Local Climate Zones for Urban Temperature Studies.*Bull. Amer.Meteor. Soc., 93*, 1879–1900.

Tan, J., Yang, L., Grimmond, C. S. B., Shi, J., Gu, W., Chang, Y. & Han, Z. (2014). Urban Integrated Meteorological Observations: Practice and Experience in Shanghai, China. *Bulletin of the American Meteorological Society*. doi:10.1175/BAMS-D-13-00216.1

Van Groenigen, J. W., Siderius, W., & Stein, A. (1999). Constrained optimisation of soil sampling for minimisation of the kriging variance. *Geoderma, 87*, 239–259.

Walvoort, D. J. J., Brus, D. J., & de Gruijter, J. J. (2010).An R package for spatial coverage sampling and random sampling from compact geographical strata by k-means.*Computers& Geosciences, 36*, 1261–1267.

Weber, D. D. & Englund, E. J. (1992). Evaluation and comparison of spatial interpolators. *Math.Geology, 24*, 381–391.

Yamato H., Takahashi H., Mikami T. (2009) New urban heat island monitoring system in Tokyo metropolis. The seventh International Conference on Urban Climate, 29 June-3 July 2009,

Yokohama,Japan.http://www.ide.titech.ac.jp/~icuc7/extended_abstracts/pdf/375391-1-090516000046-006.pdf, Last accessed 10 August 2014.

Zimmerman, D., Pavlik, C., Ruggles, A., & Armstrong, M. P. (1999). An experimental comparison of ordinary and universal kriging and inverse distance weighting. *Mathematical Geology, 31*, 375-390.