Network optimization of urban heat island measurements -Effect of reduction of observation points-

Tsuyoshi Honjo[1], Kiyoshi Umeki[1], Hiroaki Yamato[2], Takehiko Mikami[3], C.S.B. Grimmond[4]

1 *Department of Environmental Science and Landscape Architecture, Chiba University*

2 Chuo University

3 Tokyo Metropolitan University

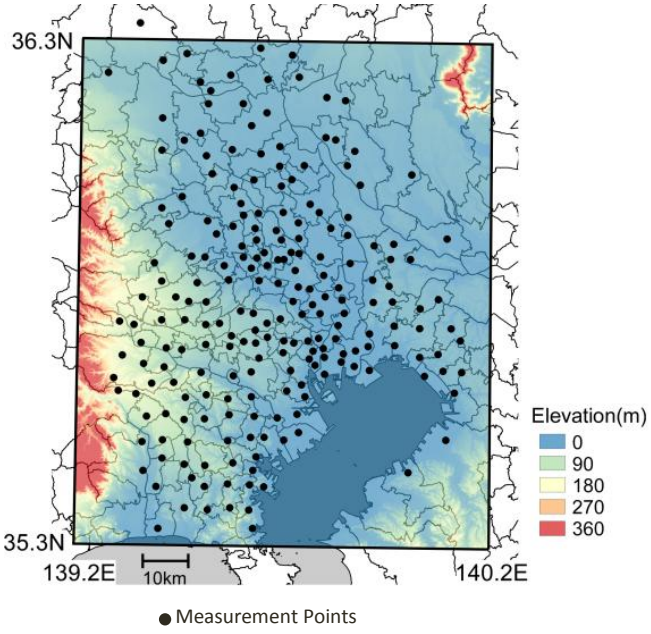
4 Department of Meteorology, University of Reading

Network for heat island measurement

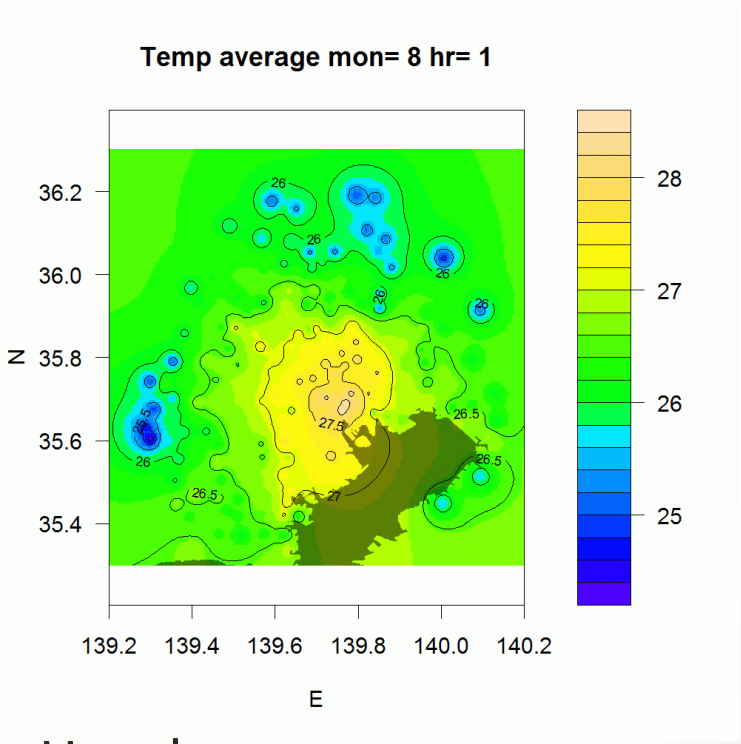
- A larger number of measurement points are better. But considering the labor and cost, a smaller number of points would be better if the same result can be obtained.
- What is the optimal number of measurement points and what is their optimal distribution?
- Whether sensors should be substituted or not, if instruments are out of order.

Objective of the study

- The effect of reducing observation points
- We analyze an existing meteorological measurement network around the Tokyo metropolitan area (Extended METROS).



Measurement points of Extended METROS.



Hourly average temperature in August, 2007.

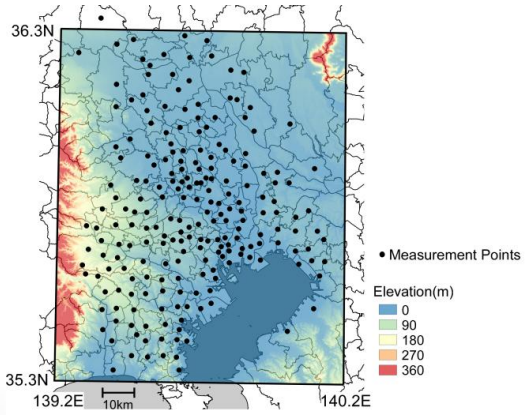
Methods

1. Data of measurement Points
2. Selection of points (10% - 90%) by clustering.
3. Interpolation by IDW (Inversed Distance Weighing)
4. Similarity between the original data and the interpolated data

Data of measurement Points

- The data obtained from May 2007 to October 2008 (18 months, every hour) were used

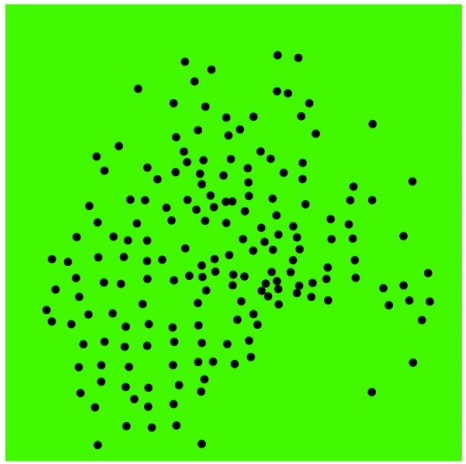
Number of measurement points where no missing data exist in each month.



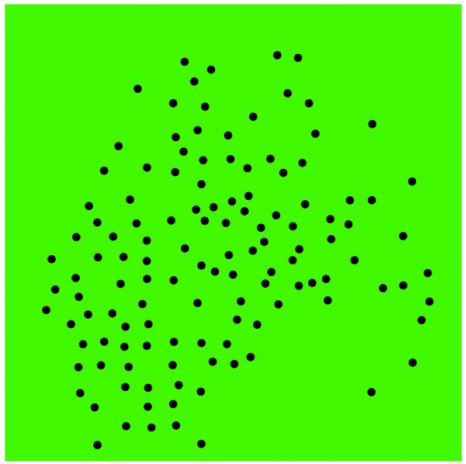
month/year	no. of points	month/year	no. of points
05/2007	193	01/2008	178
06/2007	178	02/2008	190
07/2007	165	03/2008	194
08/2007	180	04/2008	179
09/2007	176	05/2008	165
10/2007	145	06/2008	177
11/2007	180	07/2008	174
12/2007	176	08/2008	163
		09/2008	155
		10/2008	129

Selection of points by clustering

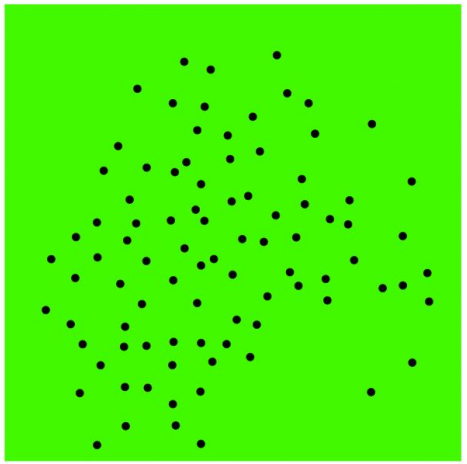
points= 180/180 (100%)



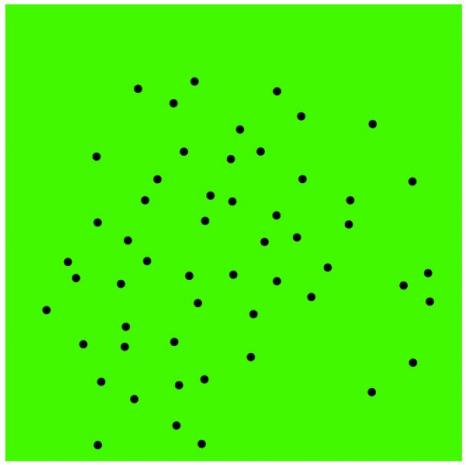
points= 126/180 (70%)



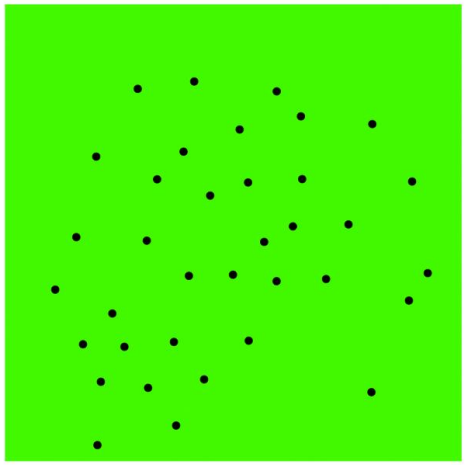
points= 90/180 (50%)



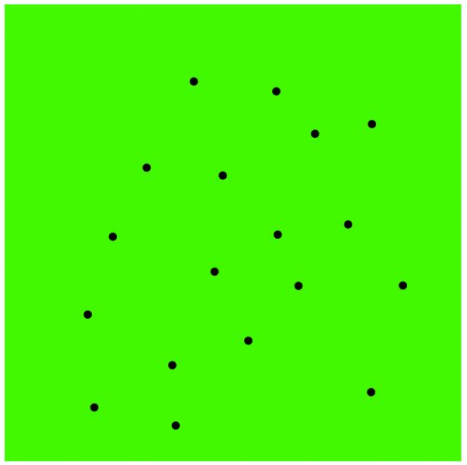
points= 54/180 (30%)



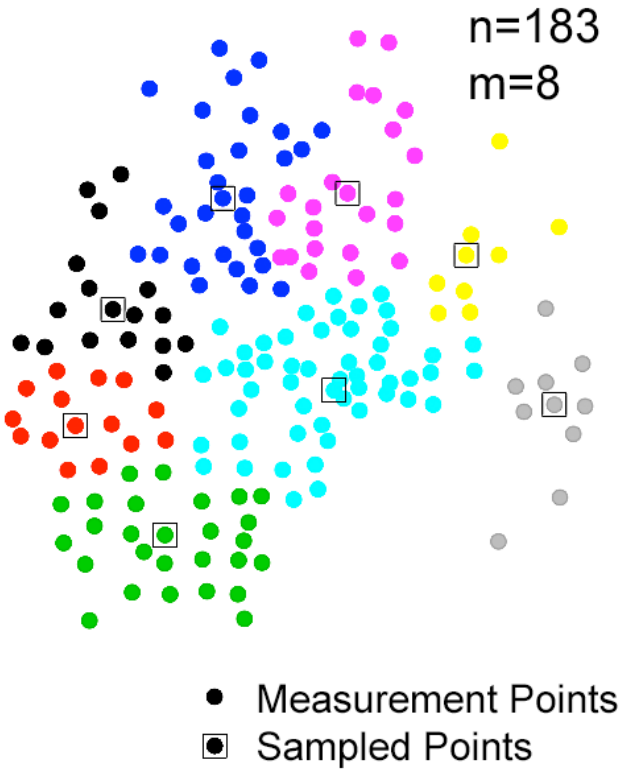
points= 36/180 (20%)



points= 18/180 (10%)



Selection of points (10% - 90%) by clustering



Example of choosing 8 points from 183 by sampling with hierarchical clustering of coodinates. First, 183 points are classified into 8 categories expressed as different colors. Second, center points of each category are selected.

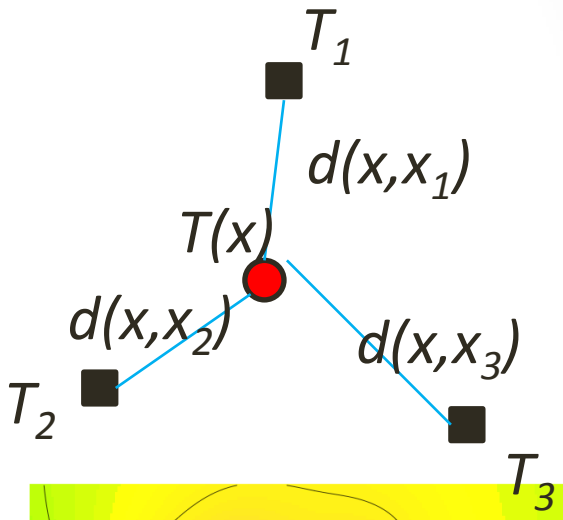
Interpolation by IDW (Inversed Distance Weighing)

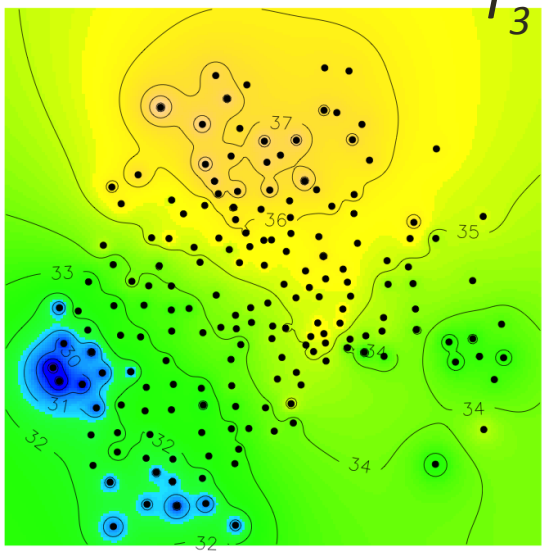
Interpolated temperature *T(x)* was calculated from measured data T_k.

$$\mathrm{T}(x) = \frac{\sum_{k=1}^{m} w_k(x) T_k}{\sum_{k=1}^{m} w_k(x)}$$

weighting function:

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





x : the coordinate vector of an interpolated point,
x_k : measurement point
$d(x,x_k)$: the distance from *x* to x_k,
m : the number of sampled points
P : parameter of IDW (*p=2* is used.)

Similarity between the original data and the interpolated data

To estimate the similarity between two interpolated temperature distribution T_1 and T_2, normalized cross-correlation, R (the correlation) and root-mean-square error (*RMSE*) were used:

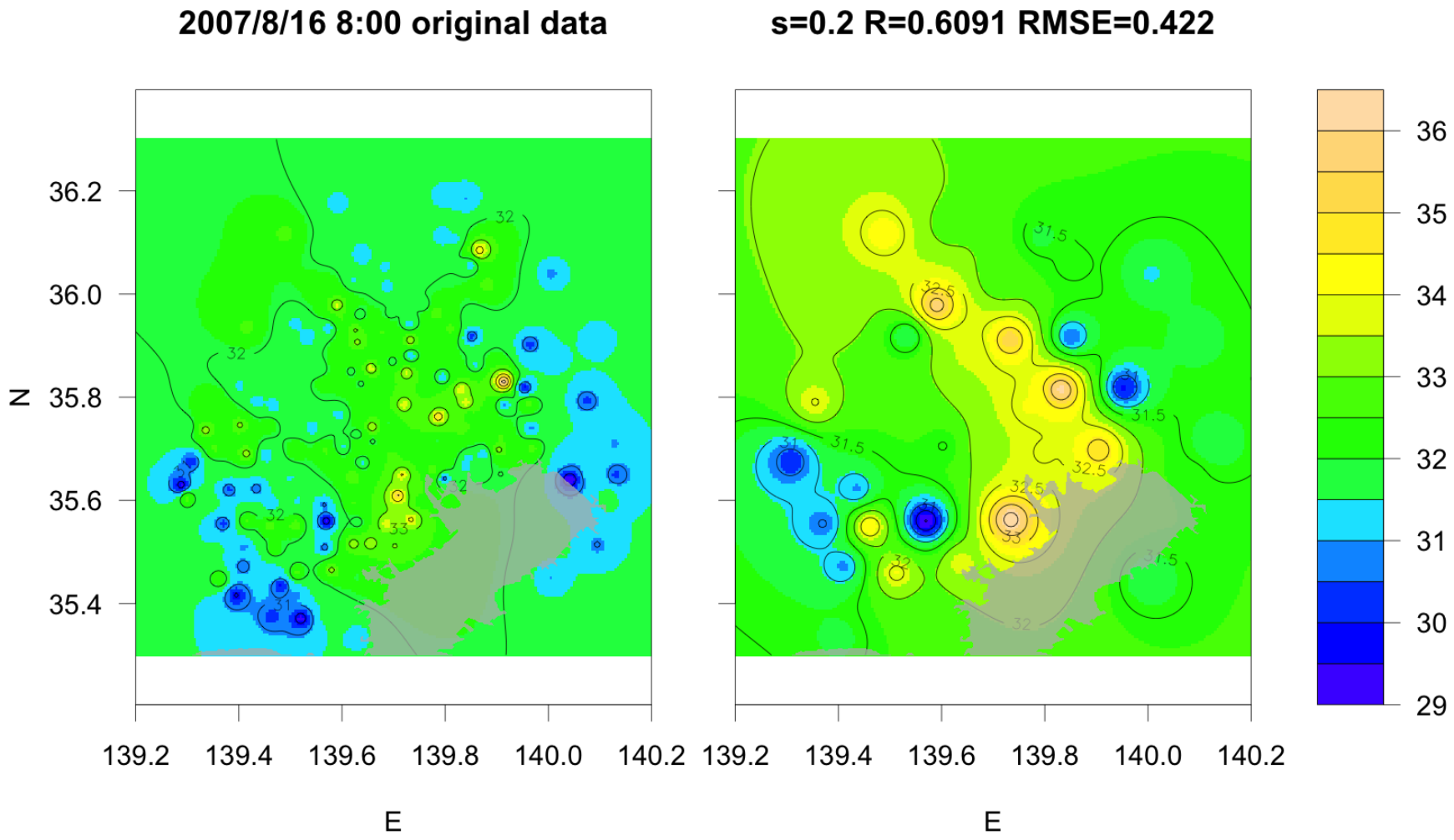
$$R = \frac{\sum_{ix=1}^{N}\sum_{iy=1}^{N}(T_1(ix,iy)-T_{1ave})(T_2(ix,iy)-T_{2ave})}{\sqrt{\sum_{ix=1}^{N}\sum_{iy=1}^{N}(T_1(ix,iy)-T_{1ave})^2\sum_{ix=1}^{N}\sum_{iy=1}^{N}(T_2(ix,iy)-T_{2ave})^2}}$$

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

T_1 *(ix,iy)* :the interpolated values of air temperature from the original data
T_2 *(ix,iy)* :the interpolated values of air temperature from sampled data
ix and *iy* :coordinates of the interpolated image
N :the number of pixels in the *x*-*y* dimension.

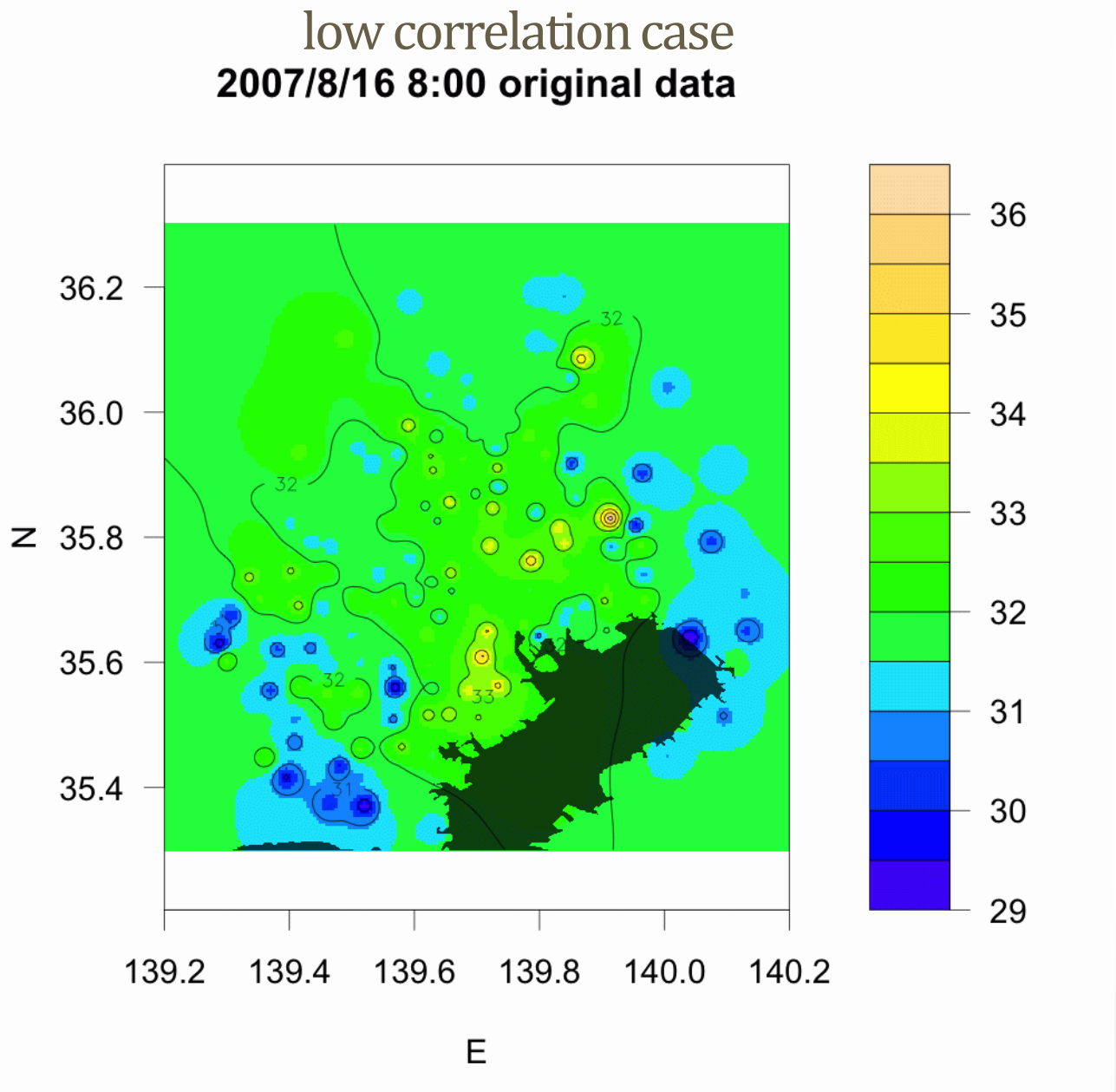
Correlation and RMSE

low correlation case



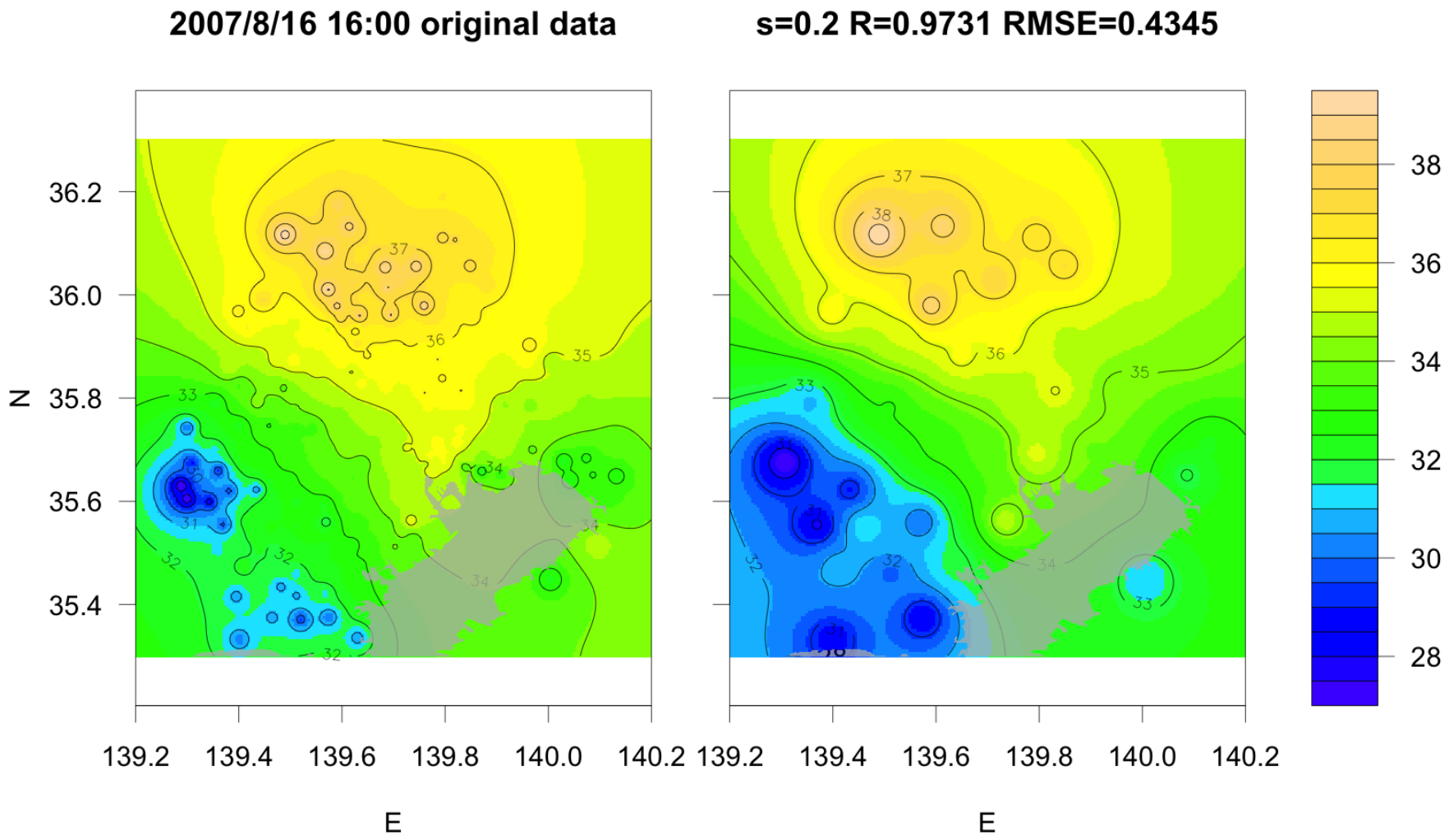
Correlation and RMSE

low correlation case



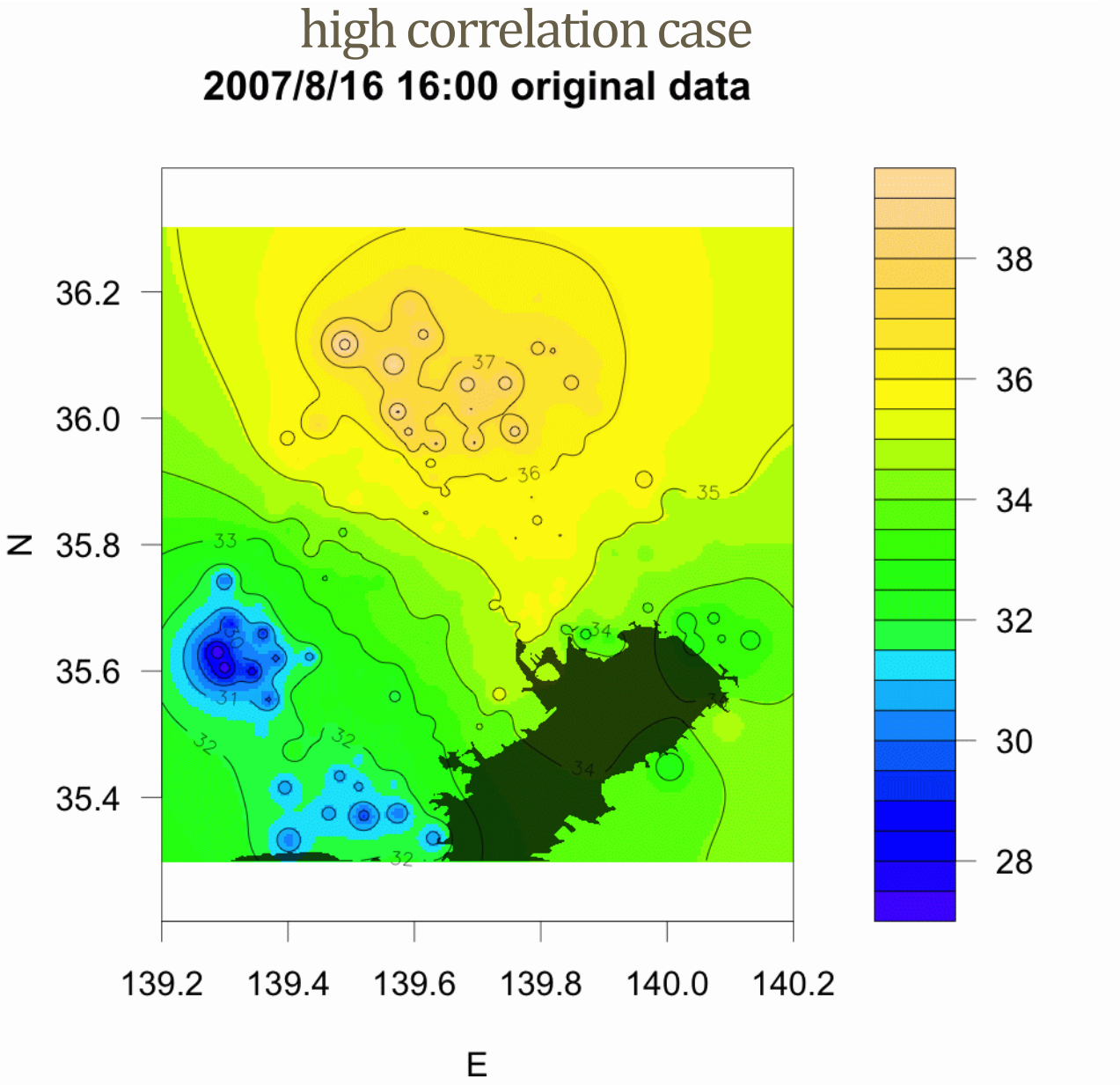
Correlation and RMSE

high correlation case

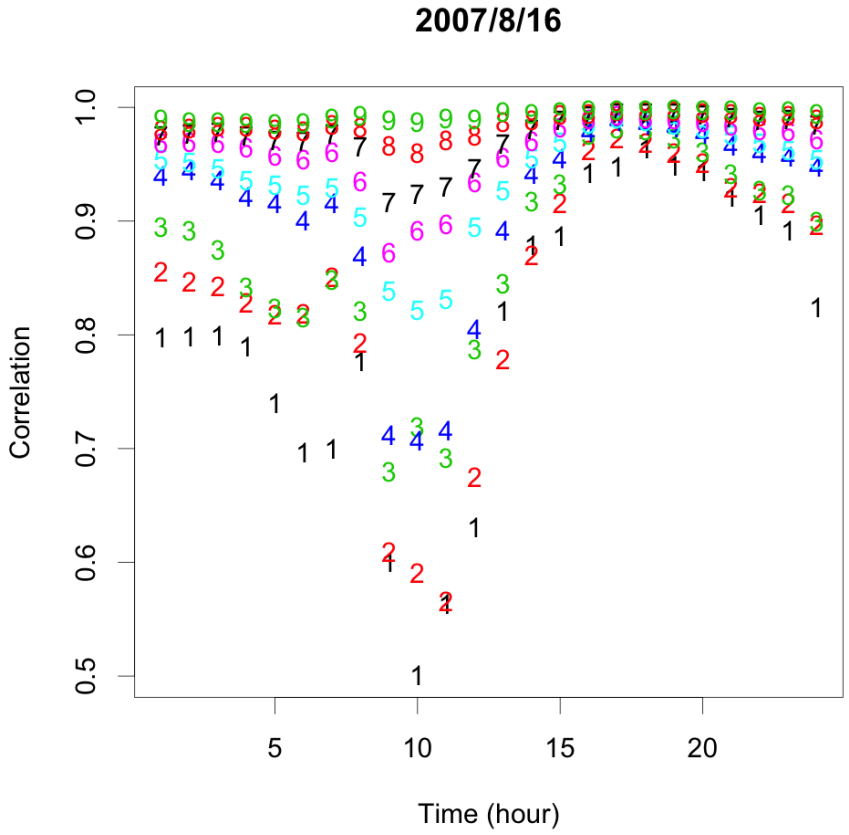


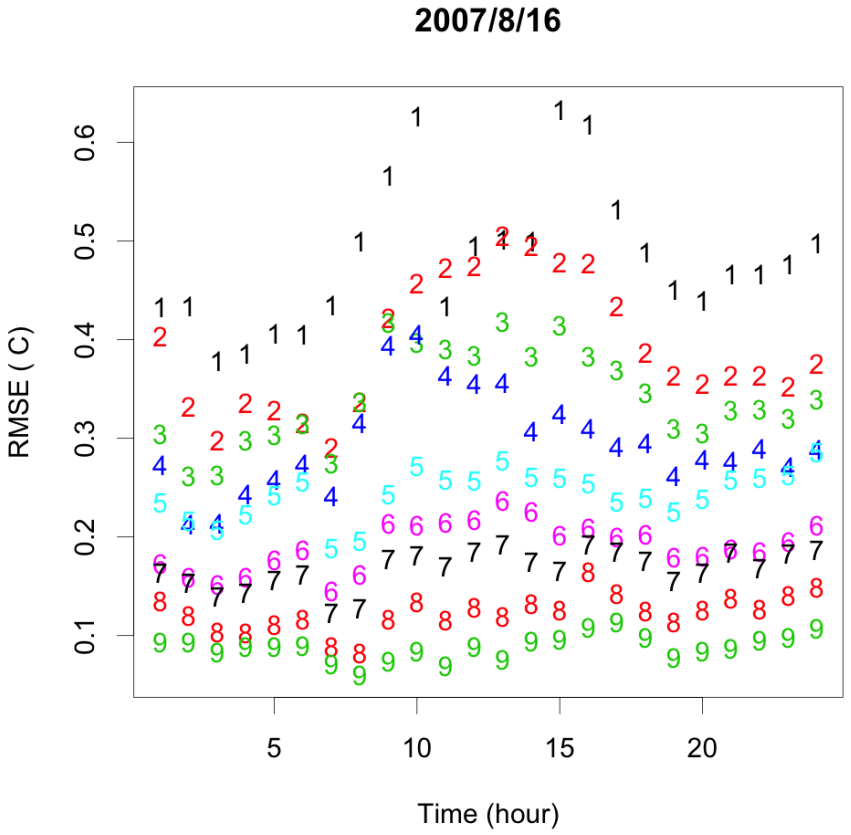
Correlation and RMSE

high correlation case

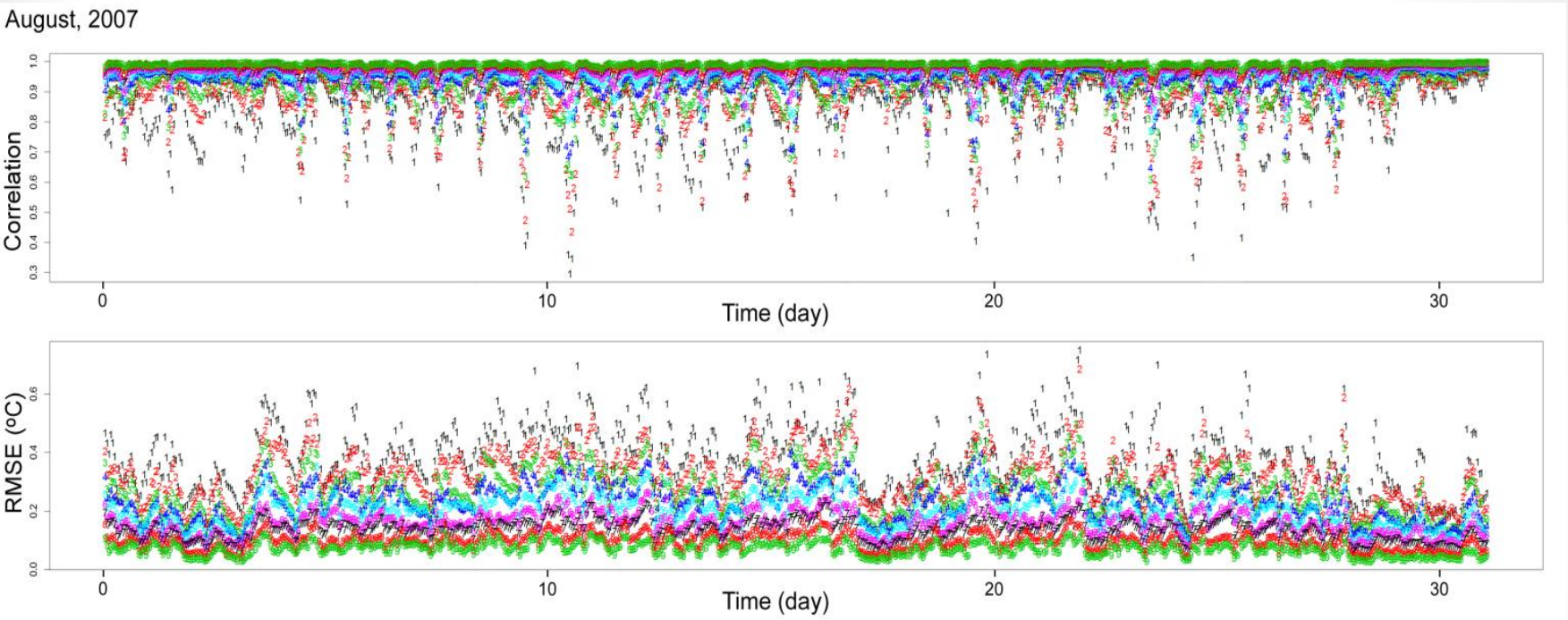


Correlation and RMSE



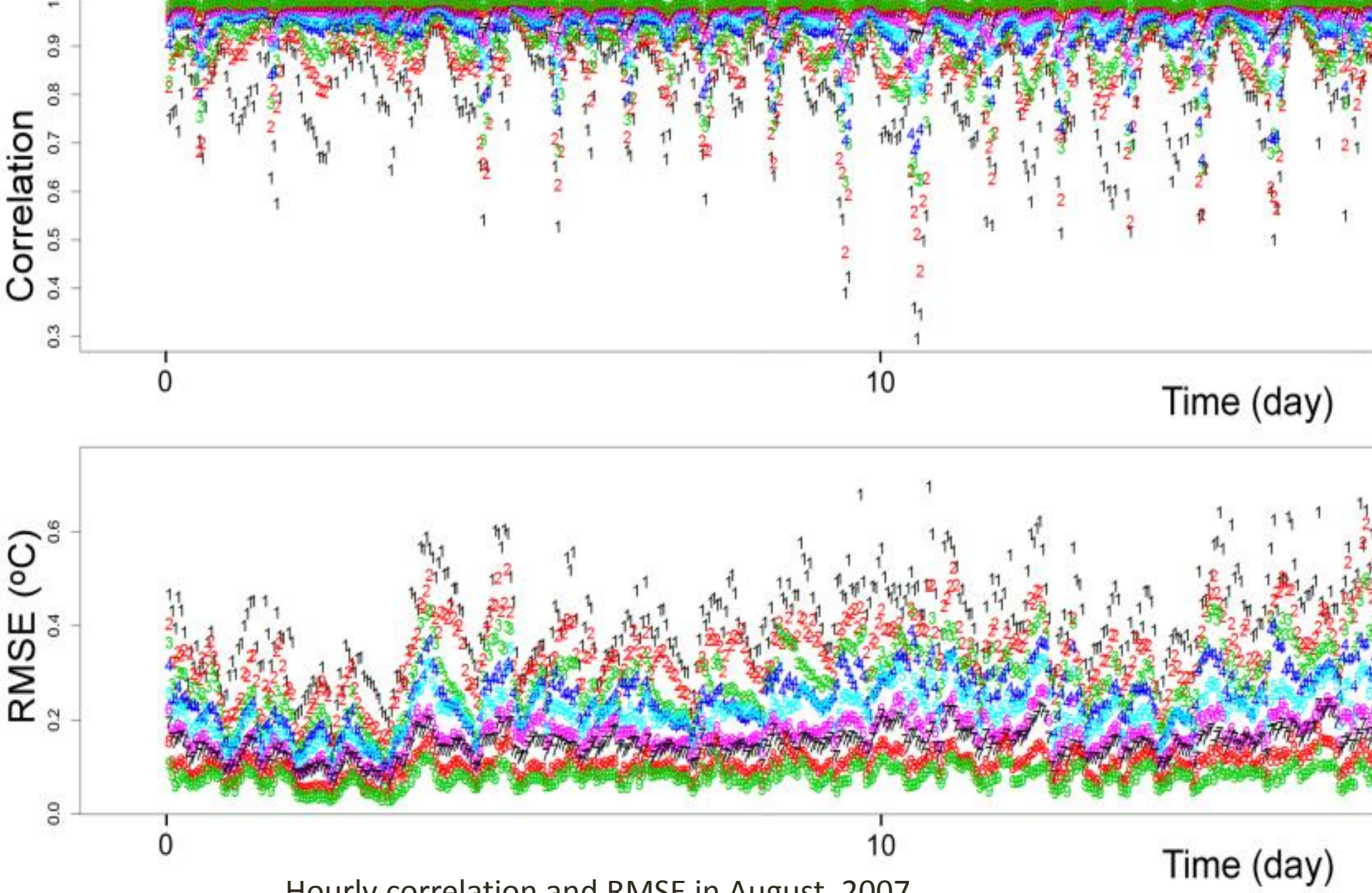


Correlations and RMSE on 16th, August, 2007. The number on the points in the figures corresponds the sampling ratio (1 is 0.1, 2 is 0.2 and 9 is 0.9).



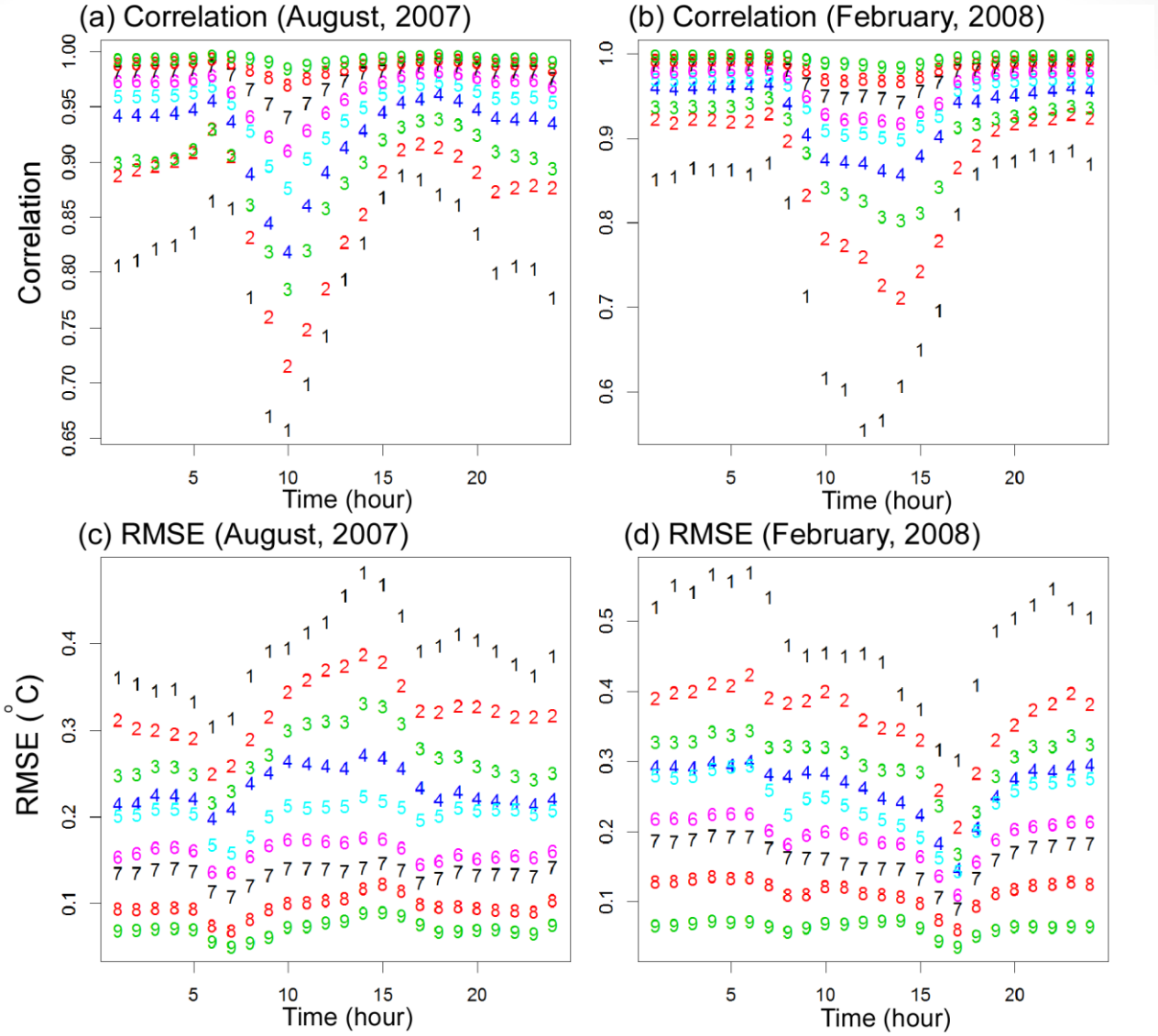
Hourly correlation and RMSE in August, 2007.

The number on the points in the figures corresponds the sampling ratio (1 is 0.1, 2 is 0.2 and 9 is 0.9).



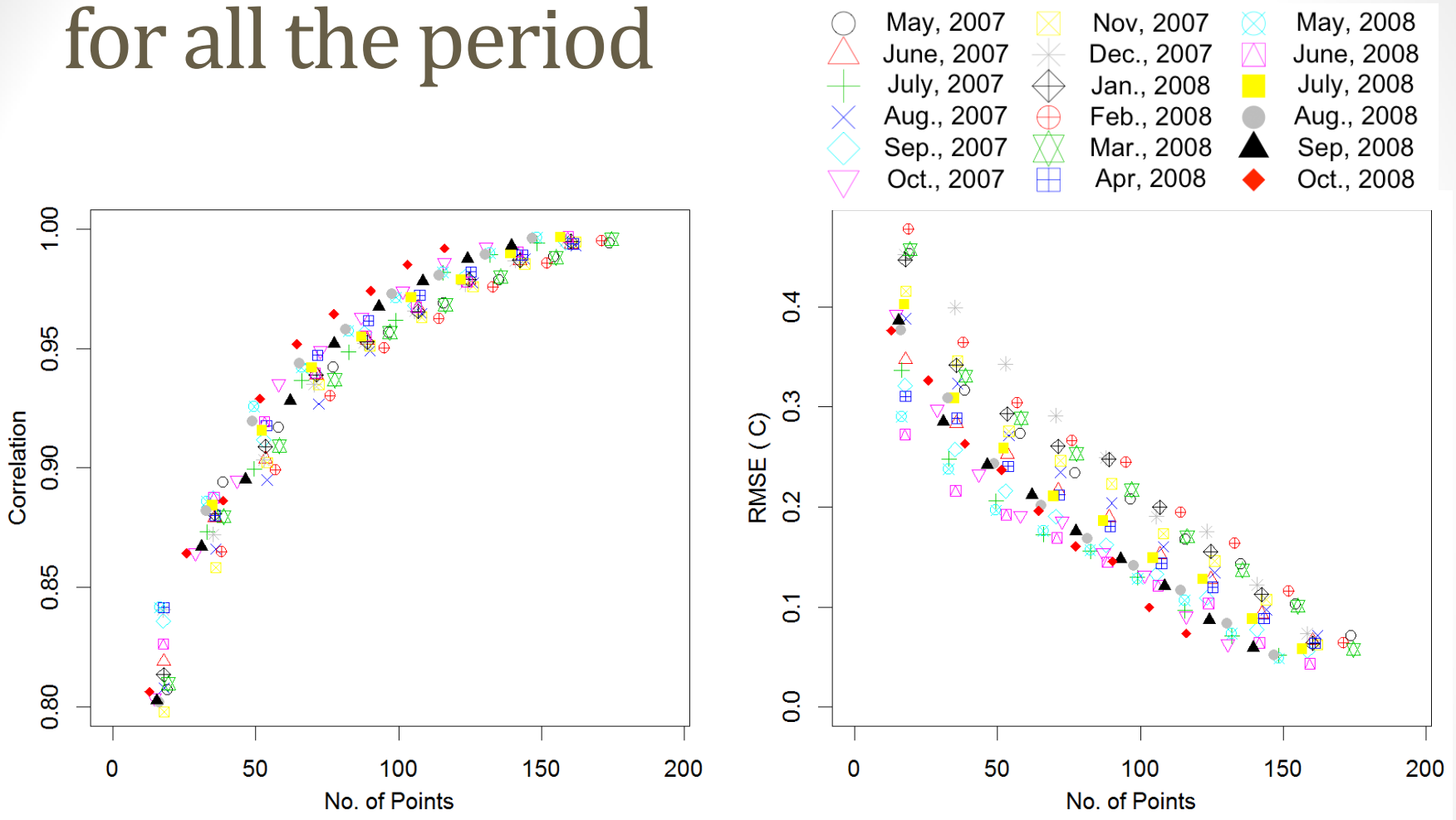
Hourly correlation and RMSE in August, 2007.

The number on the points in the figures corresponds the sampling ratio (1 is 0.1, 2 is 0.2 and 9 is 0.9).



Averaged correlations and RMSE of each hour of day in August, 2007 and February, 2008. The number on the points in the figures corresponds the sampling ratio (1 is 0.1, 2 is 0.2 and 9 is 0.9).

Correlation and RMSE for all the period

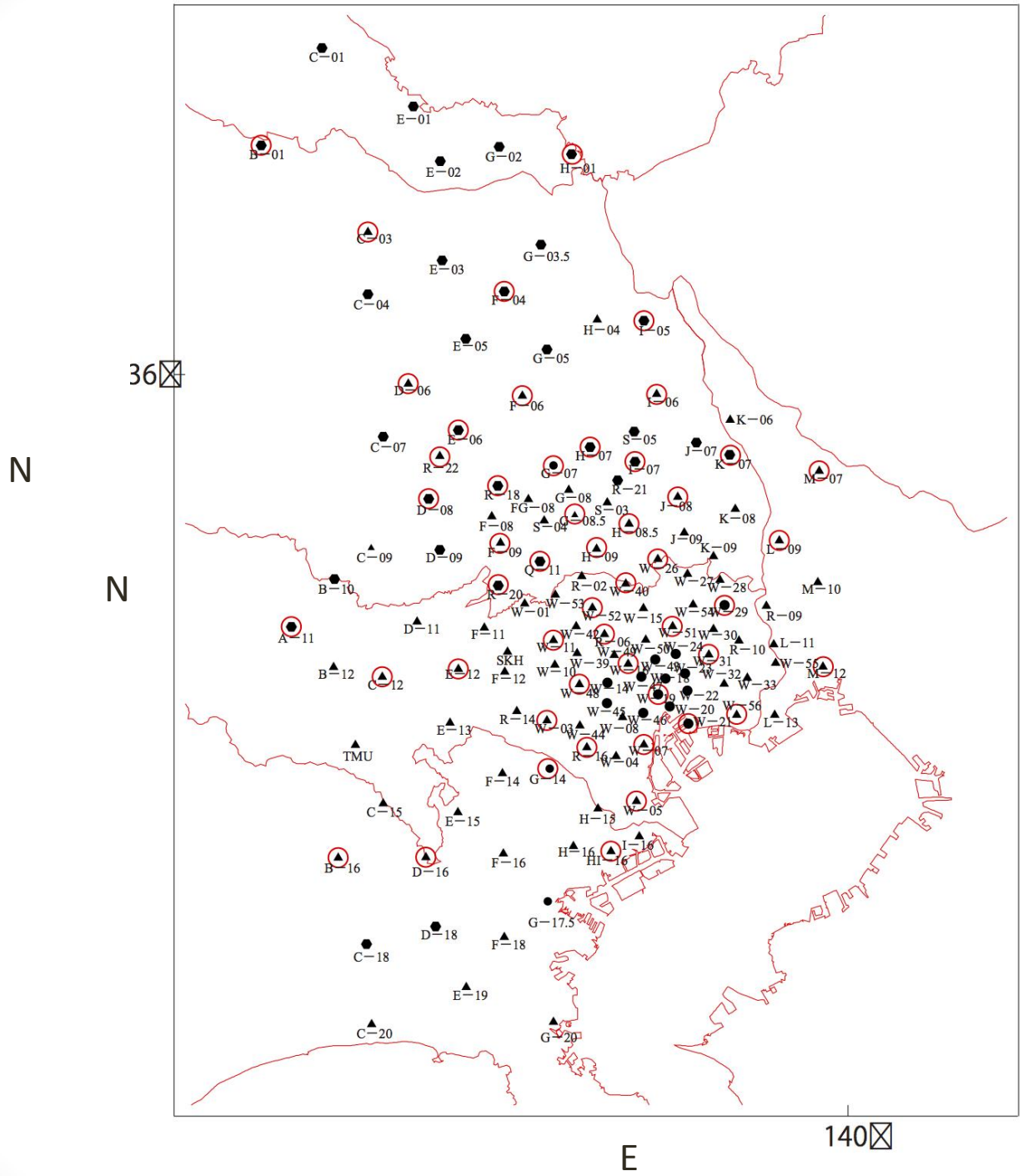


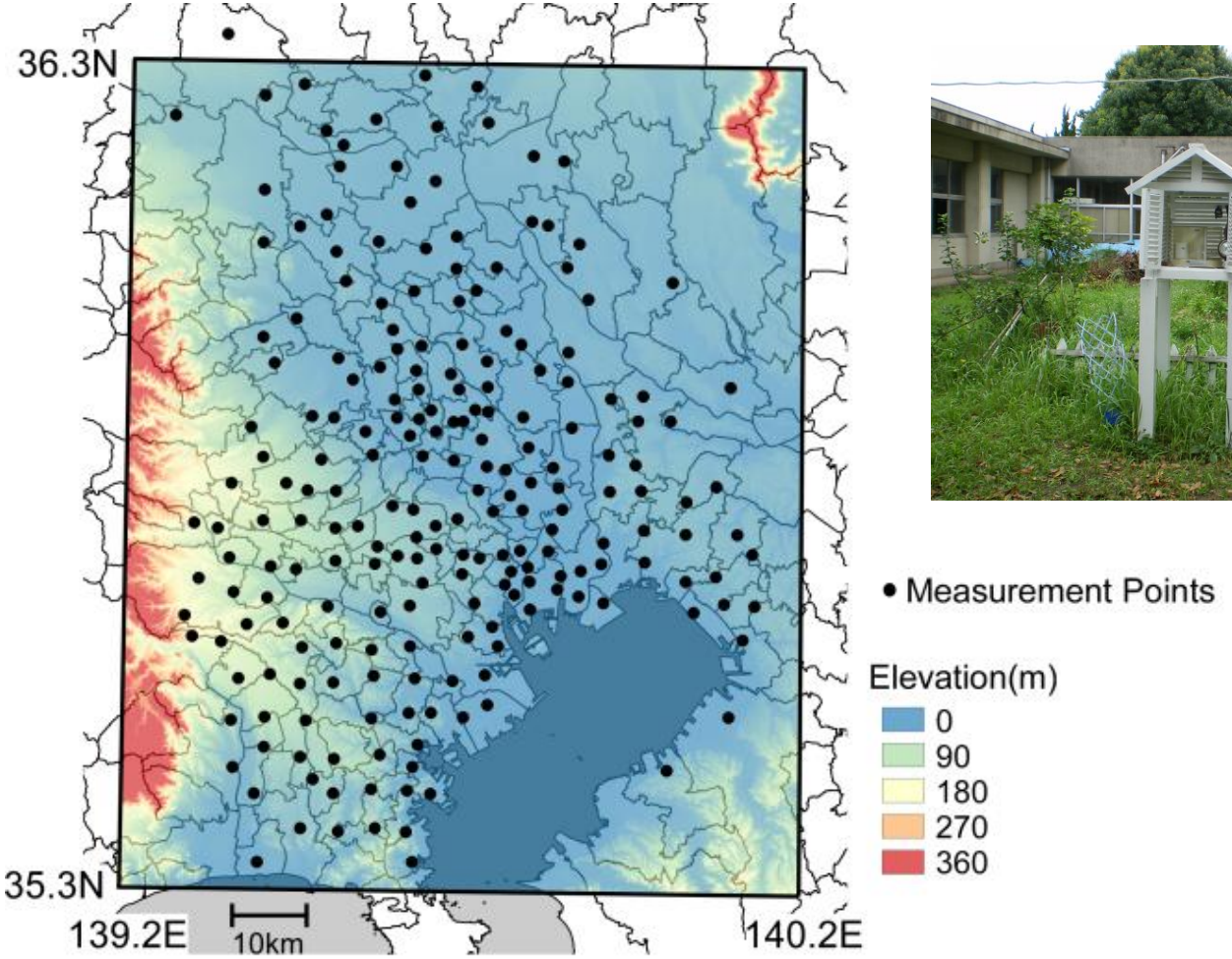
Correlation and RMSE of each month in relation with the number of points for all the period. Marks of the points represents the months.

Conculsion

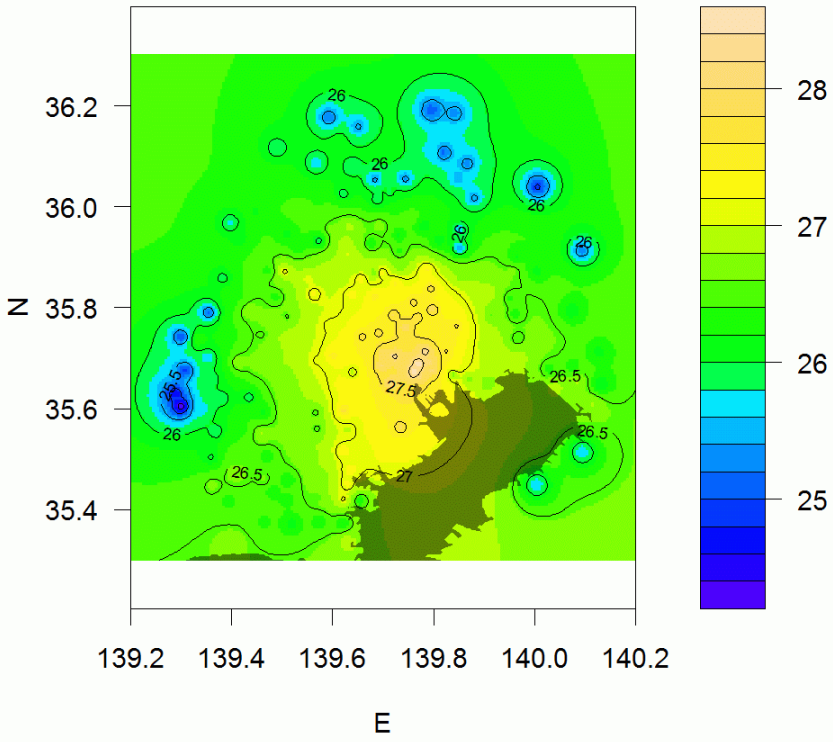
- The methods presented here clearly show the effect of reducing observation points.
- we can find the allowable points considering the limit of the correlation and RMSE.
- The methods in this study can be applied in other networks.

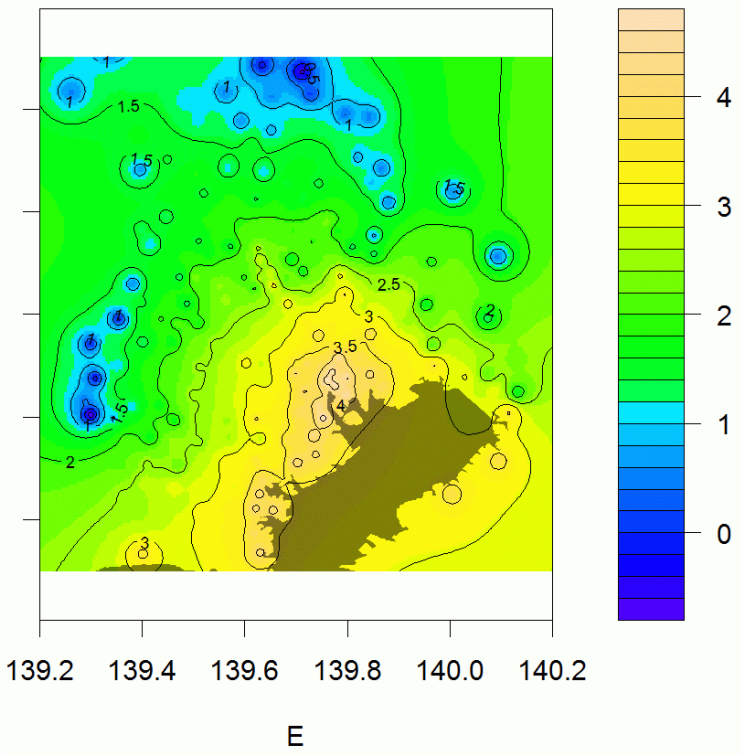
Thank you for your attention.


36
N
N
140
E



The location of measuring points of the Extended METROS network in the Tokyo metropolitan area.


Temp average mon= 8 hr= 1
36.2
36.0
35.8
35.6
35.4
N
139.2
139.4
139.6
139.8
140.0
140.2
E
28
27
26
25


Temp average mon= 2 hr= 1
139.2
139.4
139.6
139.8
140.0
140.2
E
4
3
2
1
0

Methods

Data

The data obtained from May 2007 to October 2008 (18 months, every hour) were used

Measurement Points

Monthly data of the no missing points were used for the further analysis.
10% - 90% points are selected by using the clustering.

Interpolation
Selected data were interpolated by IDW
Grid data of 201 x 201 points were made.

Similarity between the original data and the interpolated data
The correlation and Root-Mean-Square Error (*RMSE*) were used.
Similar images shows relatively high correlation and low RMSE.