



TECHNICAL REPORT NO. 19

Interpolation in laser scanned height data

Allan Aasbjerg Nielsen (Technical University of Denmark)

Allan Aasbjerg Nielsen (Technical University of Denmark)
Interpolation in laser scanned height data

Danish Geodata Agency, Technical Report Series Volume 19

ISBN 978-87-92107-65-7

Technical Report

Published 2013-12-16

This report is available from <http://www.gst.dk>

Note: The first 15 volumes of this series were published under the series title “Technical Reports of the National Survey and Cadastre”.

Interpolation in laser scanned height data



Allan Aasbjerg Nielsen
Technical University of Denmark

Contents

1 Interpolation in laser scanned height data	5
Data	5
Interpolation methods	6
Cross validation (leave one out – LOO)	7
Joint semivariograms for all data	7
2 Conclusions	7
3 Literature	8
4 Appendix: Results from 2007 data	11
Test site A 2007	12
Test site B 2007	17
Test site C 2007	22
Test site D 2007	27
Test site E 2007	32
Test site F 2007	37
Test site G 2007	42
Test site H 2007	47
Test site X 2007	52
Test site Y 2007	57
5 Appendix: Results from 2013 data	62
Test site A 2013	62
Test site B 2013	67
Test site C 2013	72
Test site D 2013	77
Test site E 2013	82
Test site F 2013	87
Test site G 2013	92
Test site H 2013	97
Test site X 2013	102
Test site Y 2013	107

Interpolation in laser scanned height data

This report is in fulfillment for 2013 of “Ydelsesaftale mellem Geodatastyrelsen og Institut for Rumforskning og –teknologi, DTU om myndighedsbetjening af Geodatastyrelsen (tidligere Kort og Matrikelstyrelsen) i form af forskningsbaseret rådgivning og forskning i relation til geodæsi og kortlægning i 2013-2016”, “Resultatmål 3: Metodeudvikling til dataindsamling og topografisk kortlægning, Krav 2”.

Data

The data received from GST consist of laser scanned height data. The twenty data sets delivered cover ten sub-areas sized 100 m by 100 m near Oksbøl (covering part of the military training area there). The figure on the title page shows the area in question.

The lidar reflections are classified and only the ones classified as “terrain” are worked with below.

For each sub-area there is one data set from 2007 and one from 2013. The sampling density for the 2007 data is around 0.5-1.0 points per m² and for 2013 it is 3-4 points per m². Hence a typical distance between observations for the 2007 data is 1 m; for 2013 it is 0.5 m. Also, the average number of neighbours within a 1.0 m search radius is 3-4 for 2007 and around 12 for 2013. Within a 2.0 m search radius the number is around 12 for 2007 and around 50 for 2013.

The following table shows basic statistics including measures of variability for the heights in the twenty areas, namely standard deviation in m and coefficient of variation (the ratio of standard deviation and mean which has no unit). There is little variation in heights in this region (except maybe for region OKSx).

Area	Number of observations	Mean [m]	Min [m]	Max [m]	Standard deviation [m]	Coefficient of variation [-]
OKSa_2007	2,405	10.0270	9.58	11.47	0.32496	0.032407
OKSb_2007	2,031	9.2259	8.39	9.86	0.25584	0.027731
OKSc_2007	3,609	9.0515	8.10	10.67	0.42815	0.047302
OKSd_2007	2,357	9.9004	9.29	10.64	0.21435	0.021651
OKSe_2007	2,719	9.8319	9.43	11.25	0.24181	0.024595
OKSf_2007	4,697	9.9447	9.04	13.49	0.43911	0.044155
OKSg_2007	3,698	9.9299	9.51	10.34	0.12042	0.012127
OKSh_2007	3,946	9.9344	9.49	11.25	0.21794	0.021938
OKSx_2007	2,034	12.3984	9.90	21.22	2.75357	0.222084
OKSy_2007	1,074	23.8013	21.02	26.77	0.99539	0.041821

Area	Number of observations	Mean [m]	Min [m]	Max [m]	Standard deviation [m]	Coefficient of variation [-]
OKSa_2013	36,861	10.0071	9.65	11.48	0.29150	0.029129
OKSb_2013	22,614	9.1945	8.37	10.03	0.25237	0.027448
OKSc_2013	36,560	8.9057	8.13	10.54	0.43624	0.048984
OKSd_2013	13,622	9.9237	9.38	10.50	0.20149	0.020303
OKSe_2013	17,138	9.8593	8.91	11.24	0.24678	0.025030
OKSf_2013	20,915	10.1106	9.06	13.39	0.58424	0.057785
OKSg_2013	21,953	9.9581	9.64	10.50	0.10661	0.010706
OKSh_2013	19,410	9.9144	9.46	10.89	0.18793	0.018956
OKSx_2013	24,606	12.7075	9.90	21.47	3.09036	0.243191
OKSy_2013	11,691	23.5409	21.06	26.83	1.14271	0.048542

In OKSb_2013 observations 22,440, 21,254 and 17,721 were removed manually.

In OKSc_2013 observations 35,247, 30,306, 29,226, 28,191 and 24,919 were removed manually.

Interpolation methods

In this work the following spatial interpolation methods are used

1. Local mean value (LM).
2. Inverse distance (ID).
3. Inverse distance squared (ID2).
4. Ordinary kriging (OK).

The four above methods are performed using 3, 5, 10, 20, 30, 40, 50, 60, 70 and 80 nearest neighbours, respectively. LM corresponds to OK with a full nugget effect semivariogram model.

Also

5. Nearest neighbour interpolation (NN) is used, and
6. Delaunay triangulation followed by bilinear interpolation always based on three nearest neighbours (TIN, triangulated irregular network) is carried out.

The TIN method is often used in GIS based work.

2-D semivariograms (also known as semivariogram maps) estimated individually for each of the ten sub-areas in both years are calculated based on a lag distance of 1.6 m for 2007 and 1.0 m for 2013. These semivariograms mostly exhibit isotropy for small lag distances. Hence also individually (1-D) isotropic (also known as omnidirectional) semivariograms are calculated (based on lag distances of 1.6 m for 2007 and 0.4 m for 2013), modelled and used in OK.

The semivariogram models applied are either spherical (some with the range so large that they appear linear), or modified Gaussian (or modified exponential). The modification consists of using $\exp(-3(h/R)^p)$

with $p = 1.25$, $p = 1.50$, or $p = 1.75$ and not $p = 2$ (or $p = 1$). h is the lag separation distance and R is the range (of influence).

Cross-validation (leave-one-out, LOO)

The quality of the above methods is quantified by a number of measures calculated from leave-one-out (LOO) cross-validation. In LOO cross-validation we exclude one observation x_i at a time (and retain all the other observations), and we then compare the estimated value \hat{x}_i with x_i in a number of fashions, see below. Important measures are bias and root-mean-squared error (RMSE). Bias is defined as

$\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)$, where n is the number of observations. RMSE is defined as the square root of $\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2$. Values for bias and RMSE close to zero are best. Results for the ten sub-areas and both years are shown in the appendix.

Joint semivariograms for all data

It is tempting to estimate joint semivariograms for all the data and use a joint model in all OK estimations. These are shown after the literature list. OK is performed based on individual semivariogram models.

Conclusions

The bias and the RMSE plots consistently show OK as the best or among the best interpolation methods followed by ID2, ID and LM.

The supremacy of OK is confirmed by the histograms and the scatterplots of LOO estimates vs measurements.

The scatter plots show that some of the interpolation methods (sometimes also OK) tend to overestimate low values and underestimate high values.

Runs with different numbers of neighbours used in the interpolations show that generally bias and RMSE do not improve beyond 20-30 neighbours.

Literature

A. G. Journel and C. J. Huijbregts (1978). *Mining Geostatistics*. Academic Press.

E. Isaaks and R. M. Srivastava (1989). *An Introduction to Applied Geostatistics*. Oxford University Press.

C. V. Deutsch and A. G. Journel (1998). *GSLIB: Geostatistical Software Library and User's Guide*, 2nd edition. Oxford University Press.

A. A. Nielsen (2009). [Geostatistics and Analysis of Spatial Data](#). Lecture note.

A. A. Nielsen (2012). Interpolation in laser scanned height data. Technical report.

For LiDAR and photogrammetry based construction and updating of height models in general, see also papers from the annual *Photogrammetrische Woche* and the *European LiDAR Mapping Forum*.

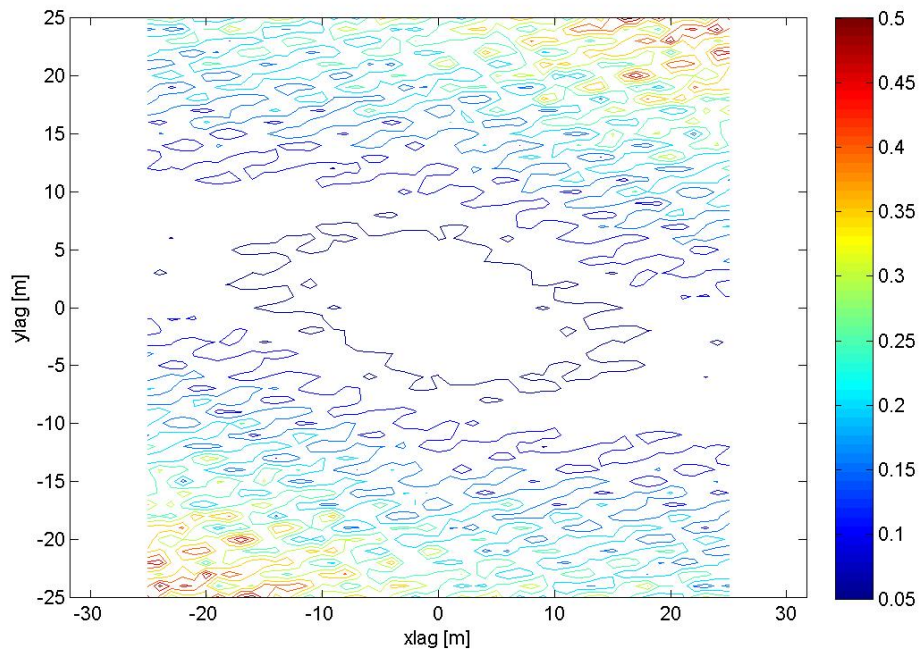


Figure 1: 2-D semivariogram for all 2007 data.

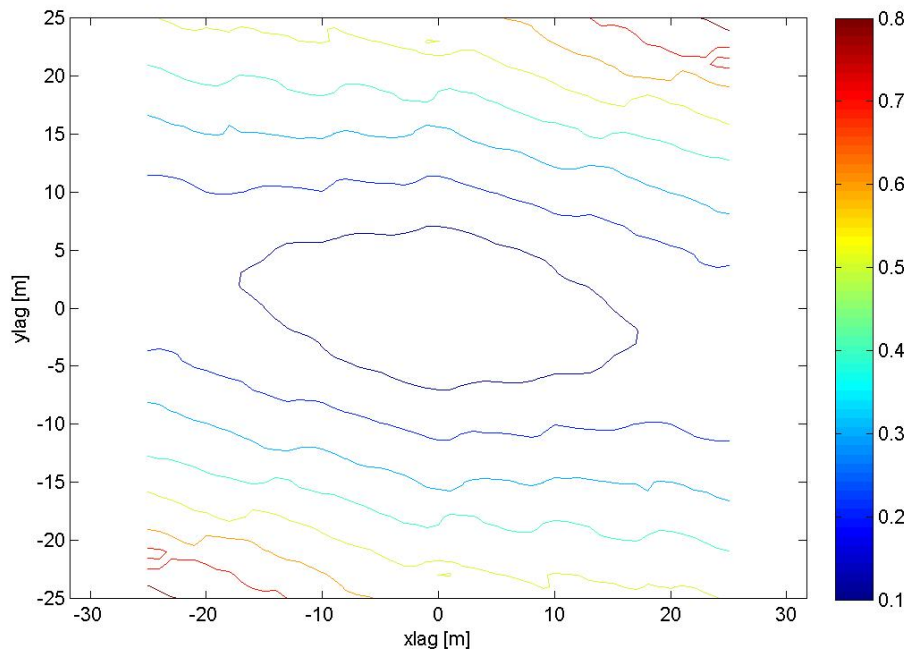


Figure 2: 2-D semivariogram for all 2013 data.

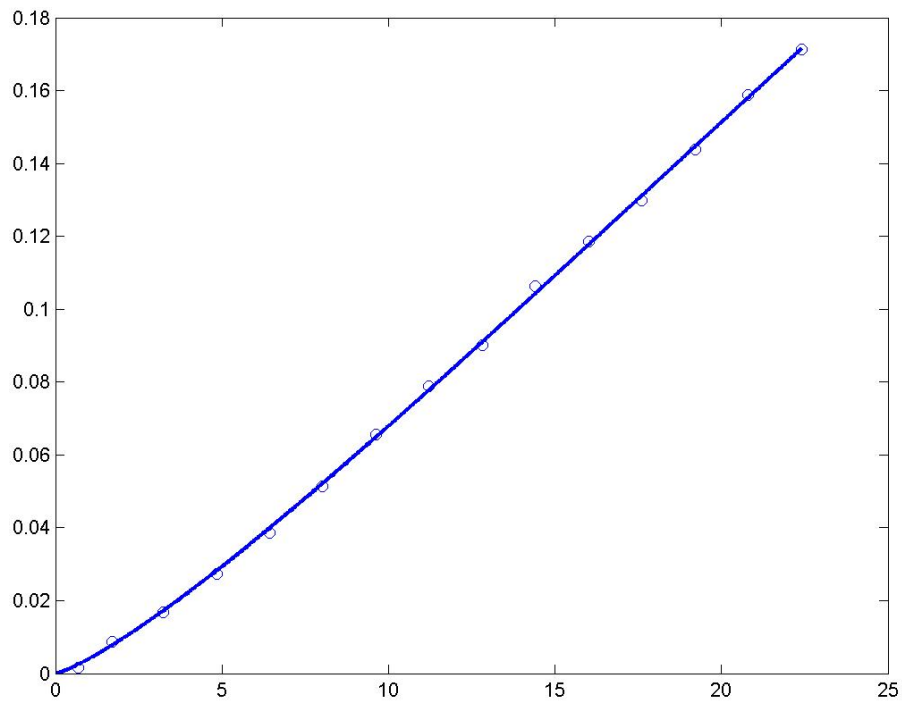


Figure 3: Isotropic semivariogram for all 2007 data and model (modified Gaussian, $p = 1.25$, no nugget).

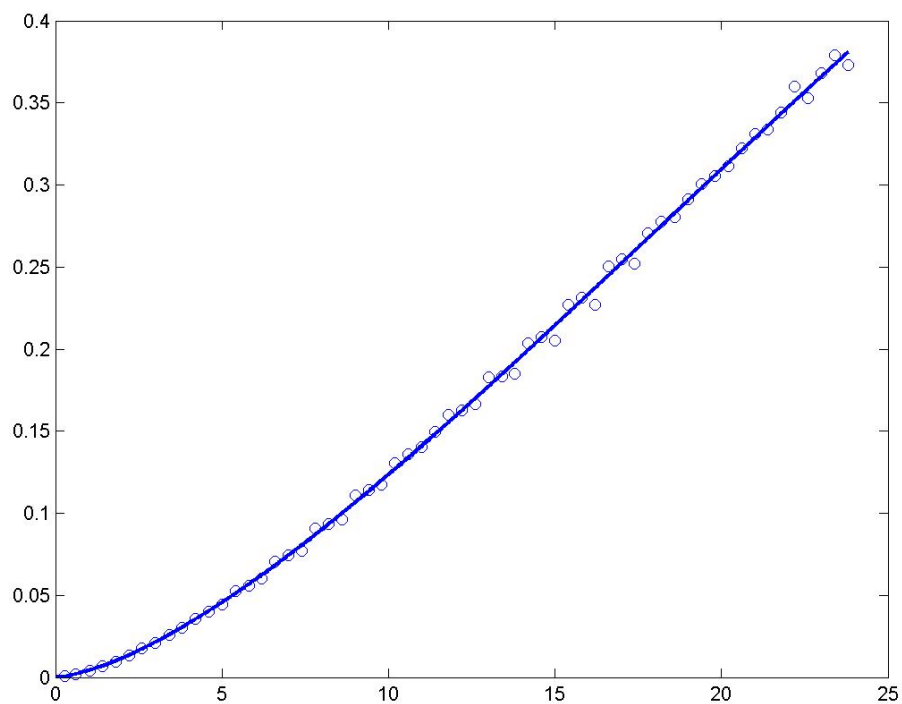


Figure 4: Isotropic semivariogram for all 2013 data and model (modified Gaussian, $p = 1.50$, no nugget).

Appendix

The following 100 pages show results from leave-one-out (LOO) cross-validations for the different interpolation methods, namely NN, TIN, ID, ID2, LM and OK.

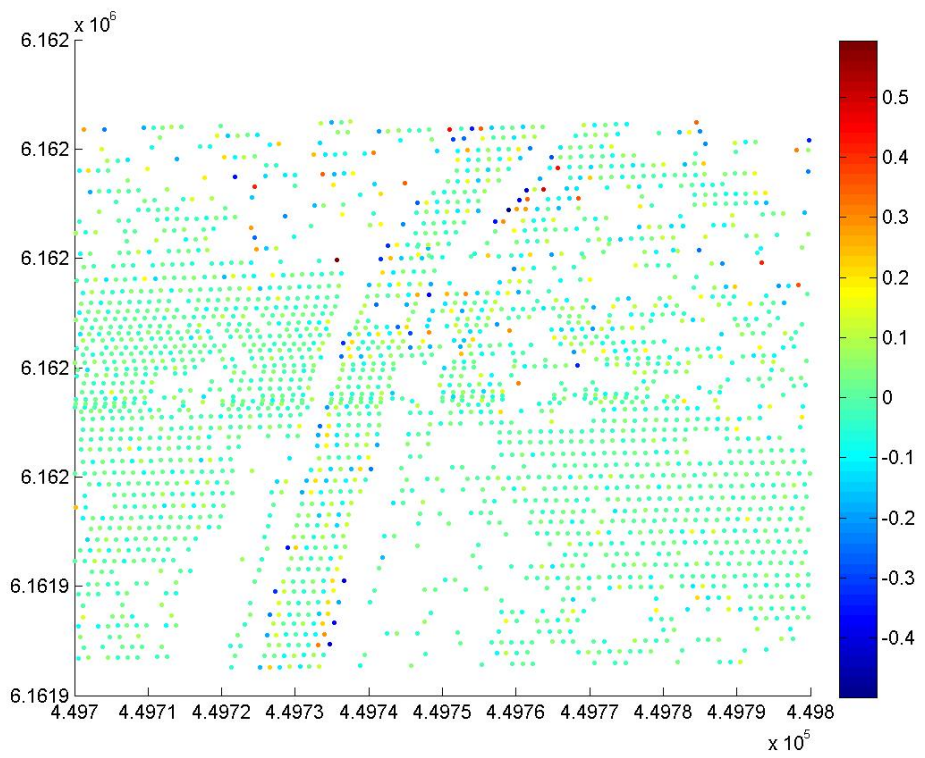
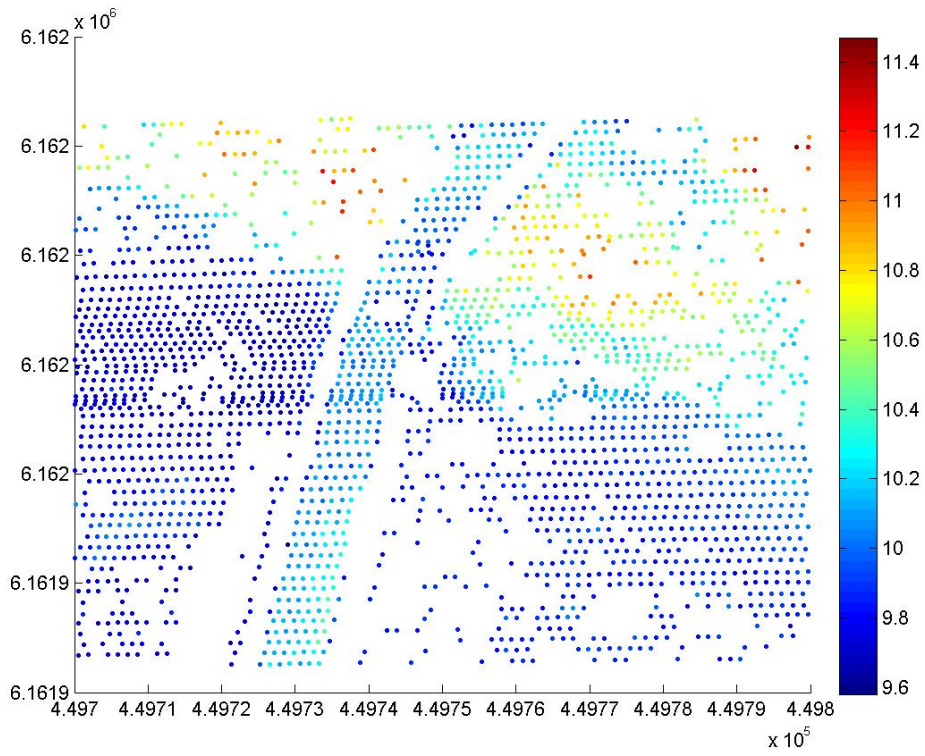
For each of the ten sub-areas in both 2007 and 2013 five pages with results are shown:

1. The first page shows 1) a posting plot of the heights; and 2) a posting plot of the differences between the heights and their LOO estimates from OK based on 30 nearest neighbours.
2. The second page shows 1) contours of a 2-D semivariogram also known as a semivariogram map to check for possible anisotropy; and 2) a (1-D) isotropic also known as an omni-directional semivariogram including a semivariogram model.
3. The third page shows histograms of the differences between the heights and their LOO estimates (based on 30 nearest neighbours for ID, ID2, LM and OK).
4. The fourth page shows bias and RMSE for LM, ID, ID2 and OK as functions of number of nearest neighbours.
5. Finally, the fifth page shows a plot of the heights (x-axis) and their LOO estimates (y-axis) from all six methods (based on 30 nearest neighbours for ID, ID2, LM and OK).

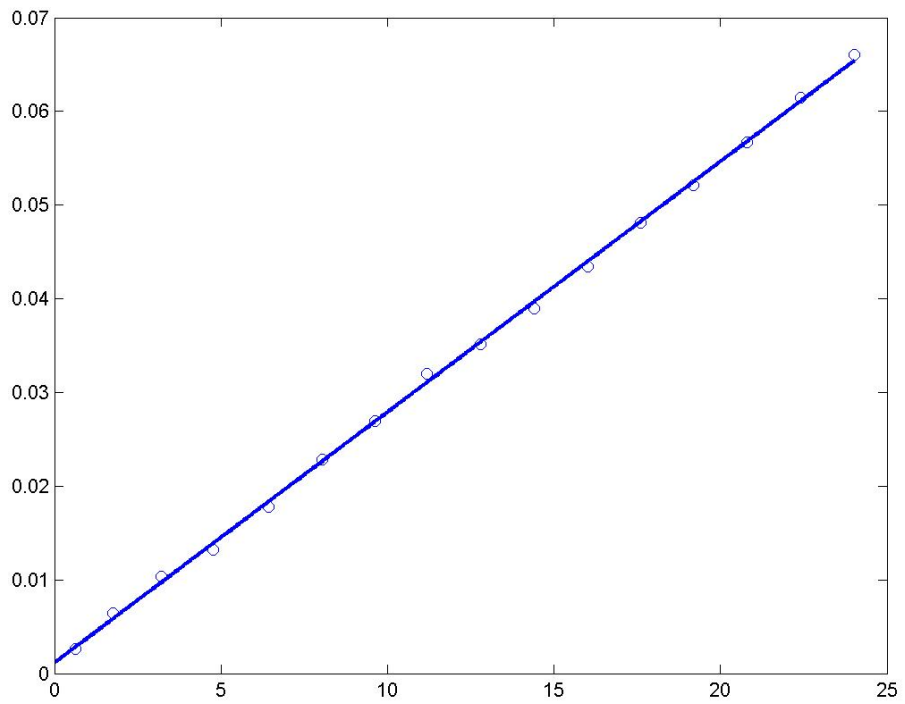
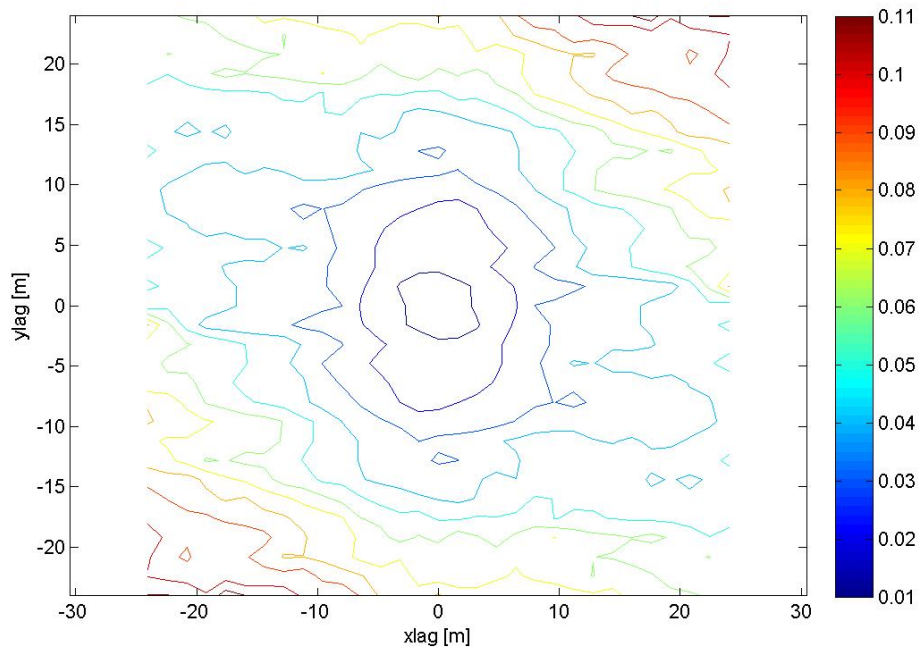
When six figures are presented on one page, the order of the figures is

NN	TIN
ID	ID2
LM	OK

OKSa_2007

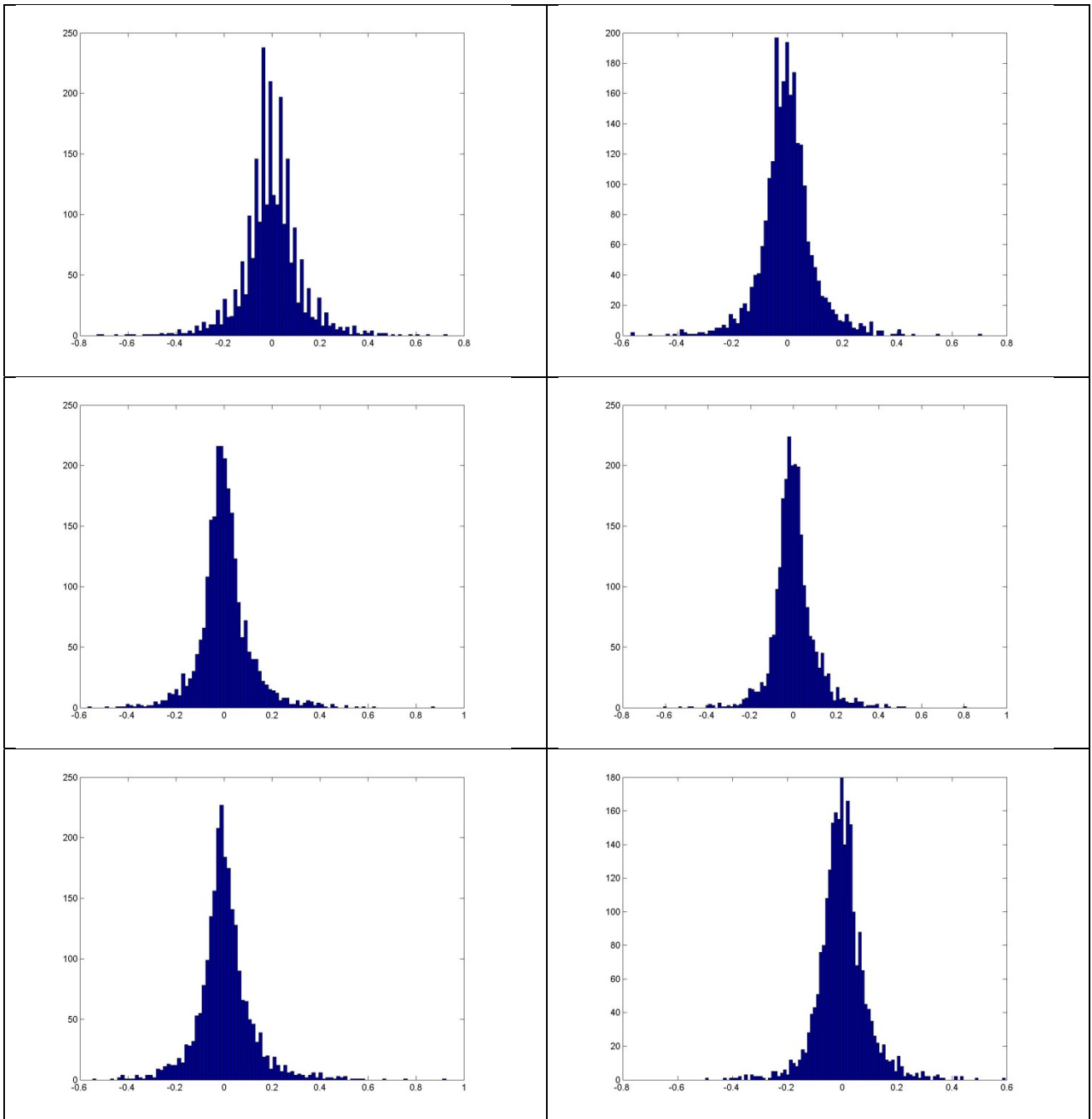


OKSa_2007

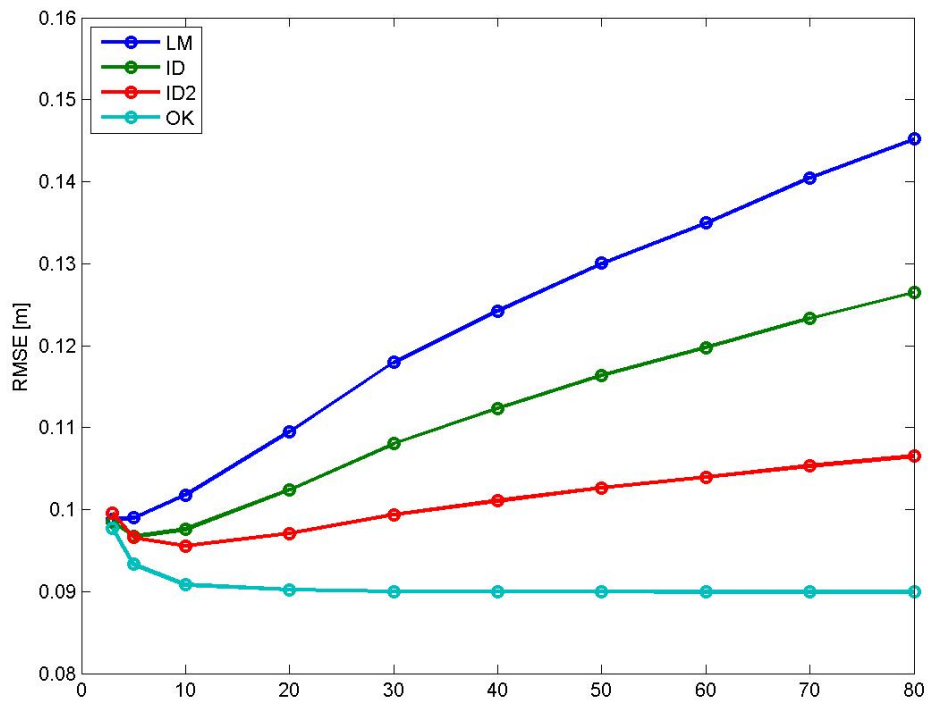
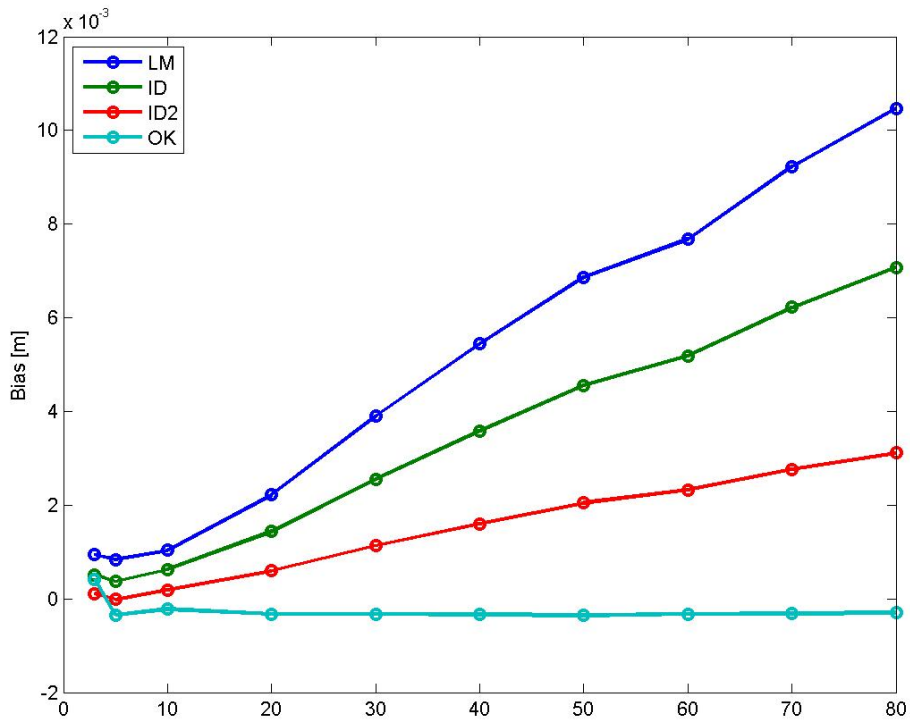


A spherical (linear) semivariogram model is used.

OKSa_2007

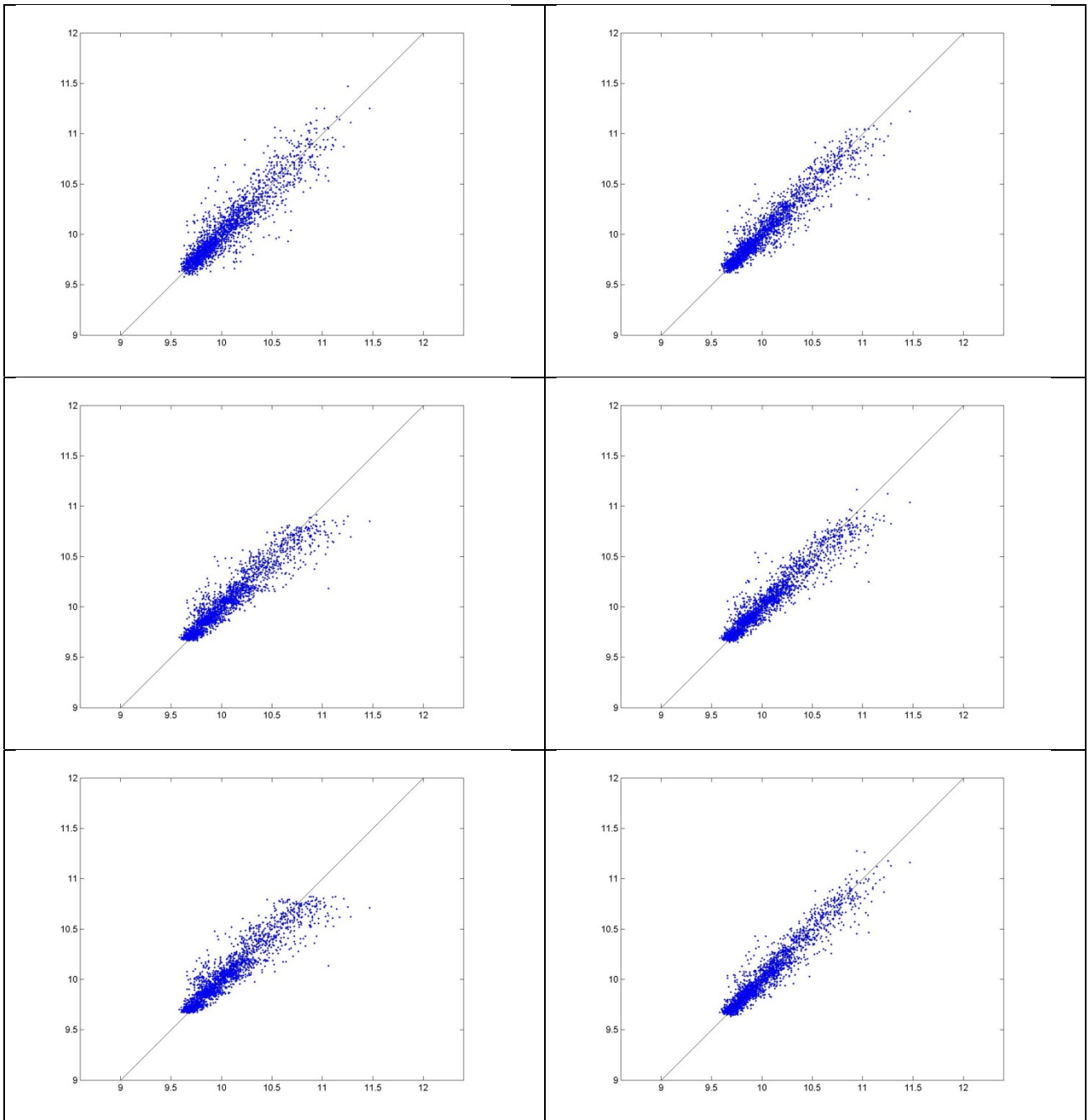


OKSa_2007

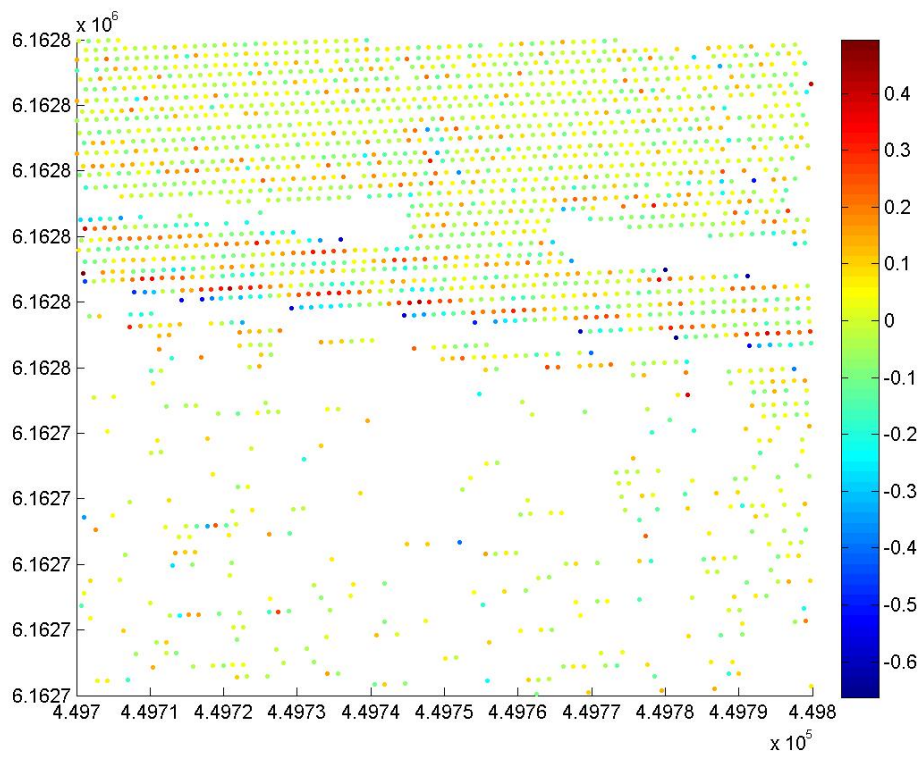
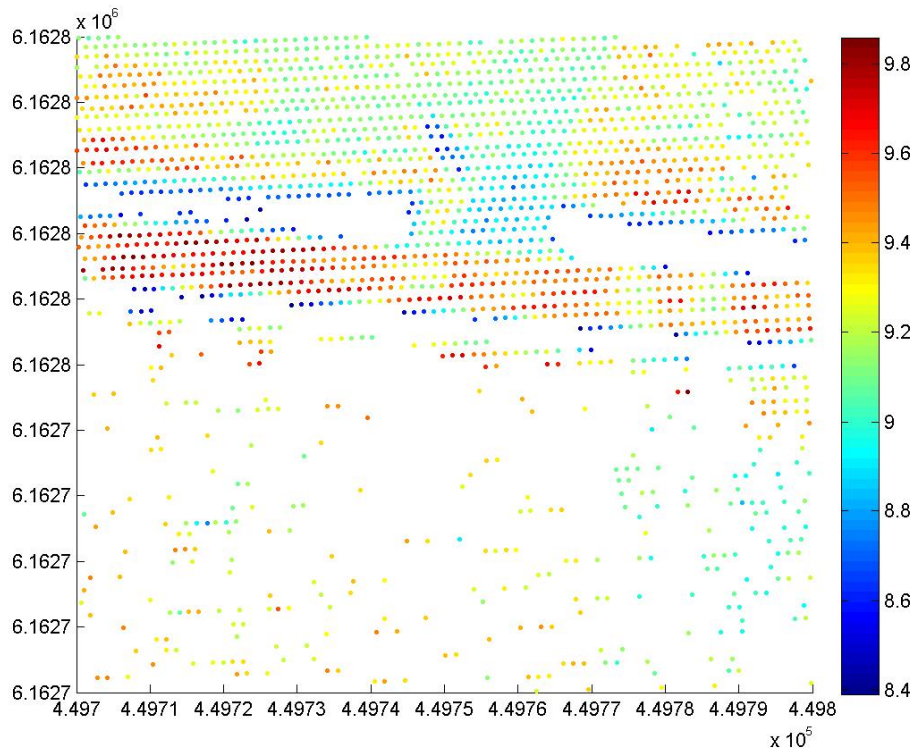


Bias for NN is -0.0029896 m, and for TIN 0.0003281 m. RMSE for NN is 0.12486 m, and for TIN 0.09863 m.

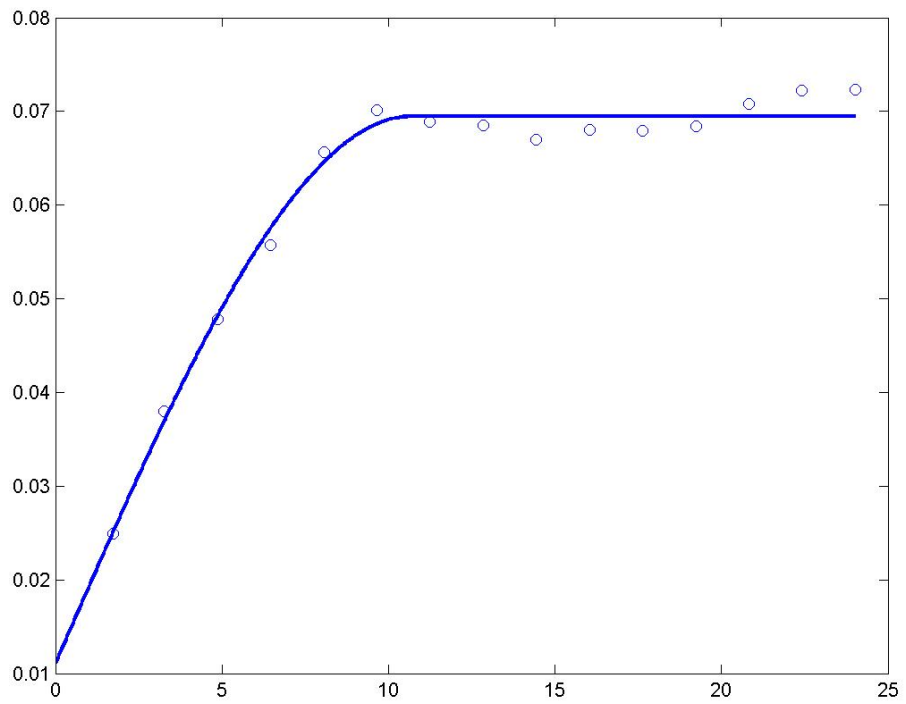
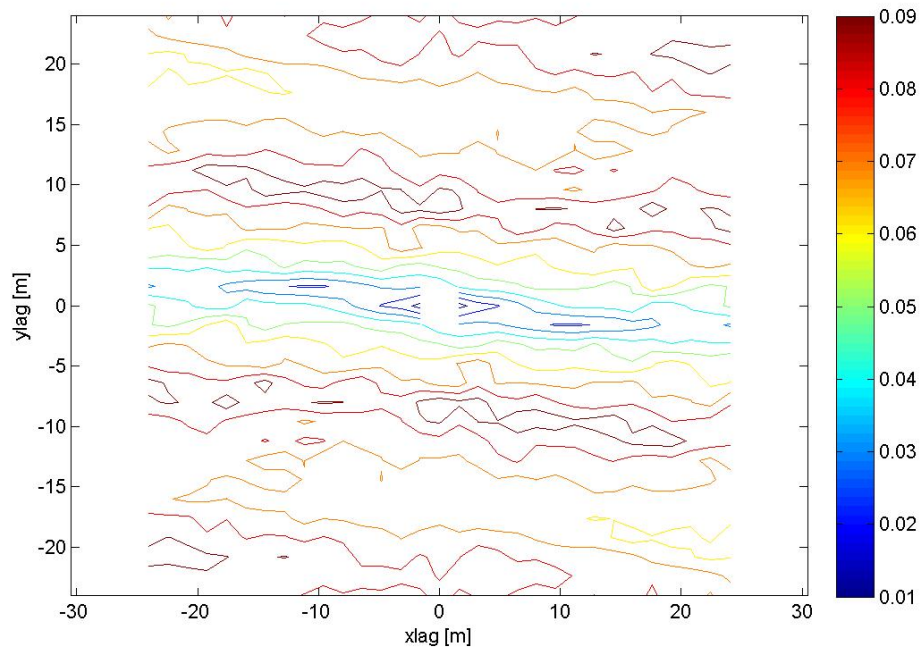
OKSa_2007



OKSb_2007

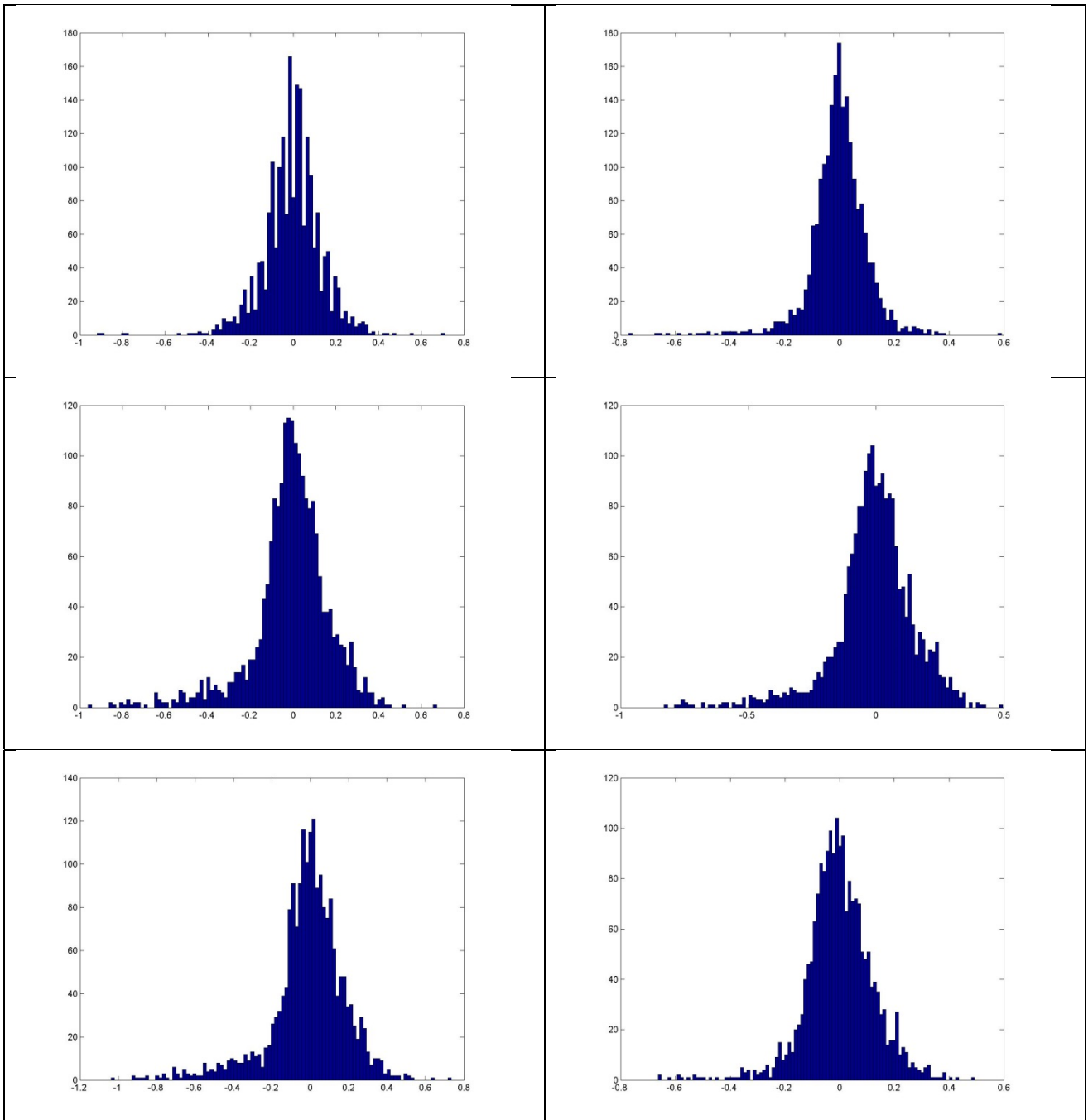


OKSb_2007

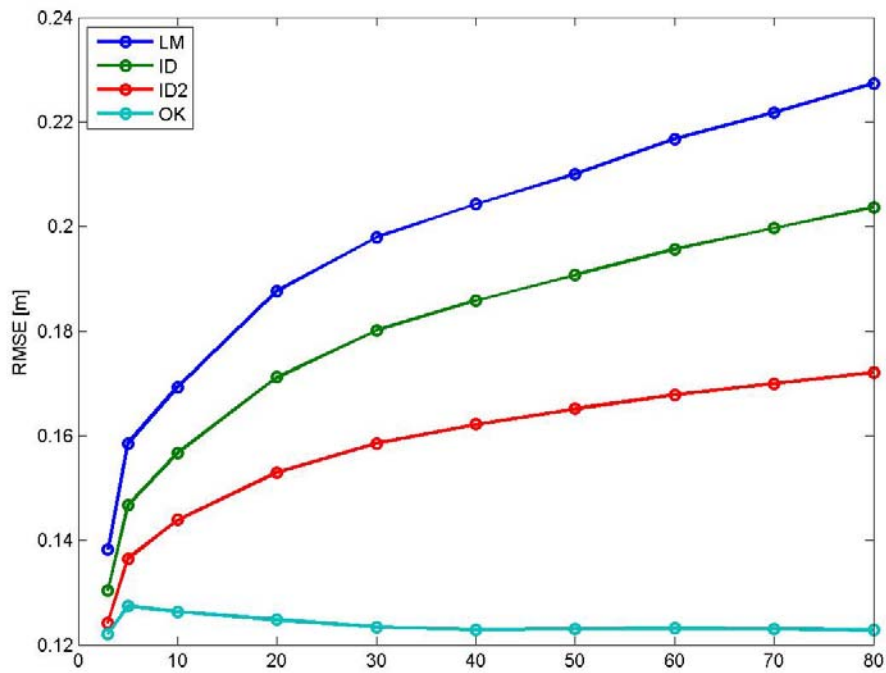
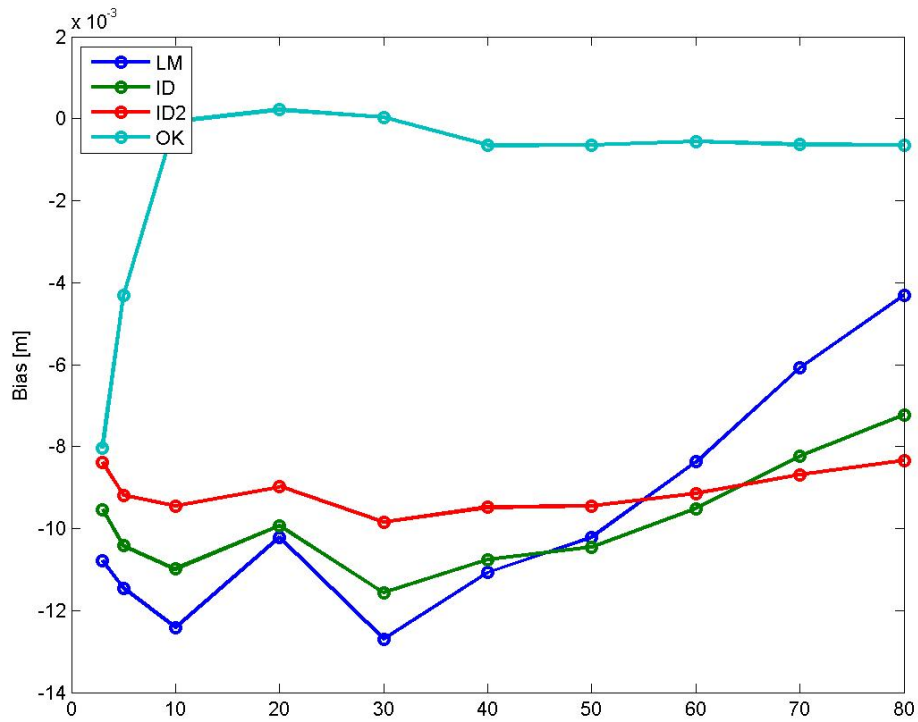


A spherical semivariogram model is used.

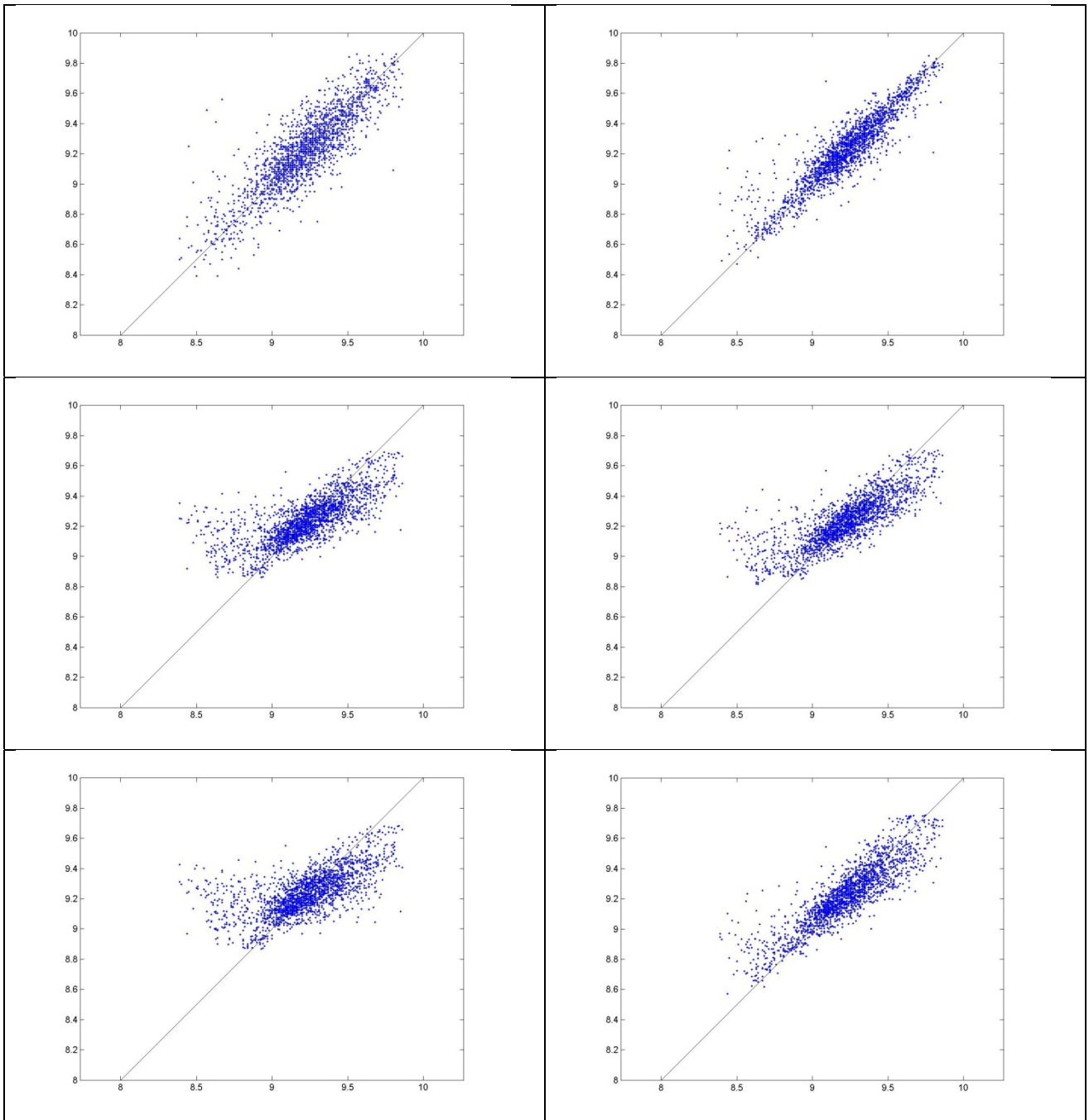
OKSb_2007



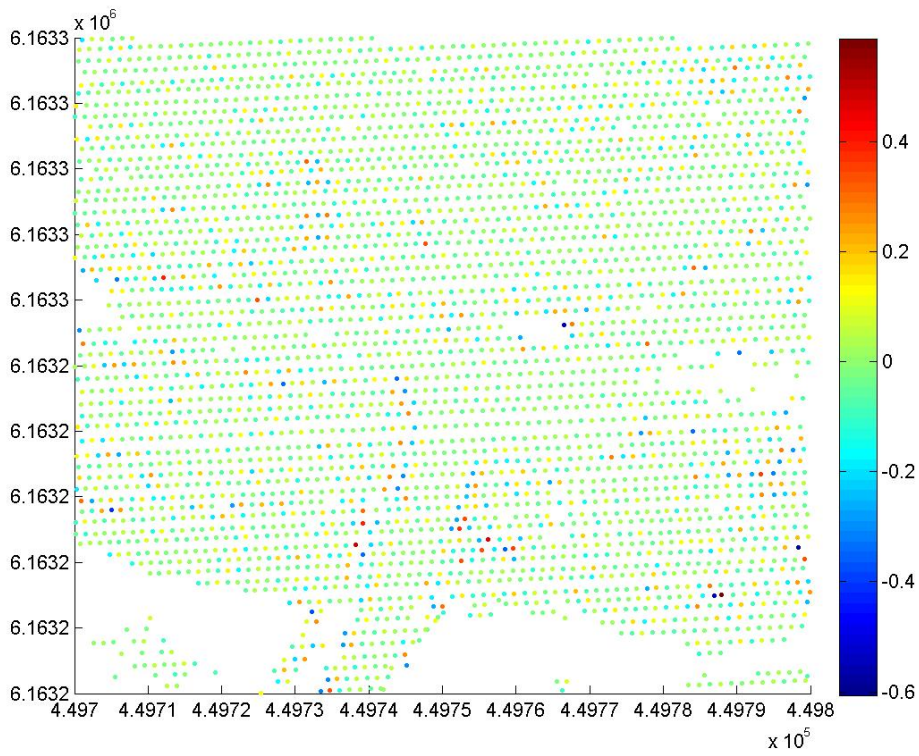
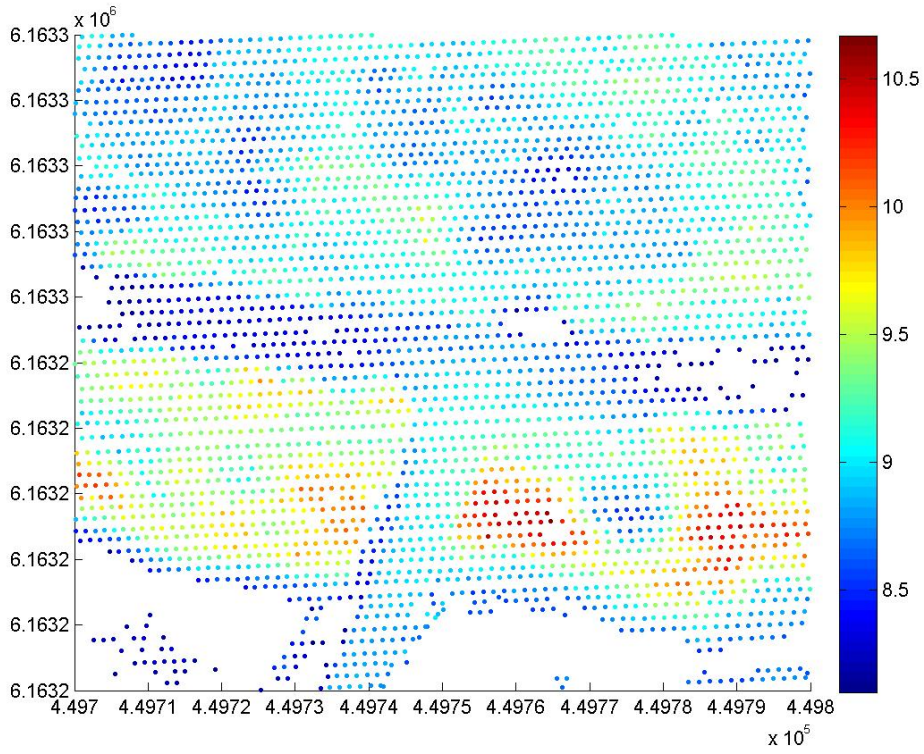
OKSb_2007



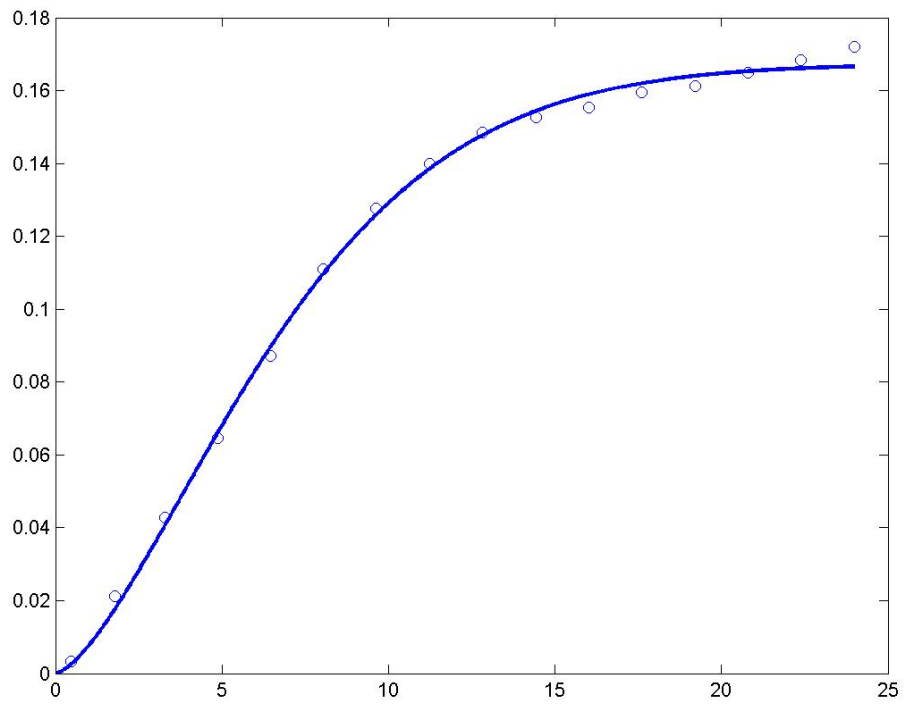
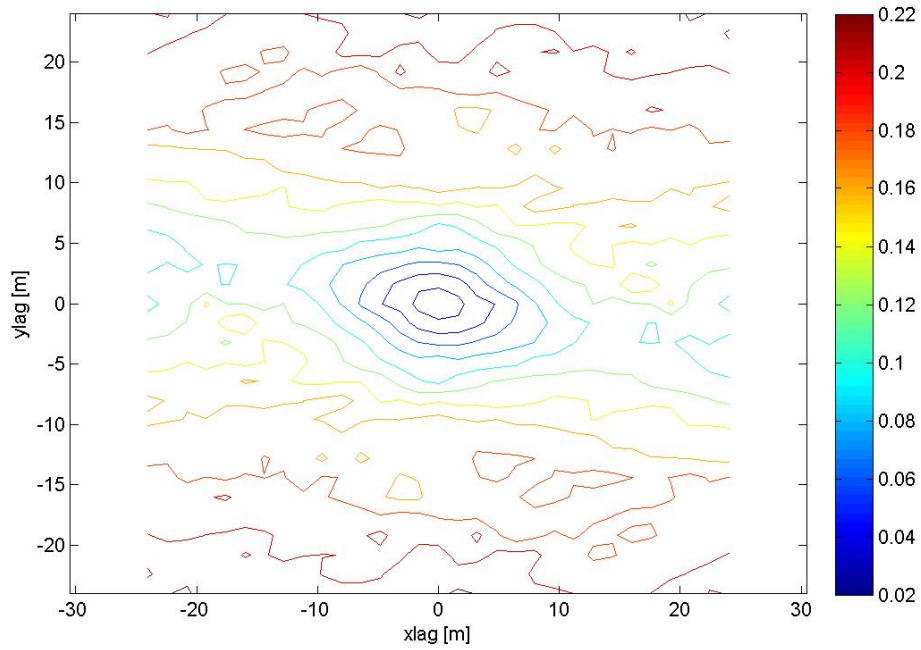
Bias for NN is -0.0031856 m, and for TIN -0.0028933 m. RMSE for NN is 0.13378 m, and for TIN 0.10383 m.



OKSc_2007

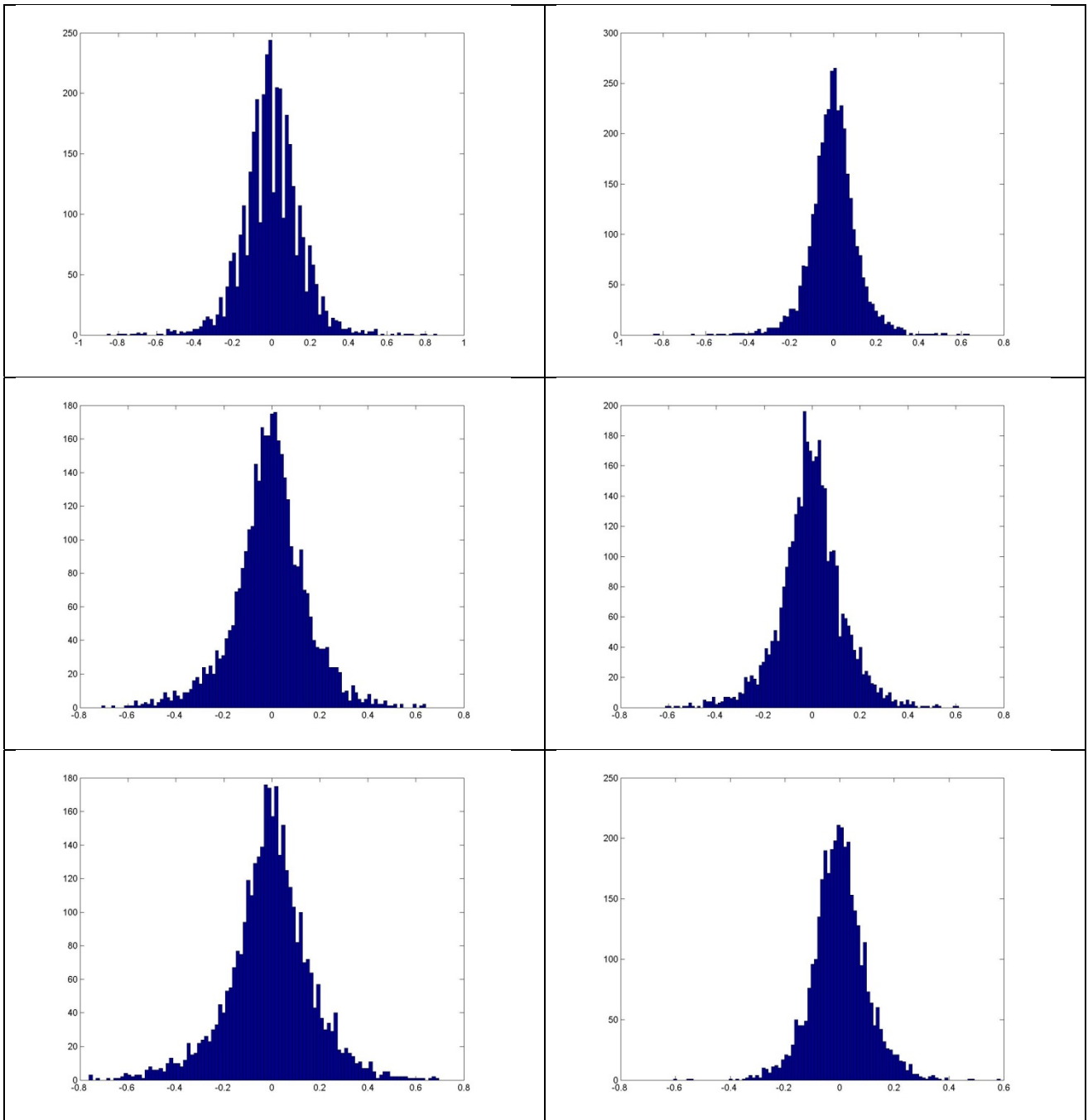


OKSc_2007

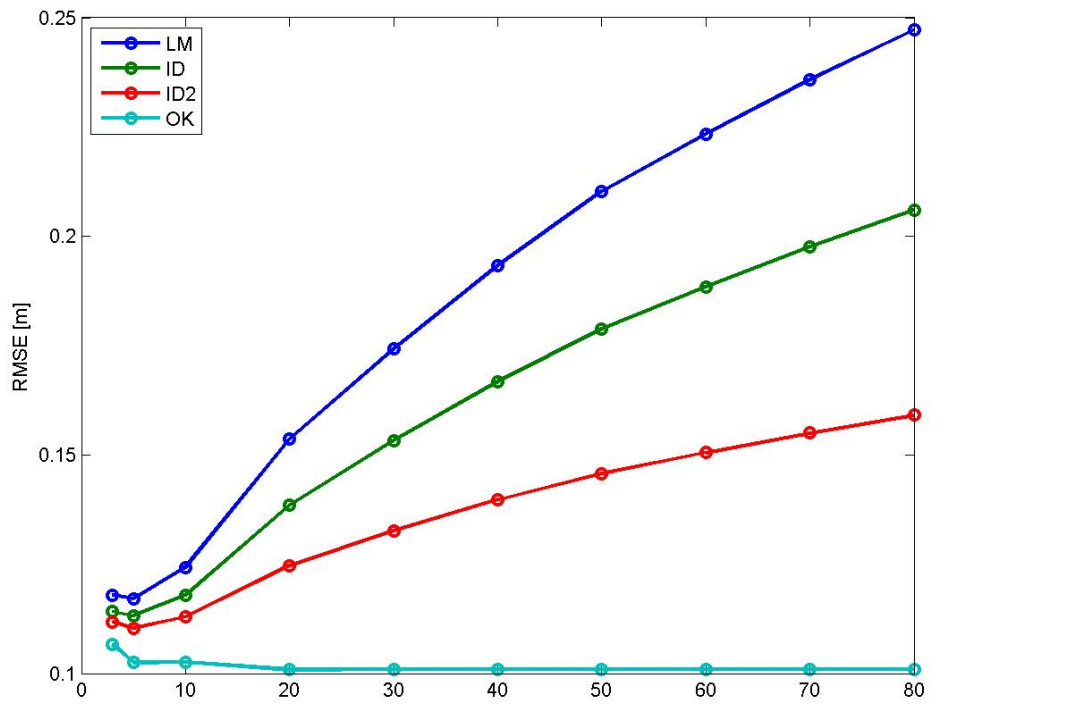
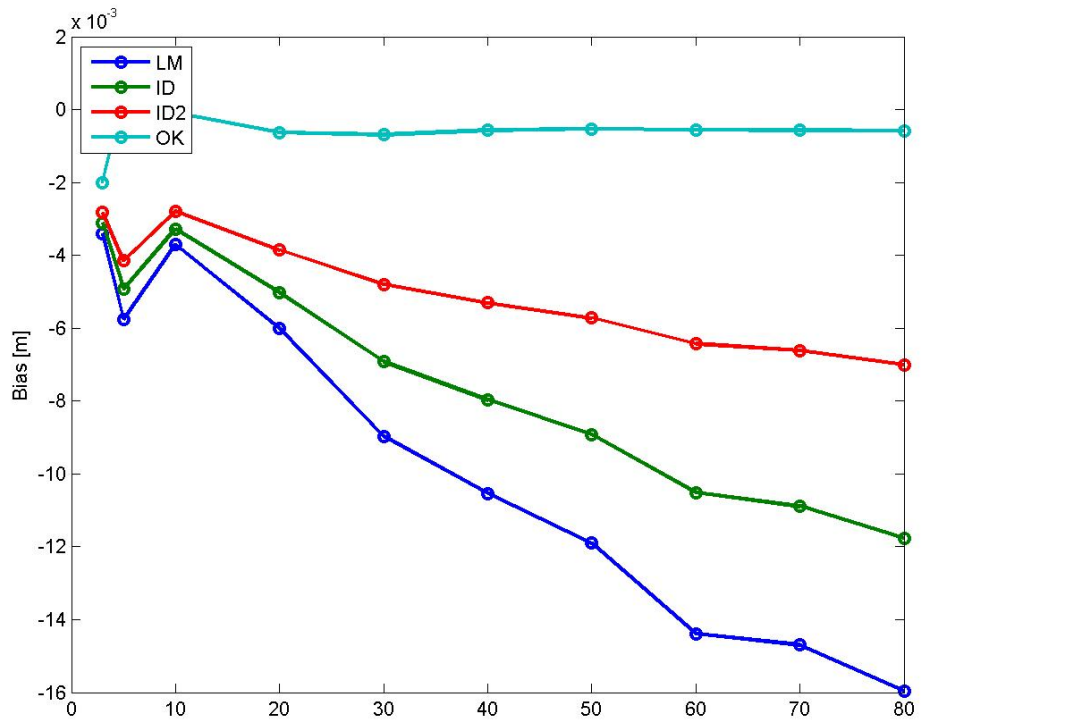


A modified Gaussian semivariogram model is used (power is 1.5 and not 2), no nugget effect.

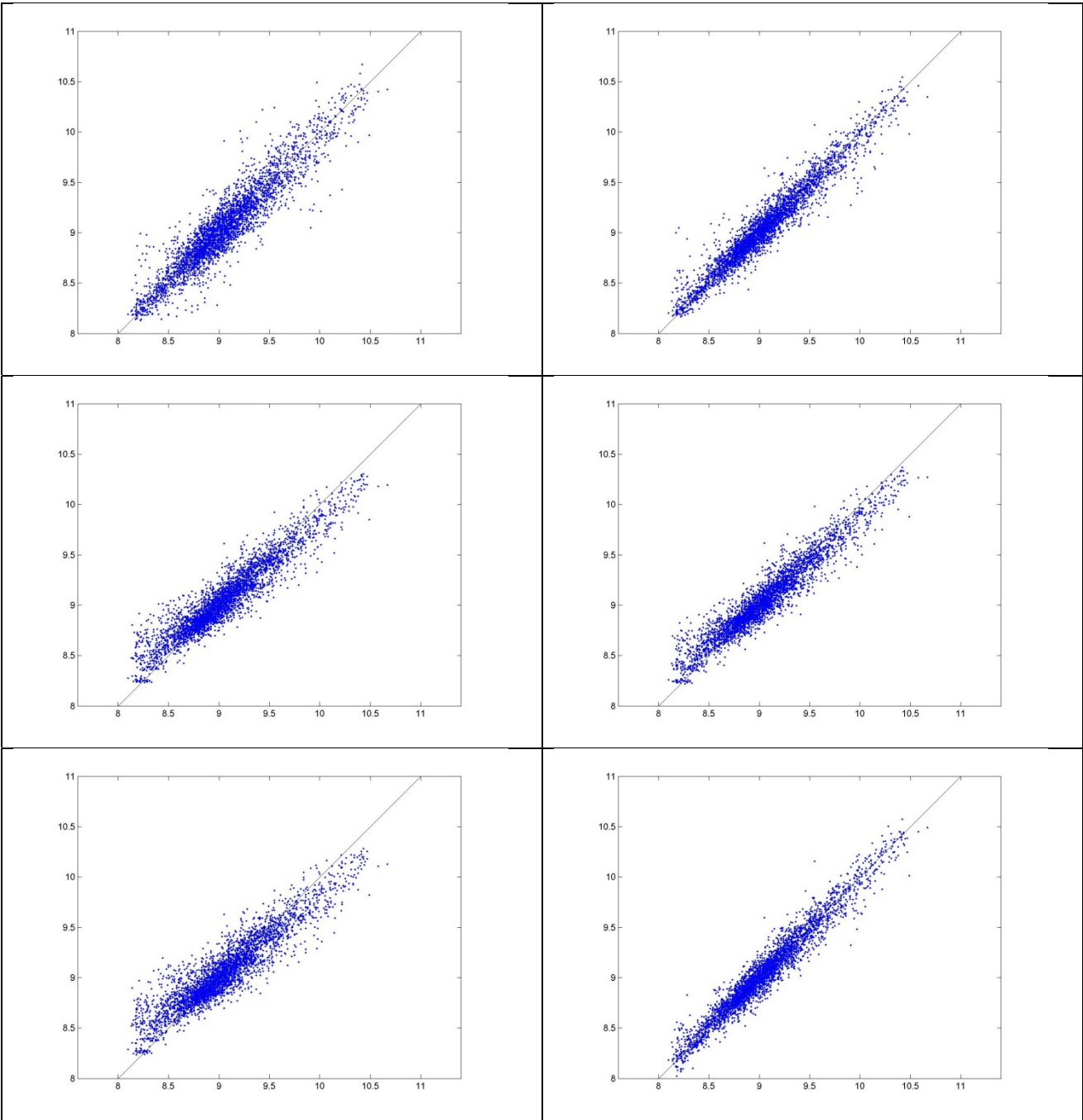
OKSc_2007



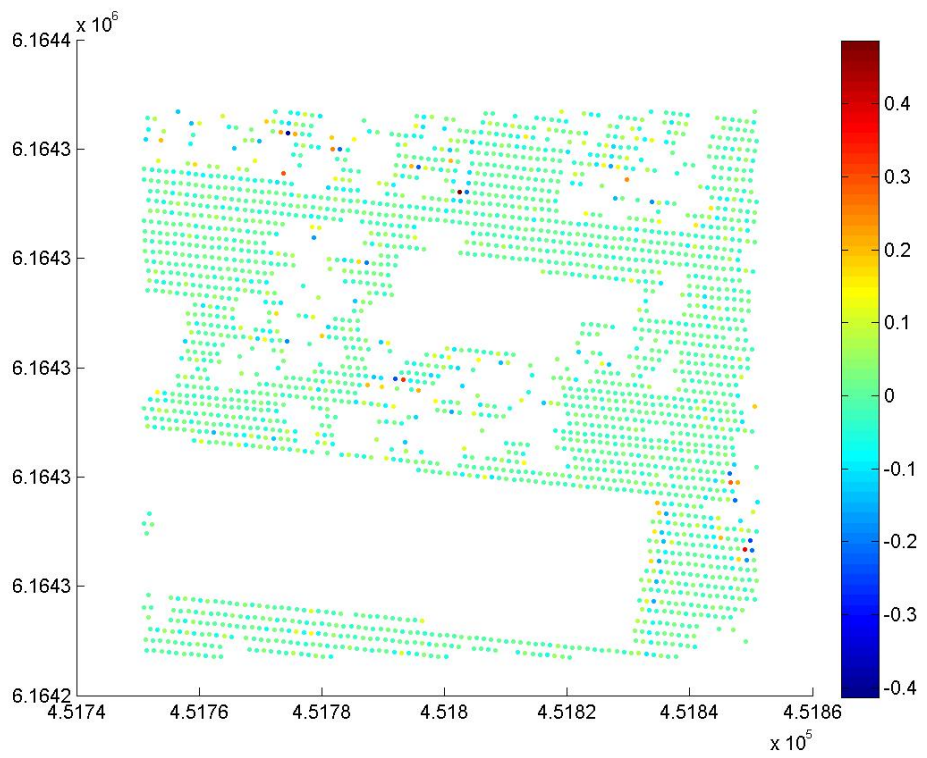
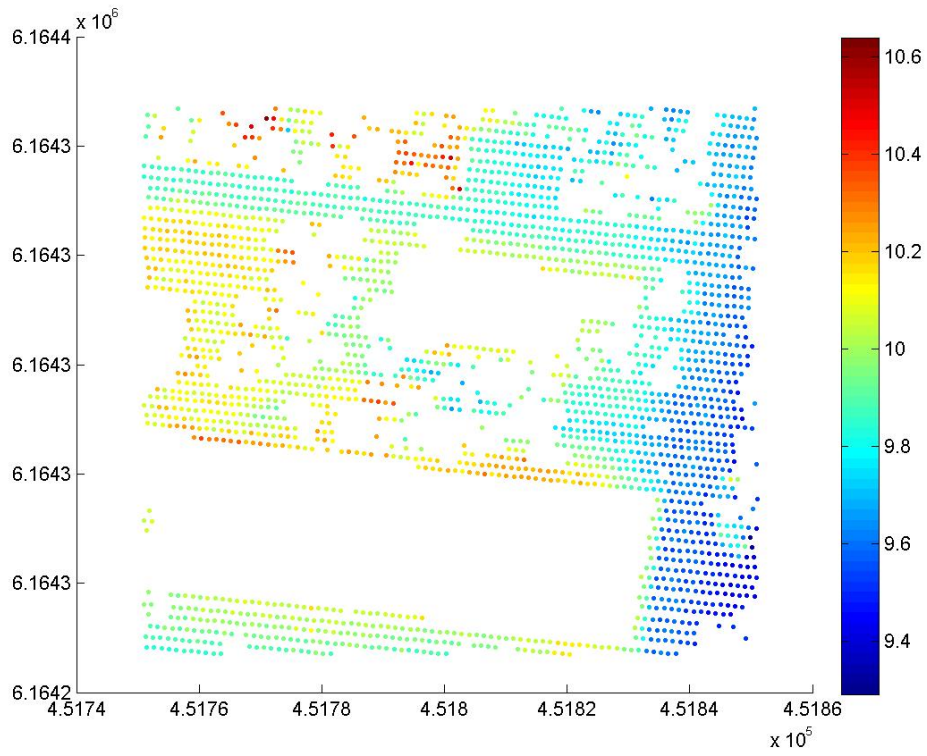
OKSc_2007



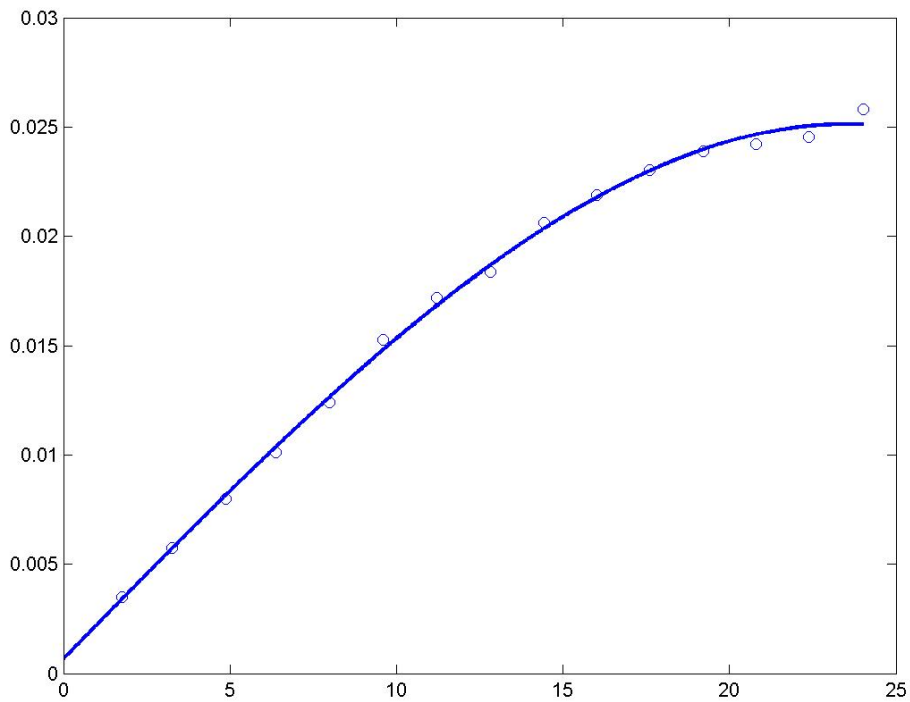
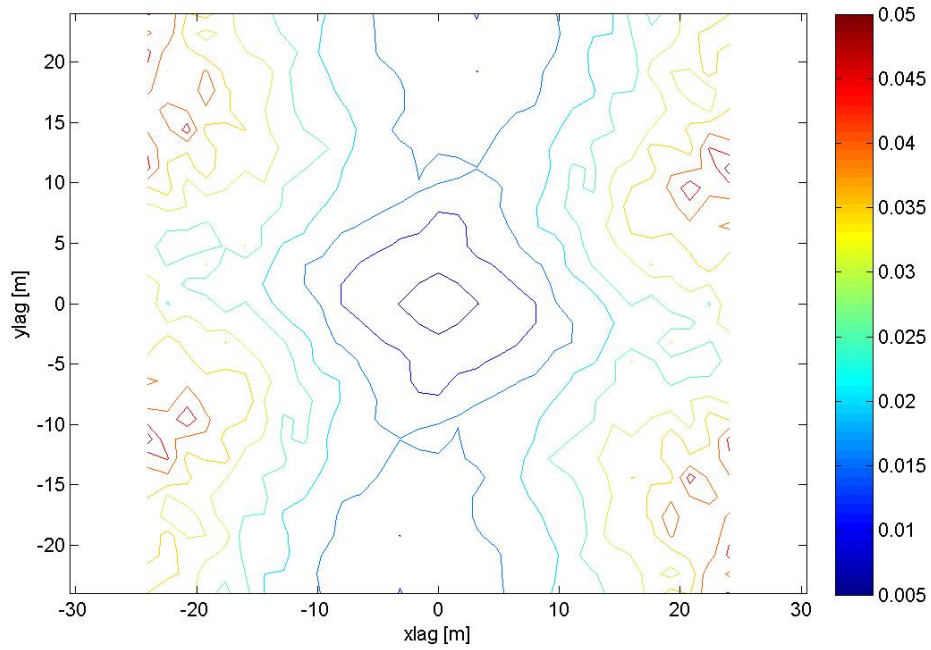
Bias for NN is -0.00179 m, and for TIN -0.00088283 m. RMSE for NN is 0.15491 m, and for TIN 0.11269 m.



OKSd_2007

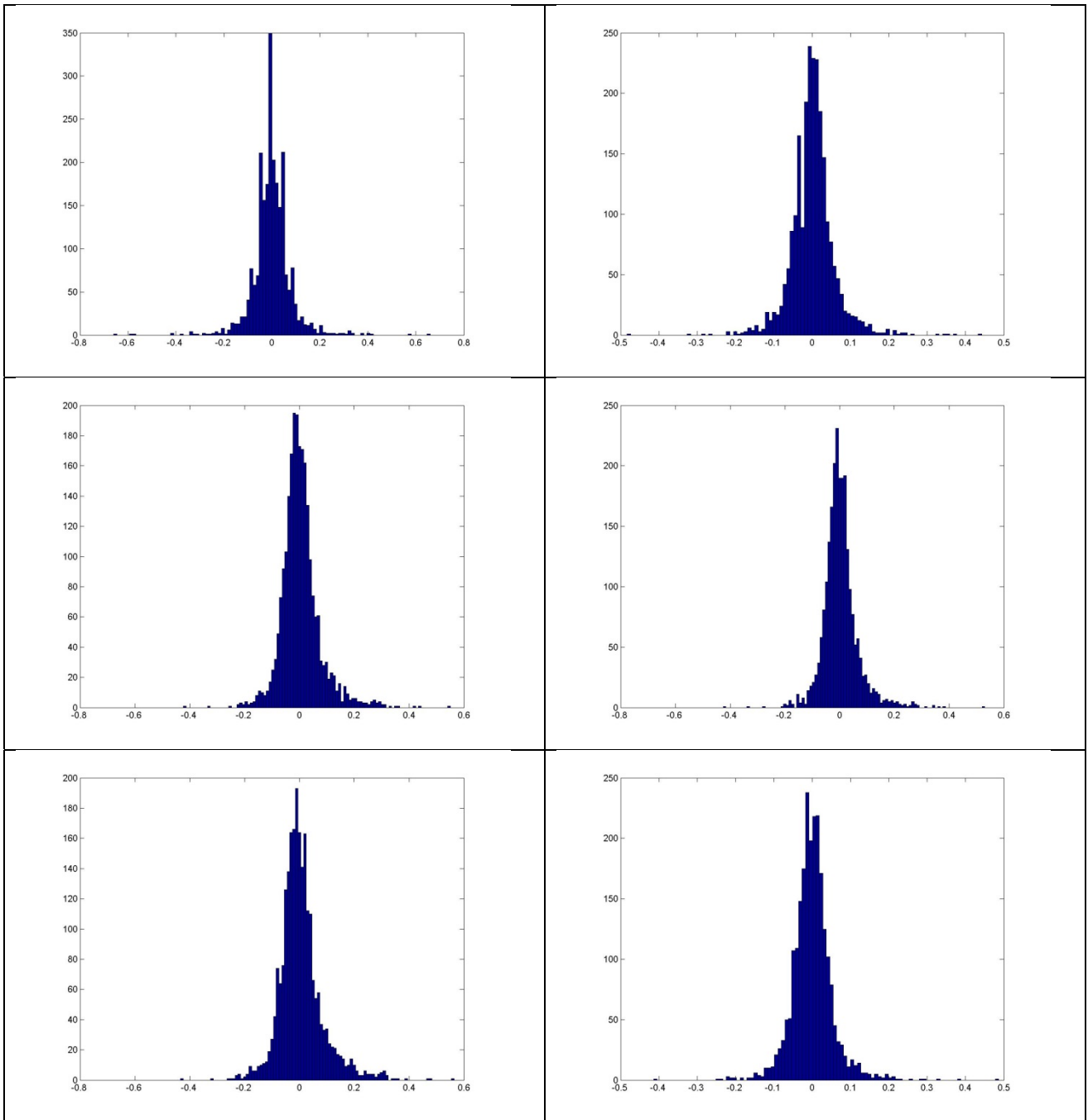


OKSd_2007

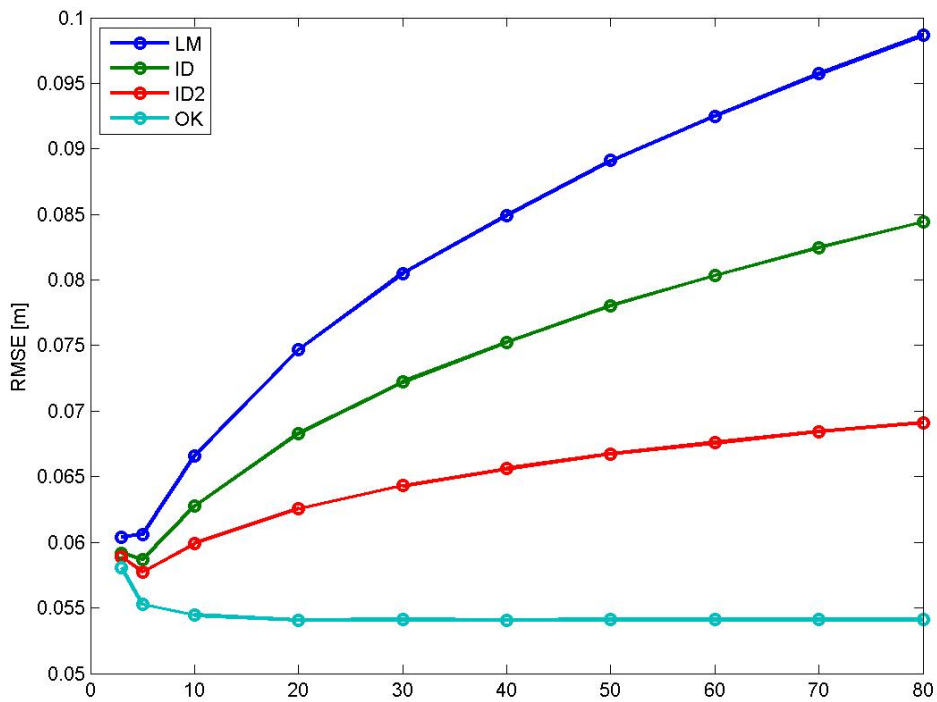
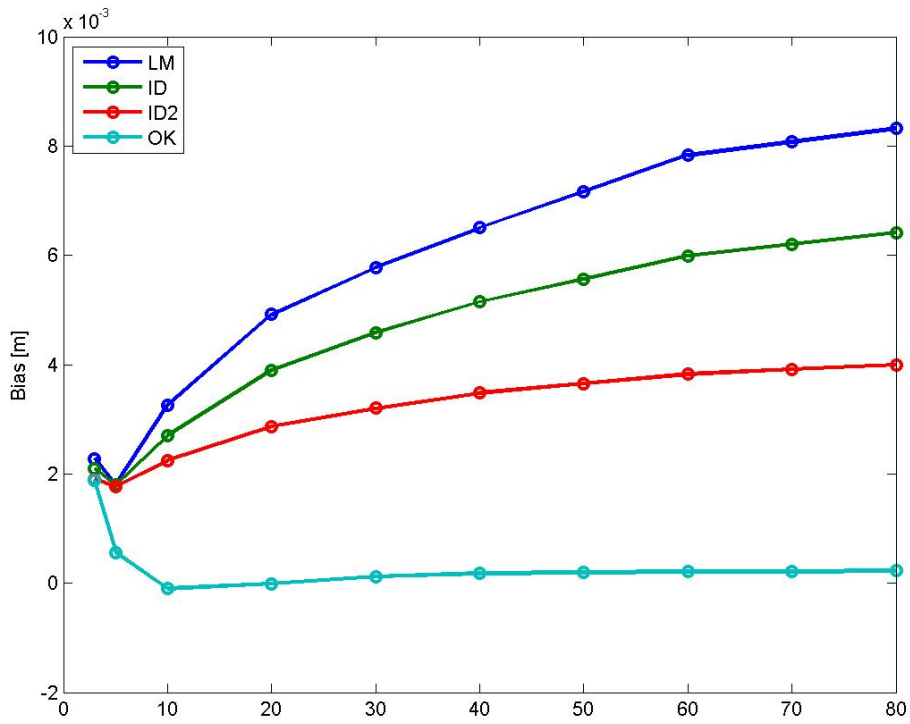


A spherical semivariogram model is used.

OKSd_2007

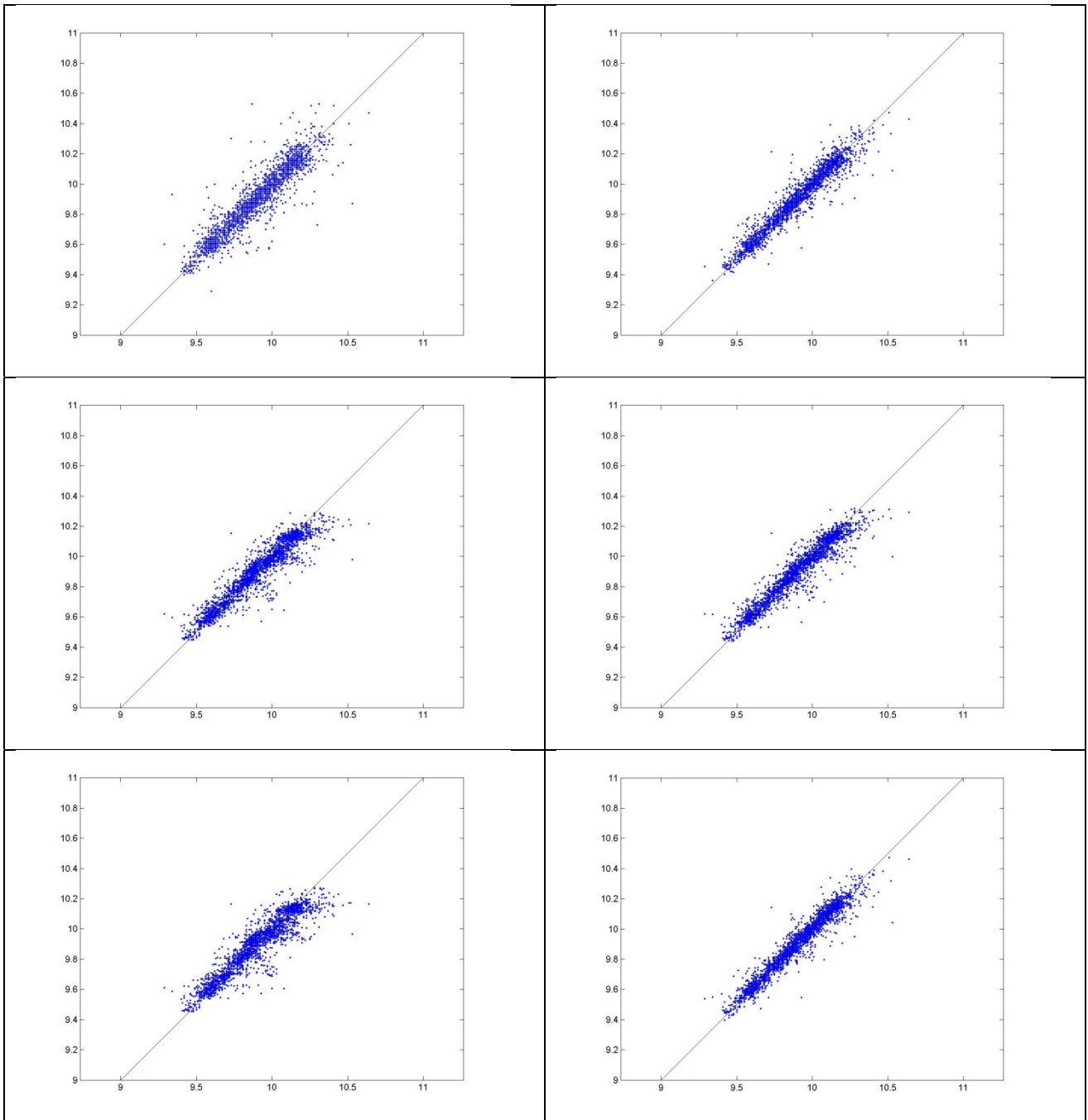


OKSd_2007

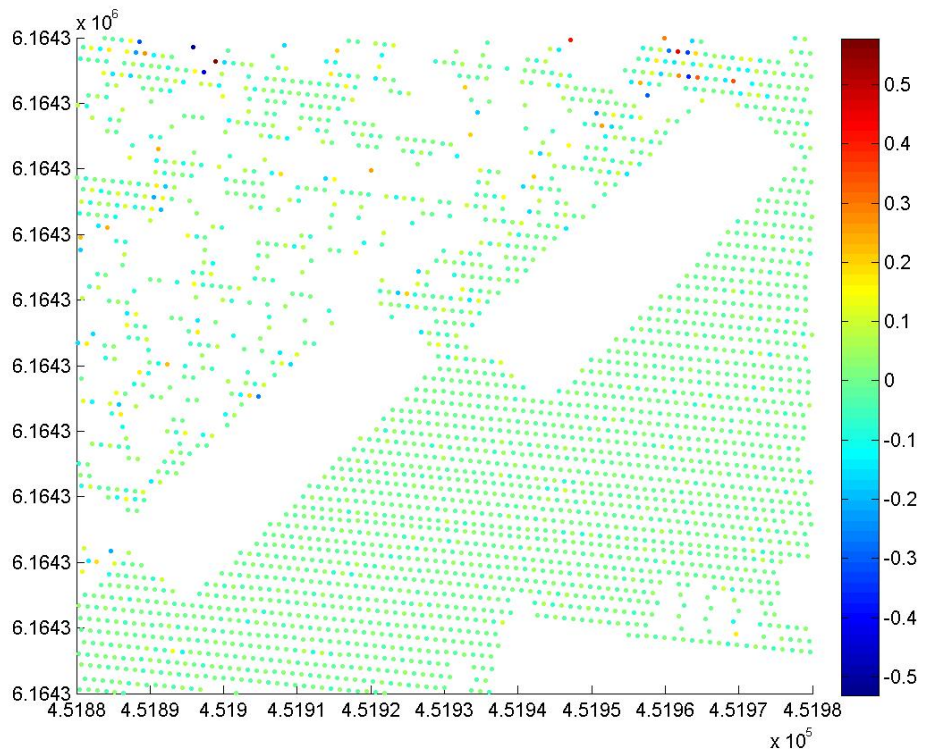
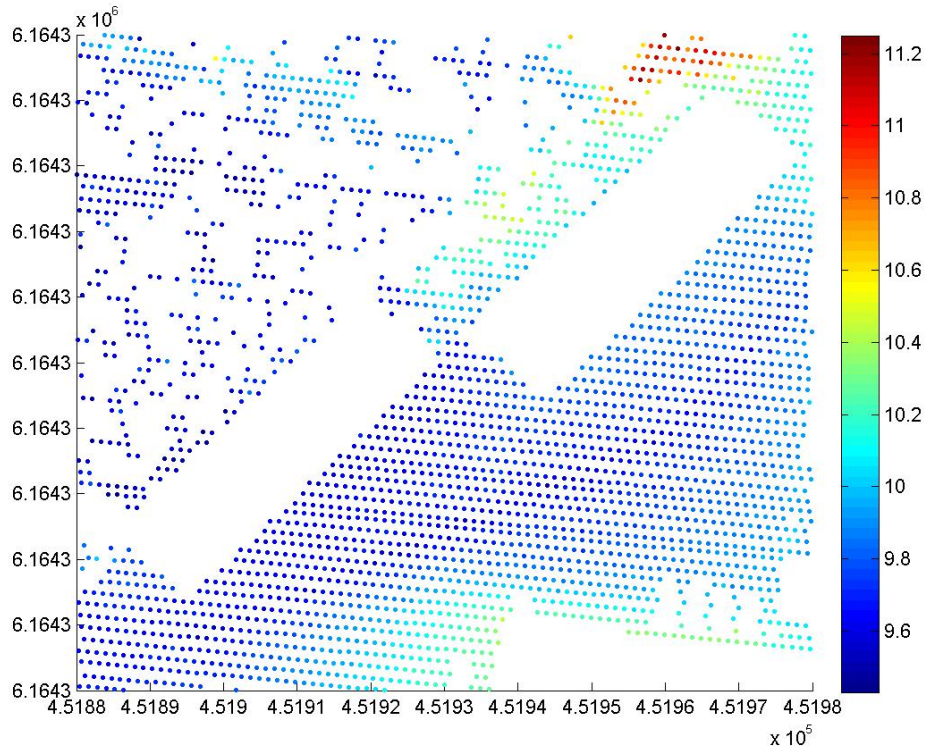


Bias for NN is 0.0002800 m, and for TIN 0.0012231 m. RMSE for NN is 0.079079 m, and for TIN 0.057797 m.

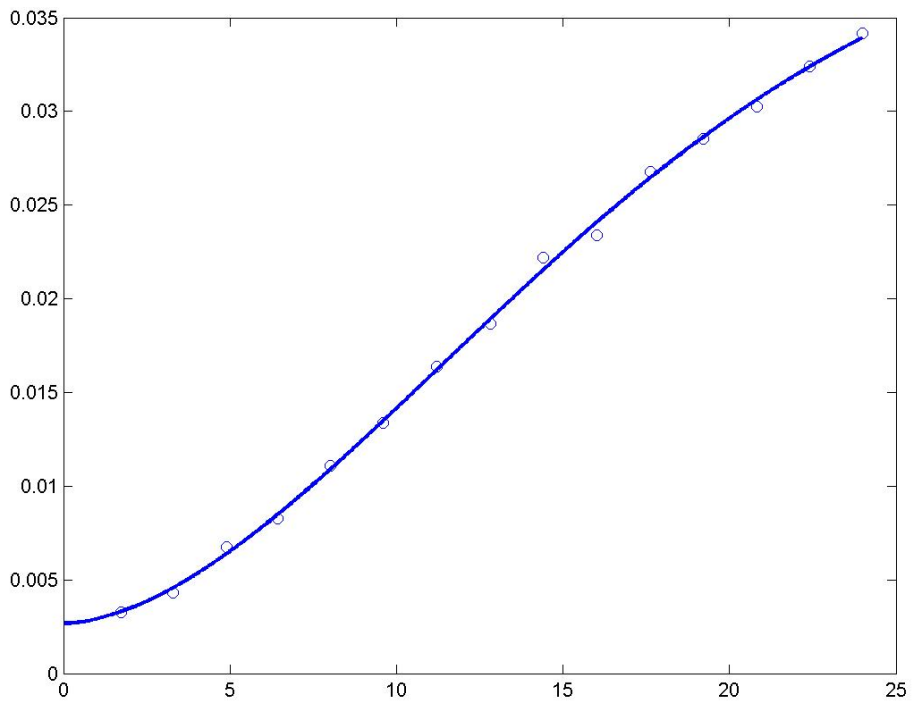
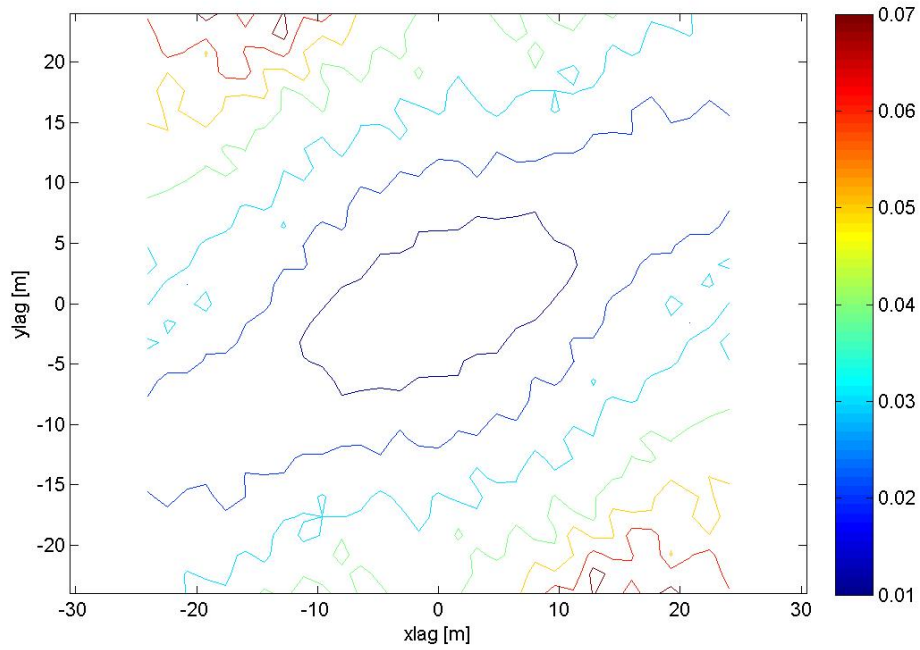
OKSd_2007



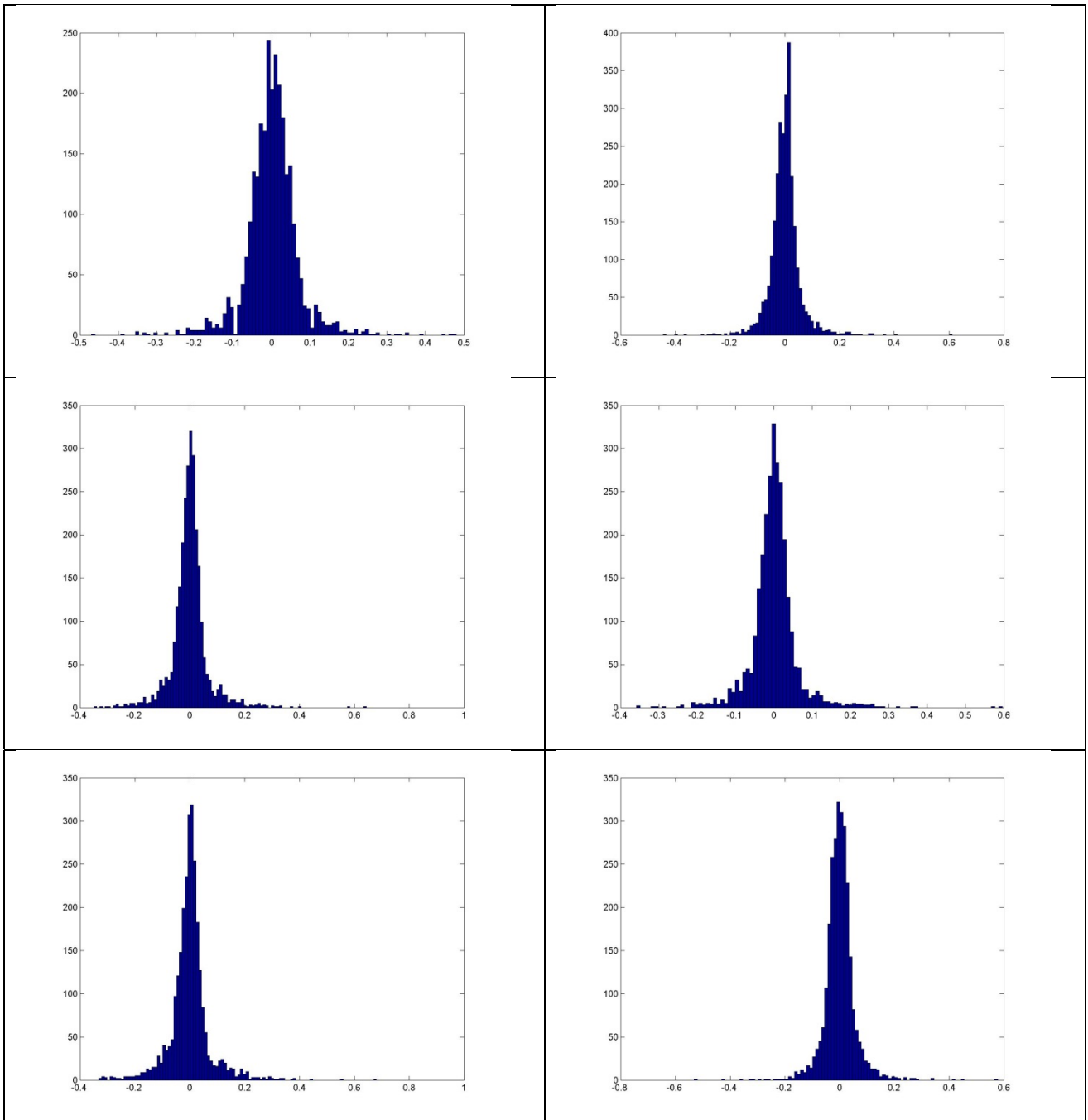
OKSe_2007



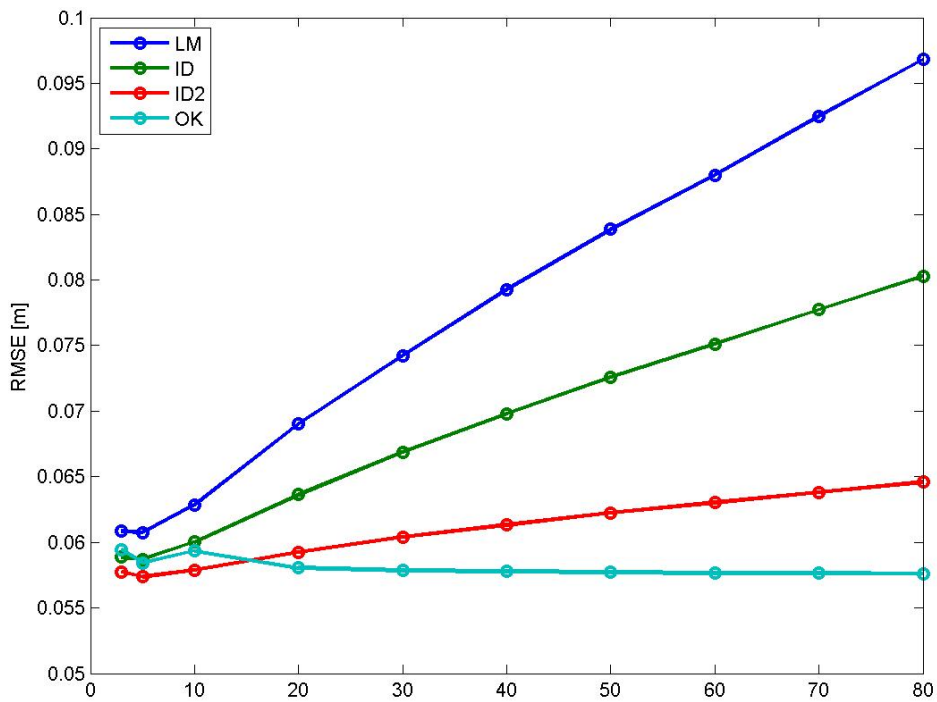
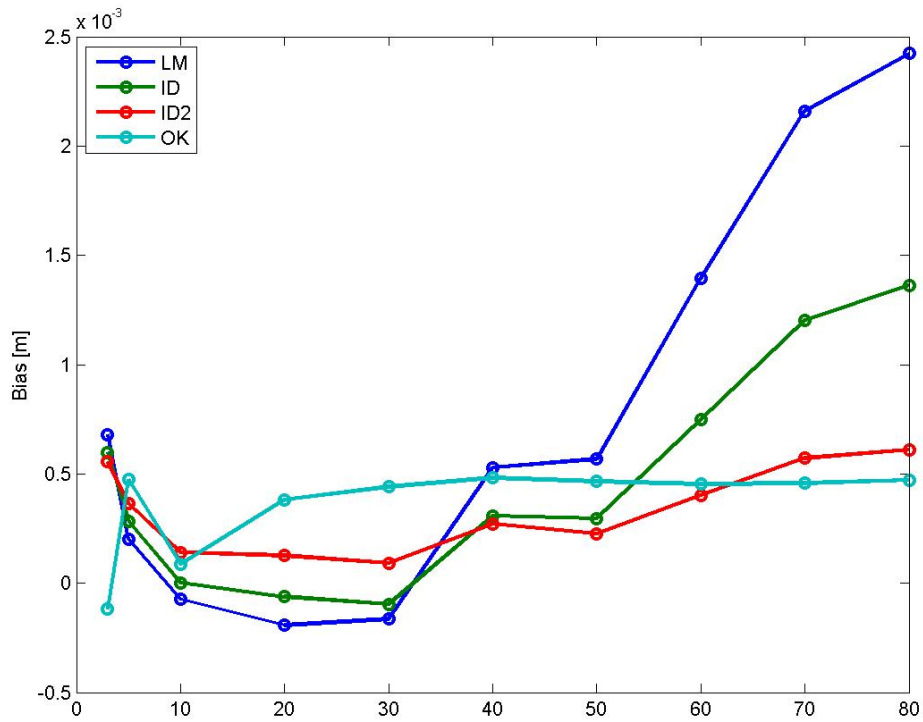
OKSe_2007



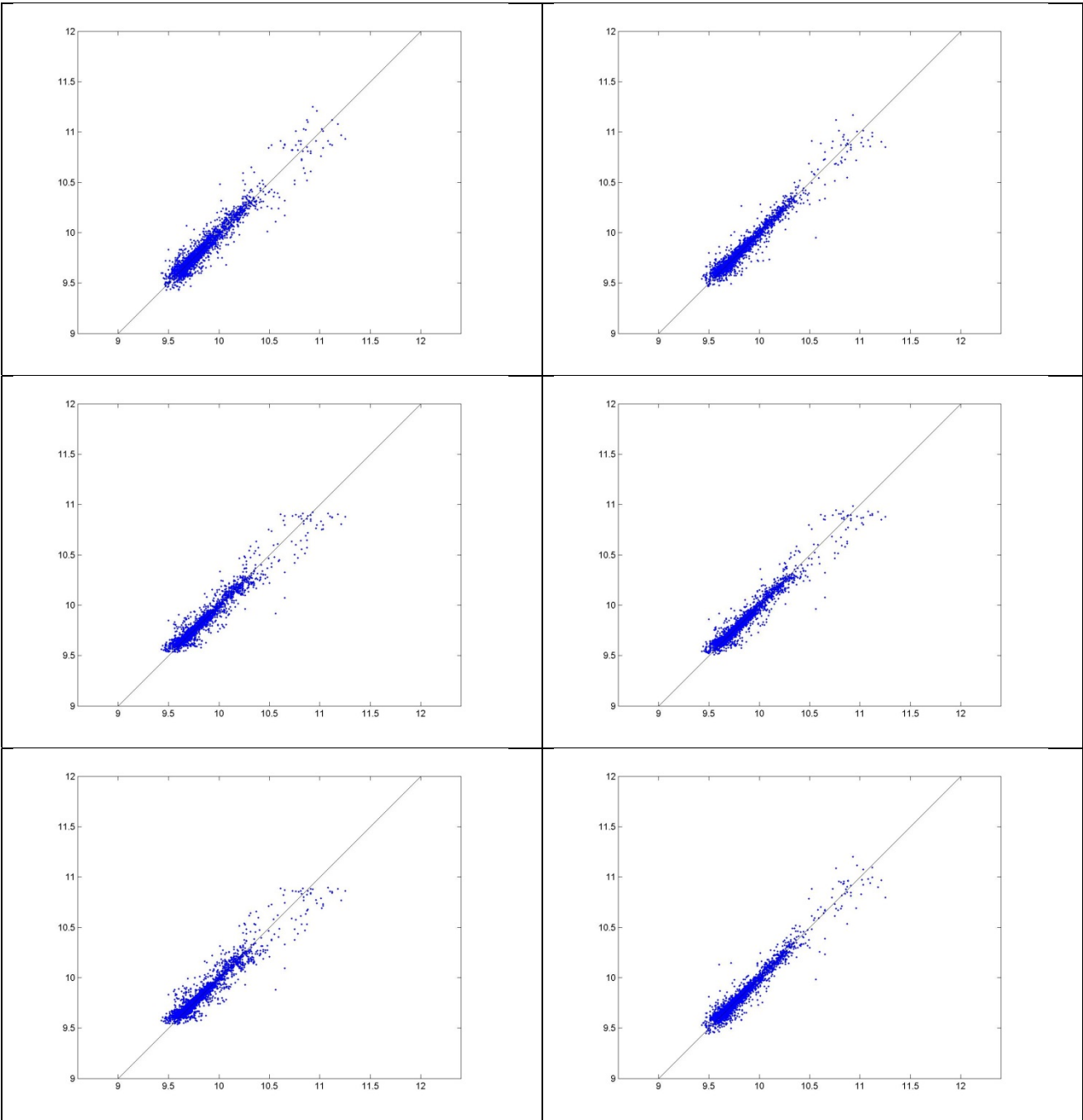
A modified Gaussian semivariogram model is used (power is 1.75 and not 2).



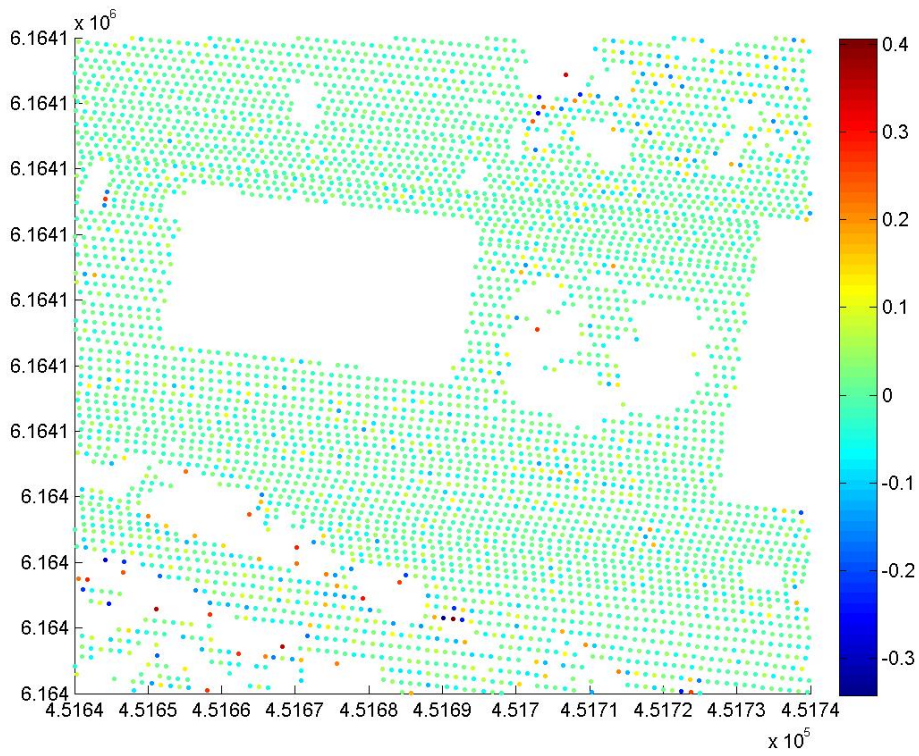
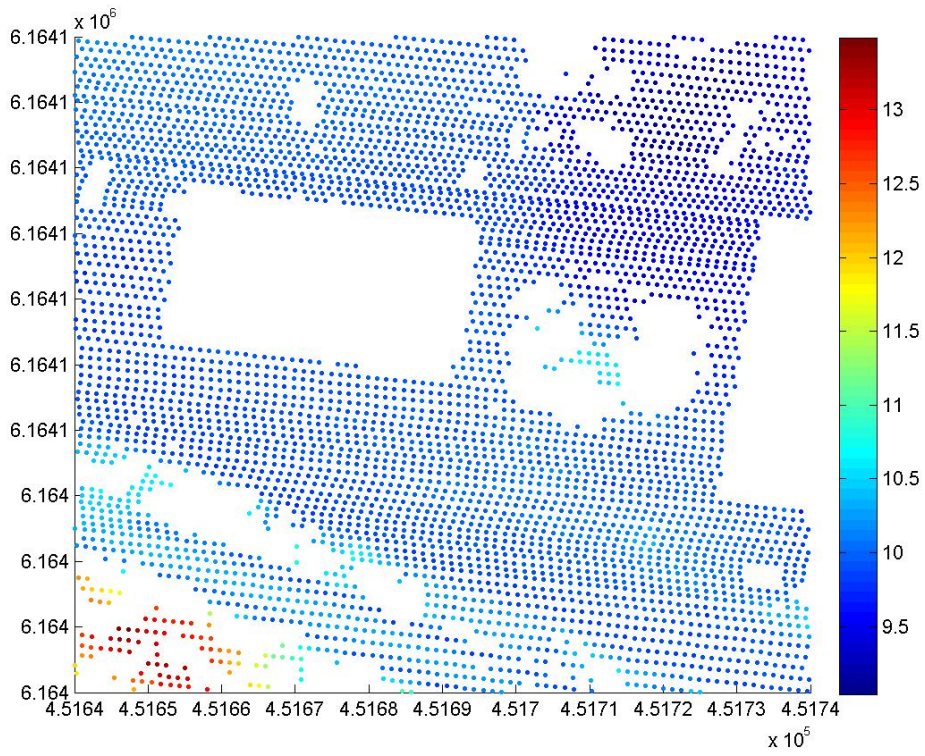
OKSe_2007



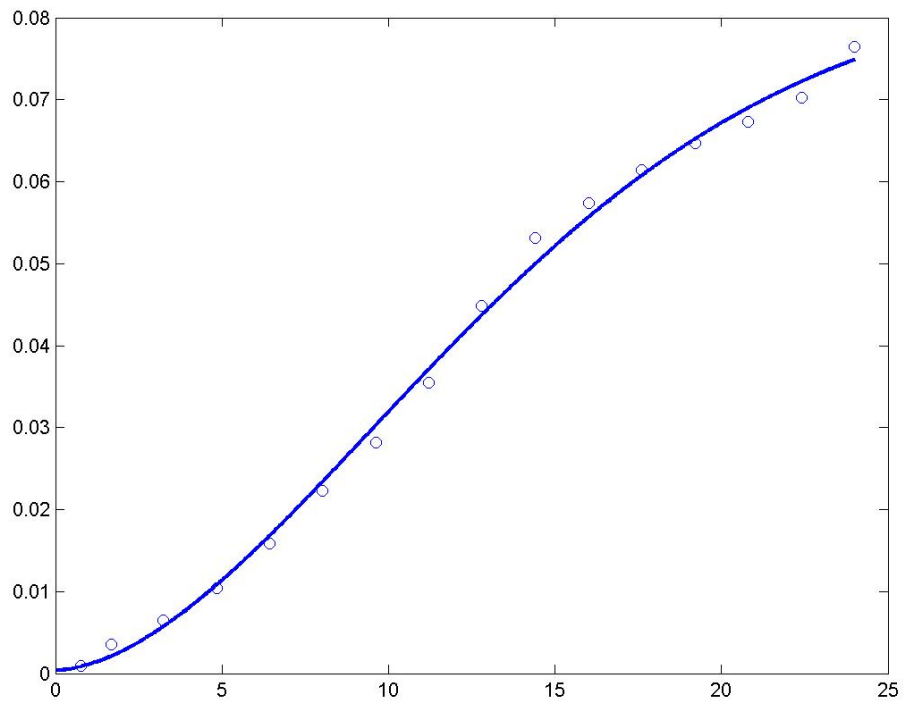
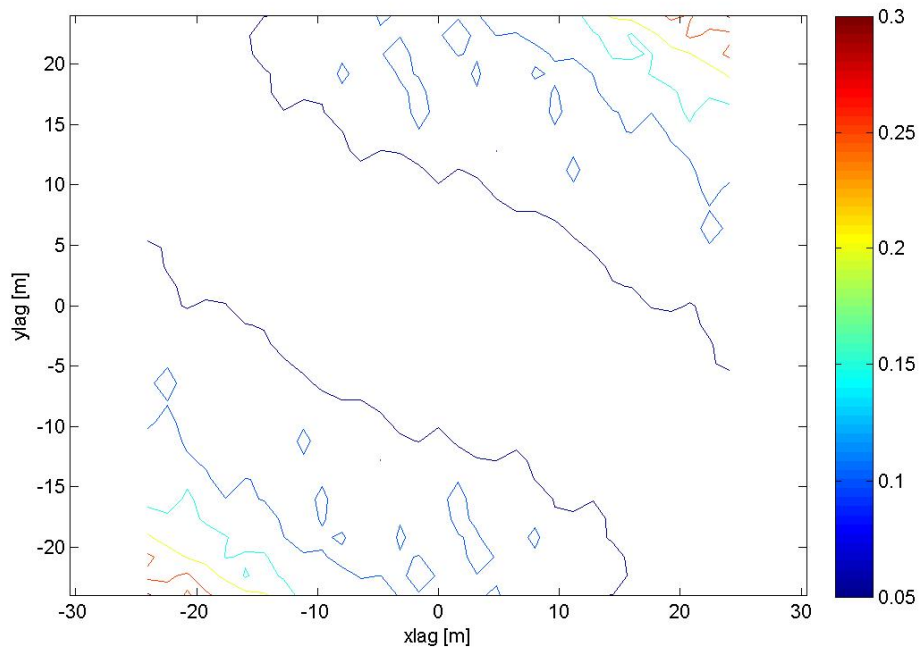
Bias for NN is -0.00179 m, and for TIN -0.00088283 m. RMSE for NN is 0.15491 m, and for TIN 0.11269 m.



OKSf_2007

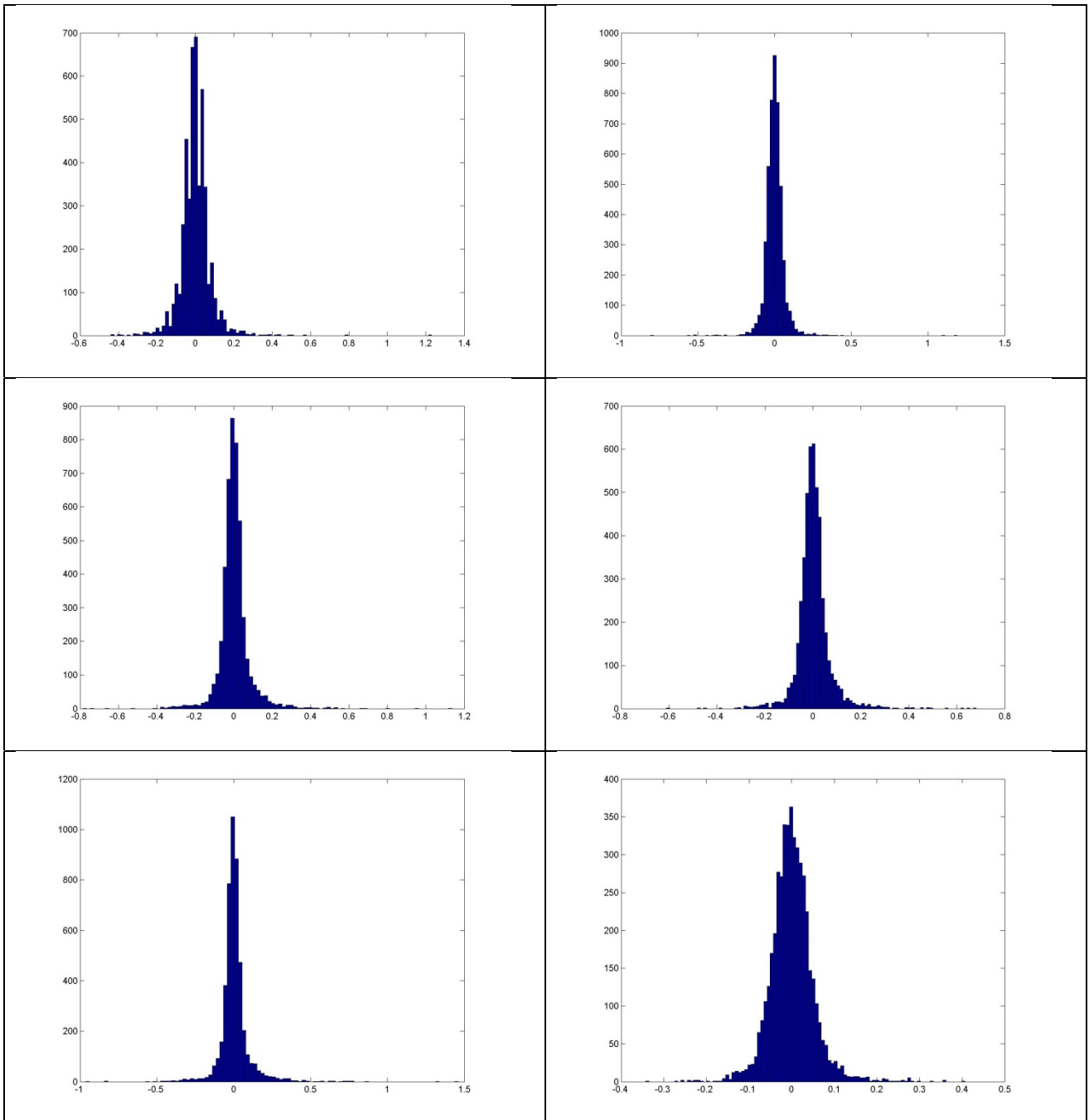


OKSf_2007

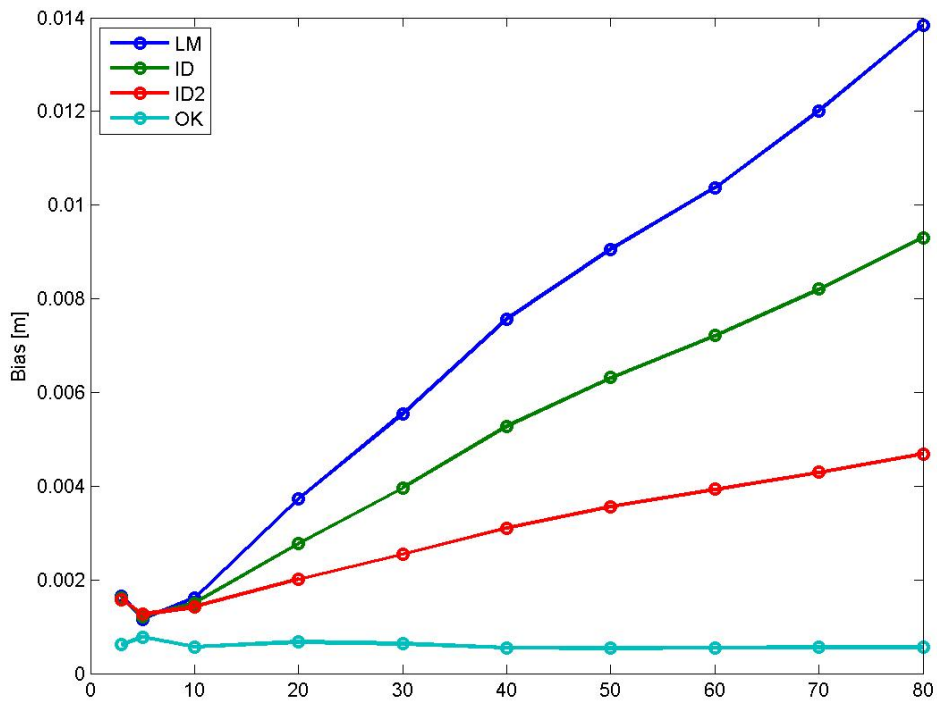
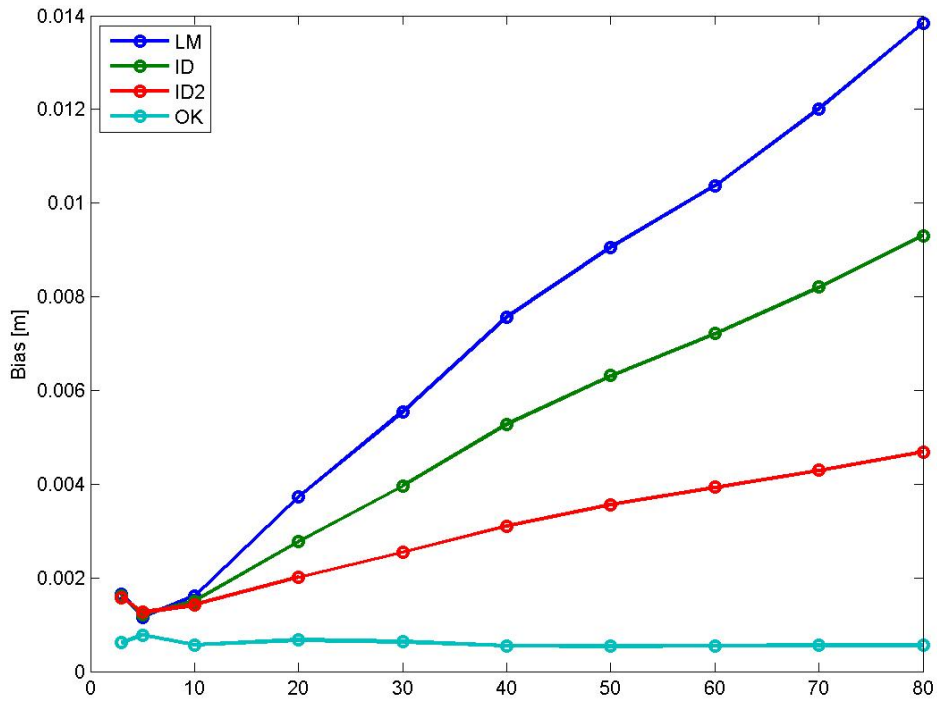


A modified Gaussian semivariogram model is used (power is 1.75 and not 2).

OKSf_2007

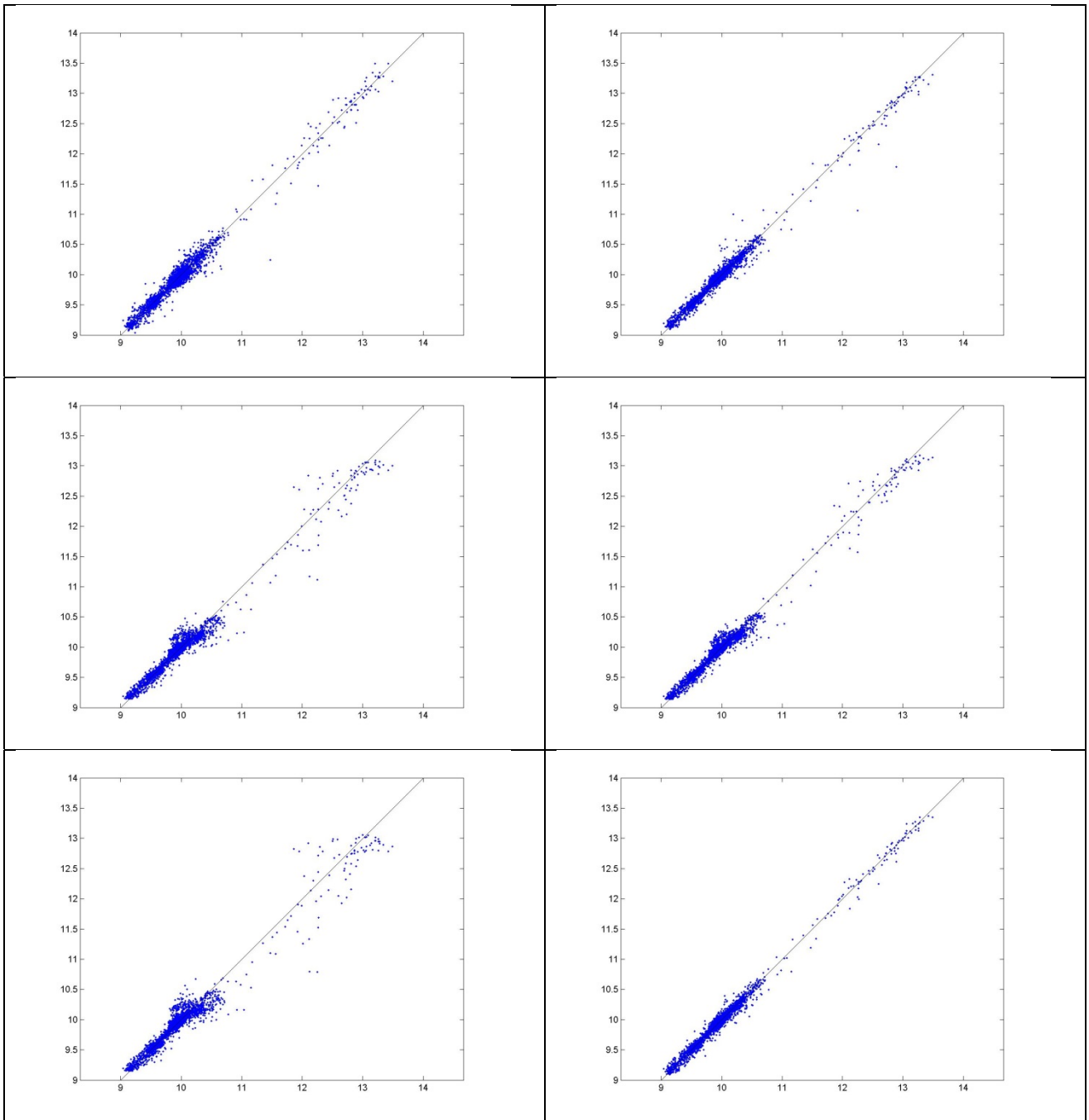


OKSf_2007

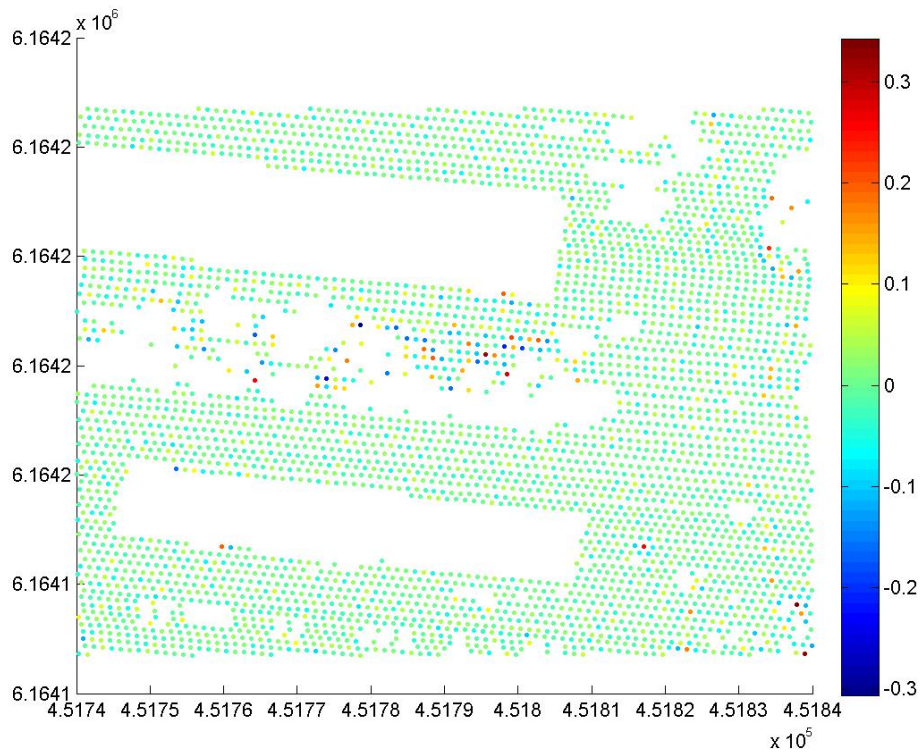
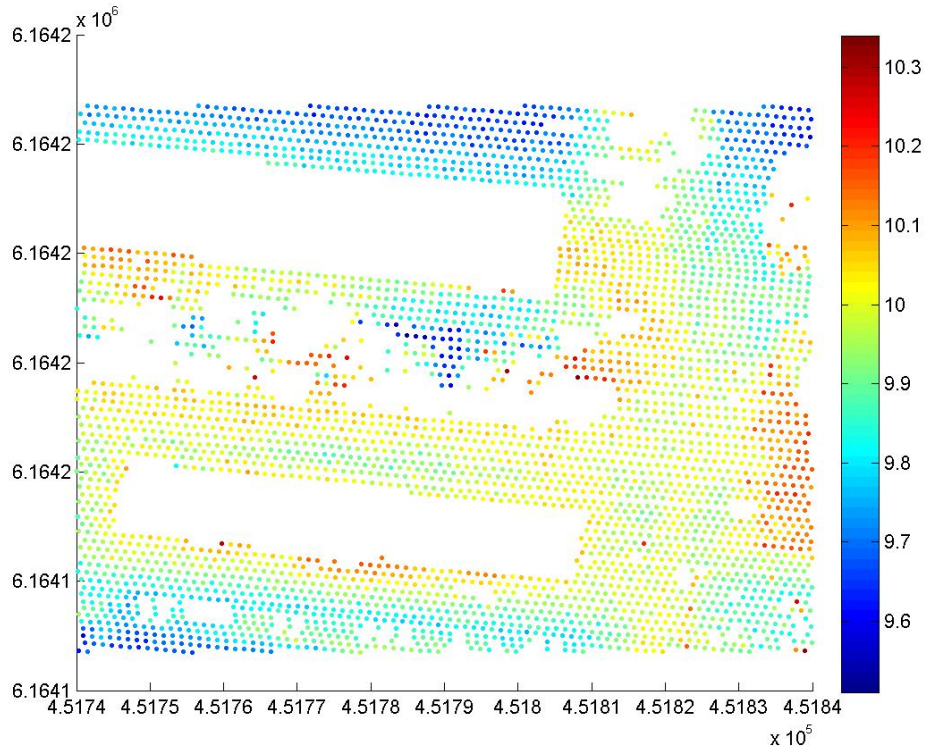


Bias for NN is 0.001618 m, and for TIN -9.9869e-05 m. RMSE for NN is 0.075206 m, and for TIN 0.06266 m.

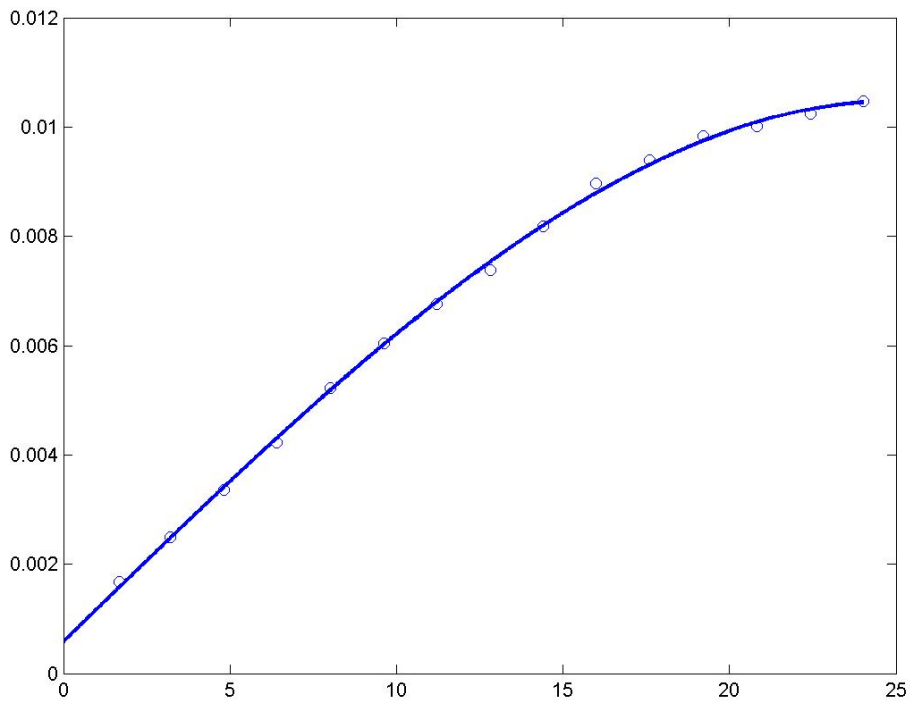
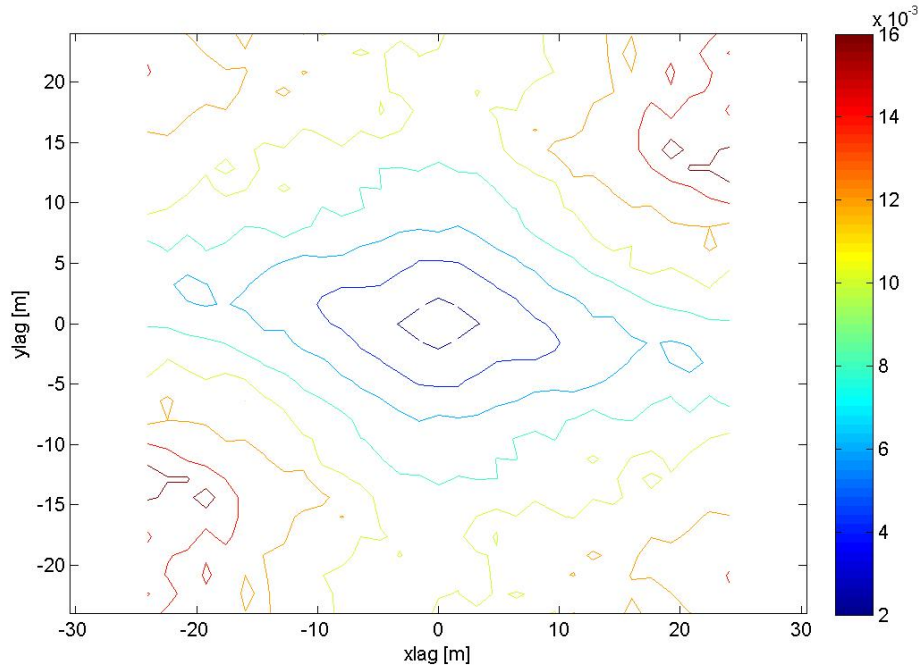
OKSf_2007



OKSg_2007

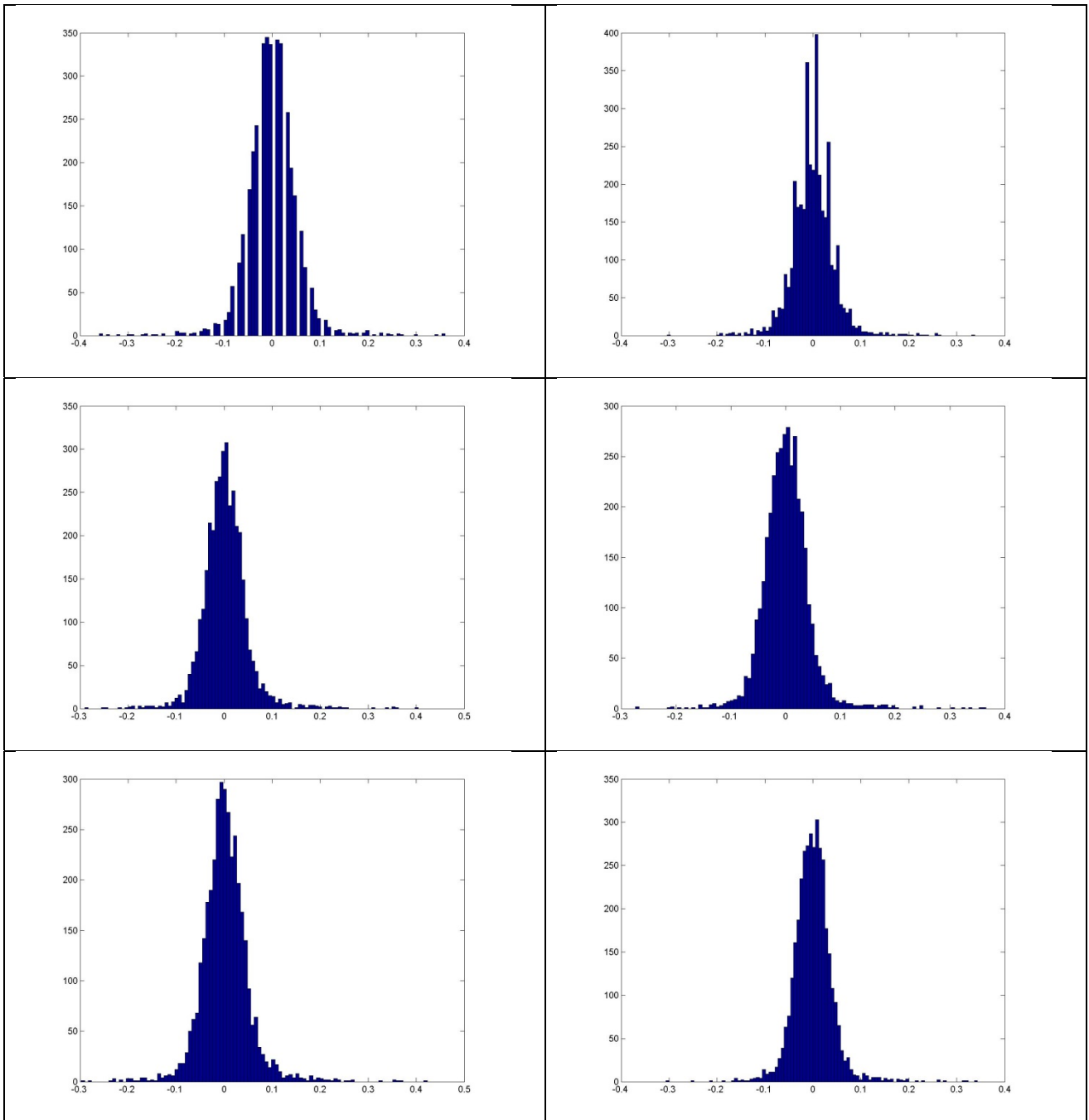


OKSg_2007

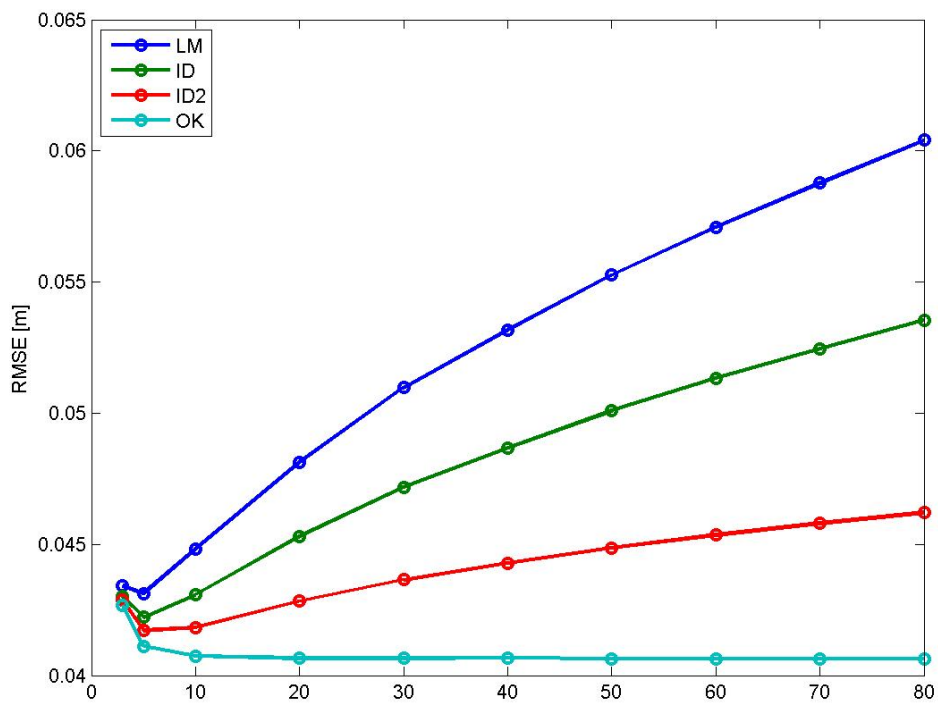
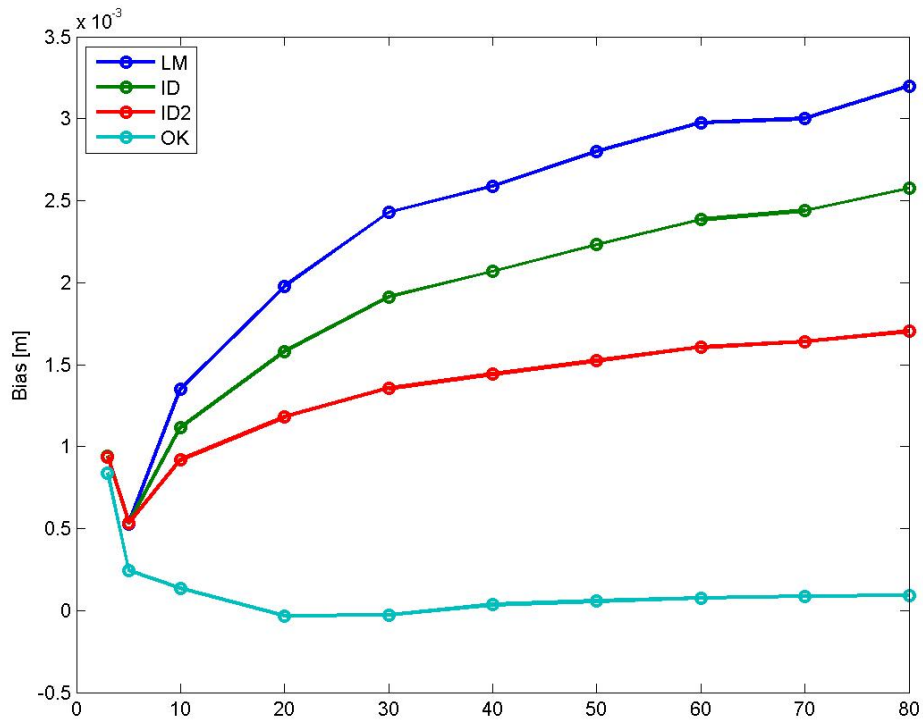


A spherical semivariogram model is used.

OKSg_2007

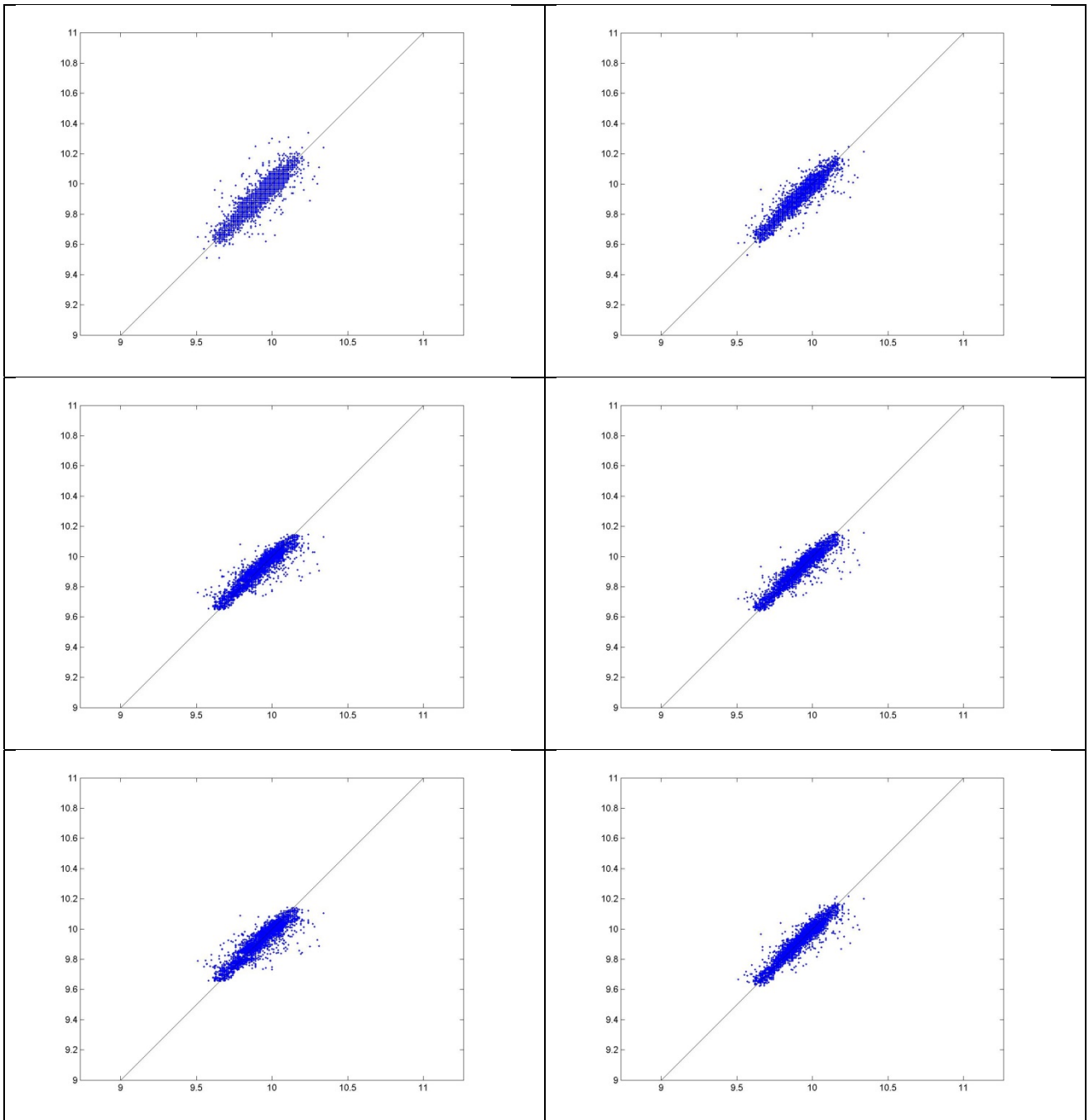


OKSg_2007

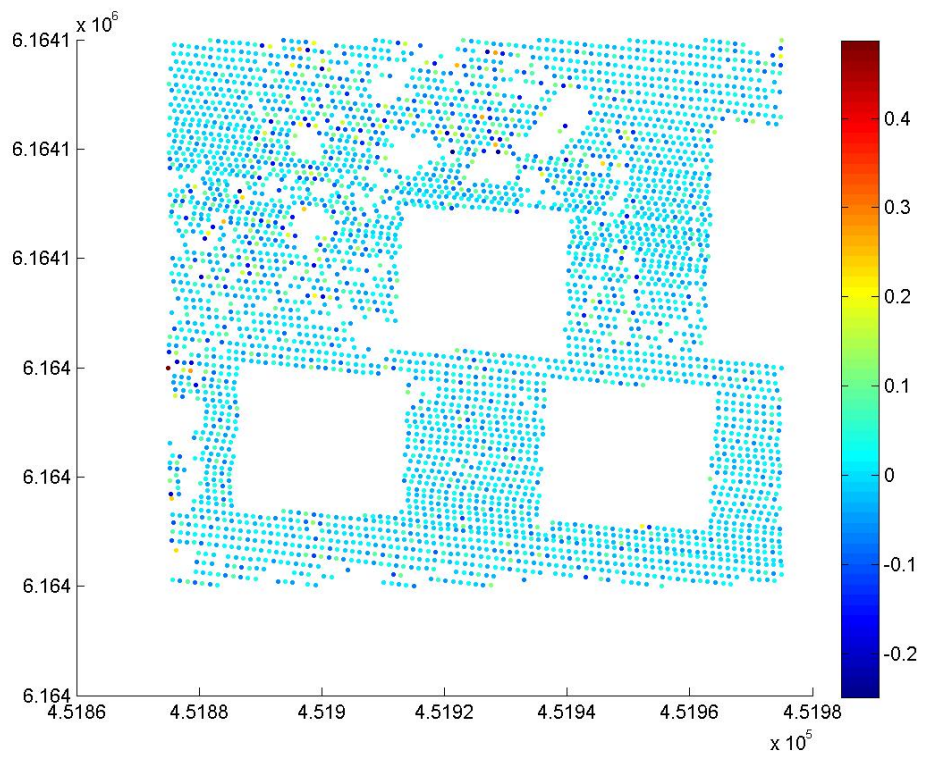
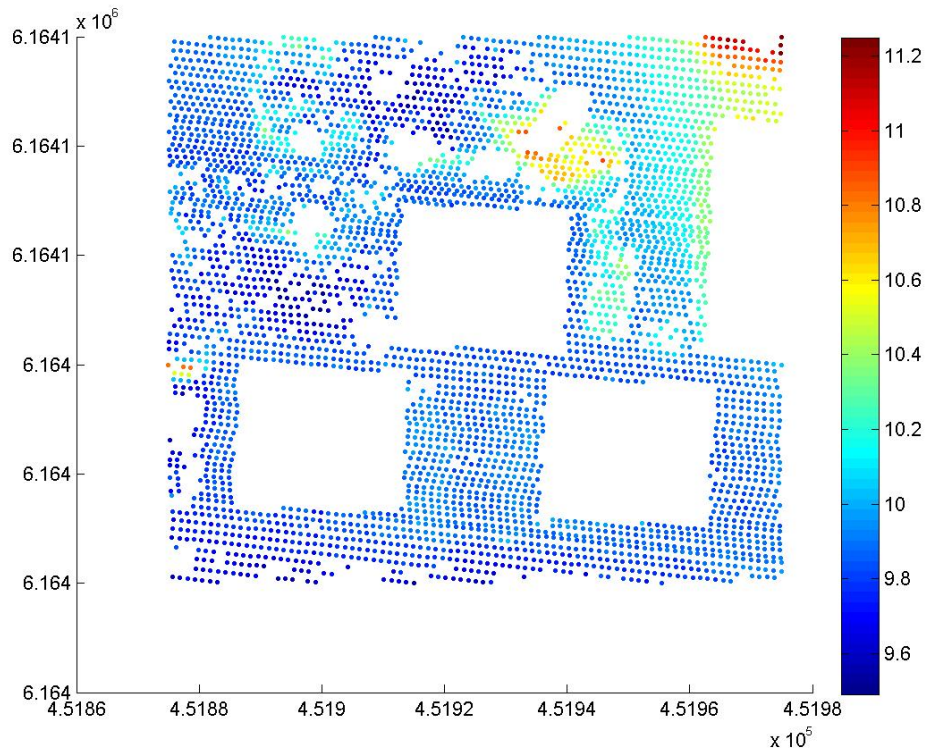


Bias for NN is -0.0001947 m, and for TIN $7.7836e-05$ m. RMSE for NN is 0.05241 m, and for TIN 0.04339 m.

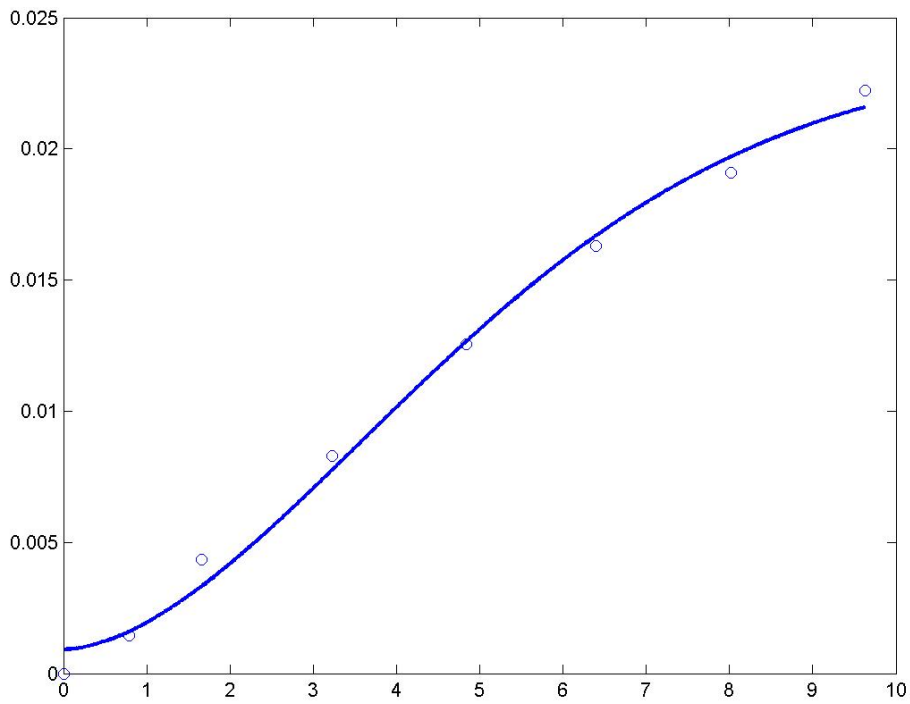
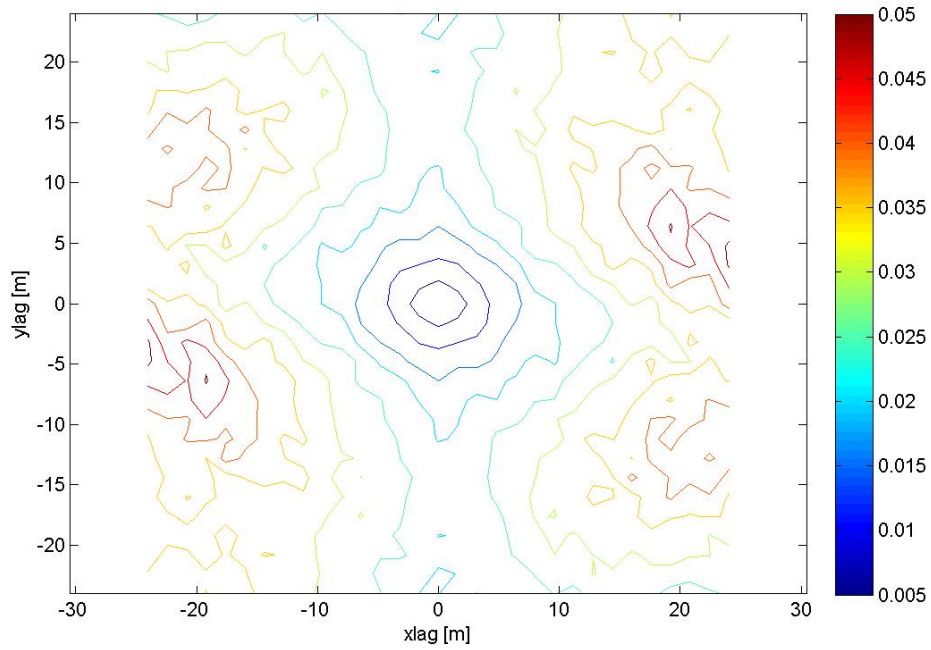
OKSg_2007



OKSh_2007

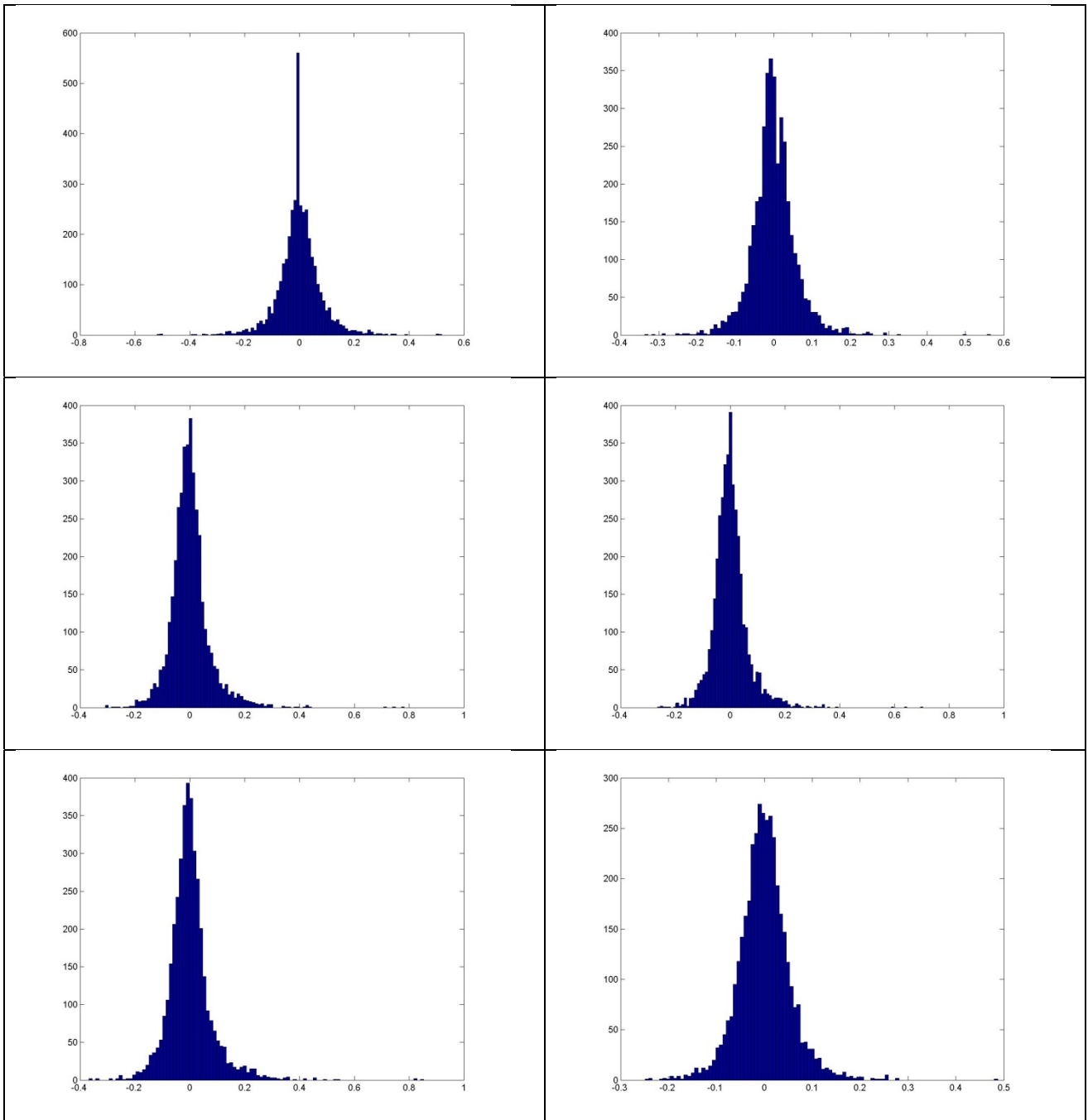


OKSh_2007

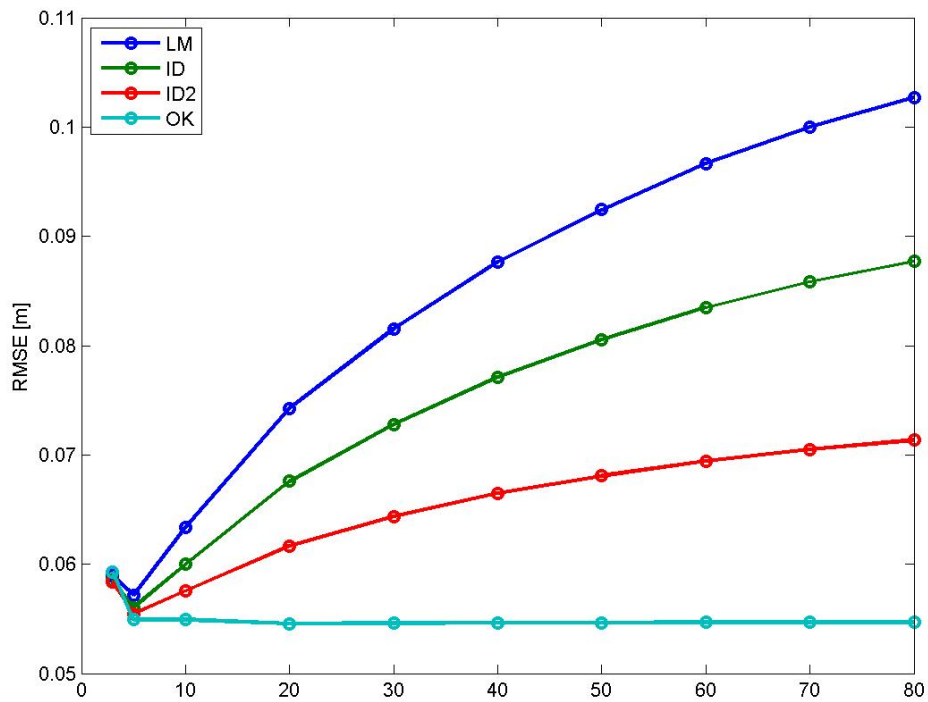
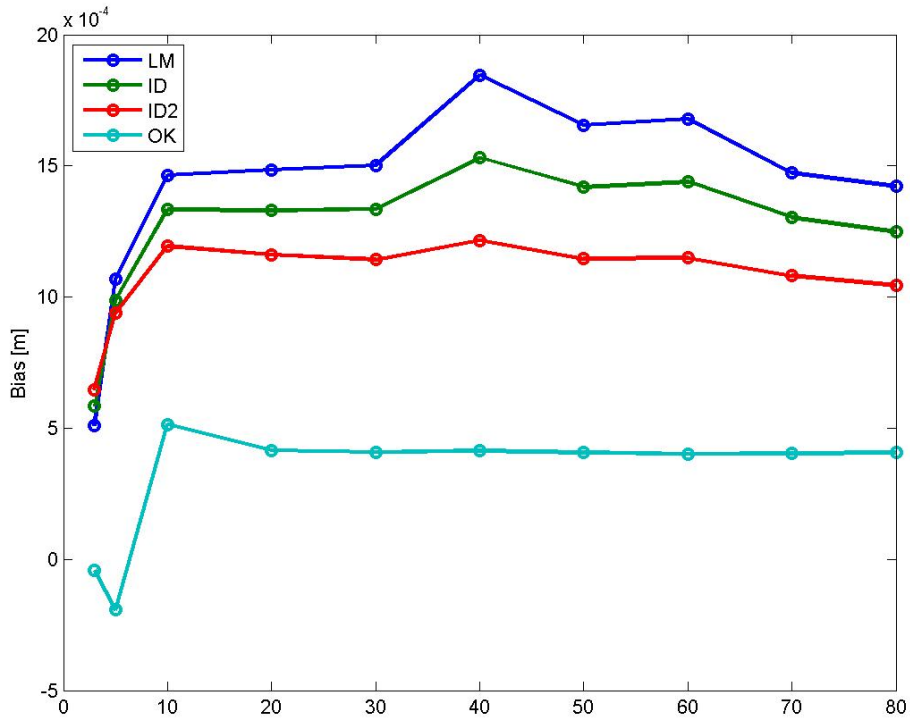


A modified Gaussian semivariogram model is used (power is 1.75 and not 2).

OKSh_2007

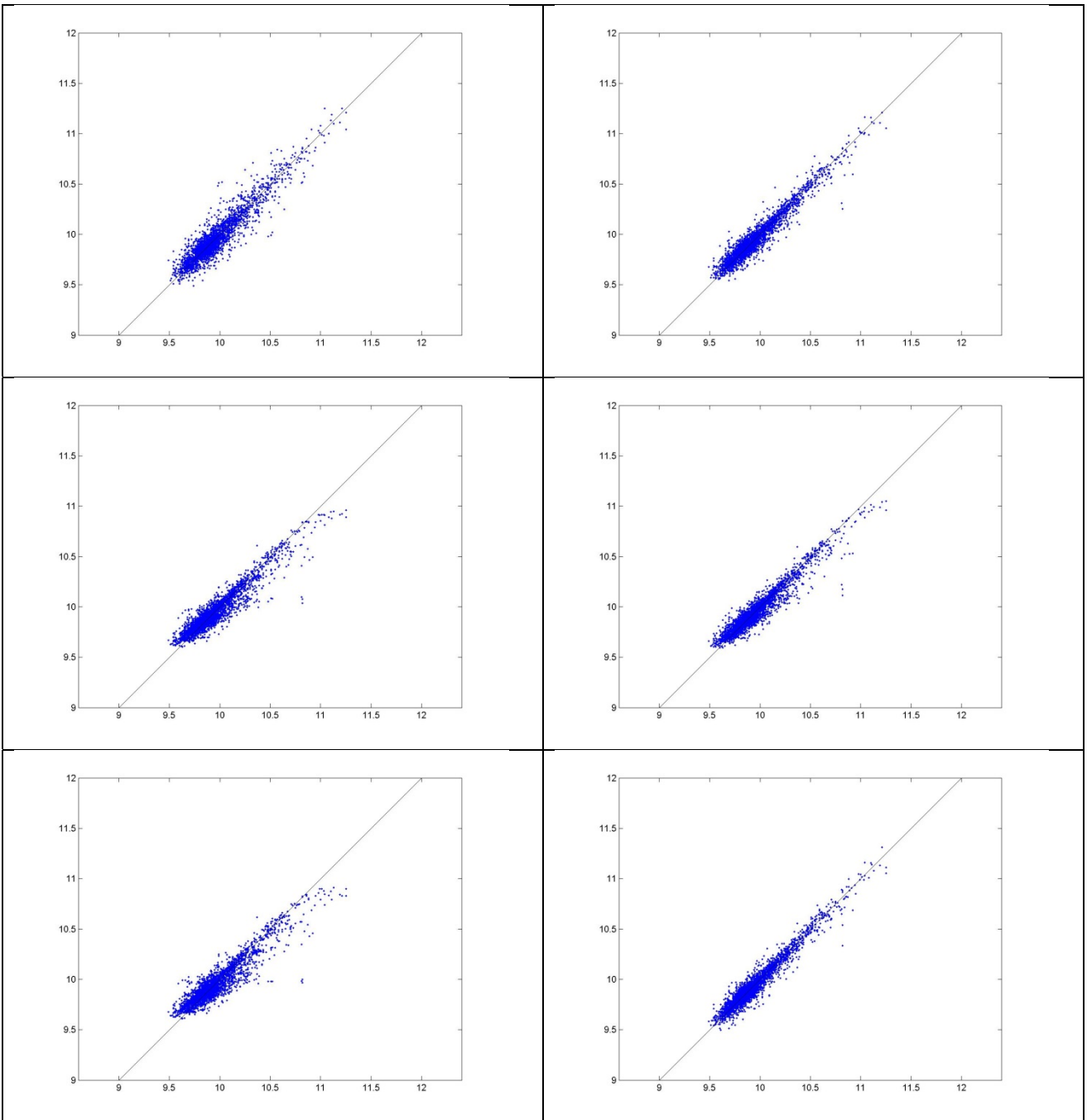


OKSh_2007

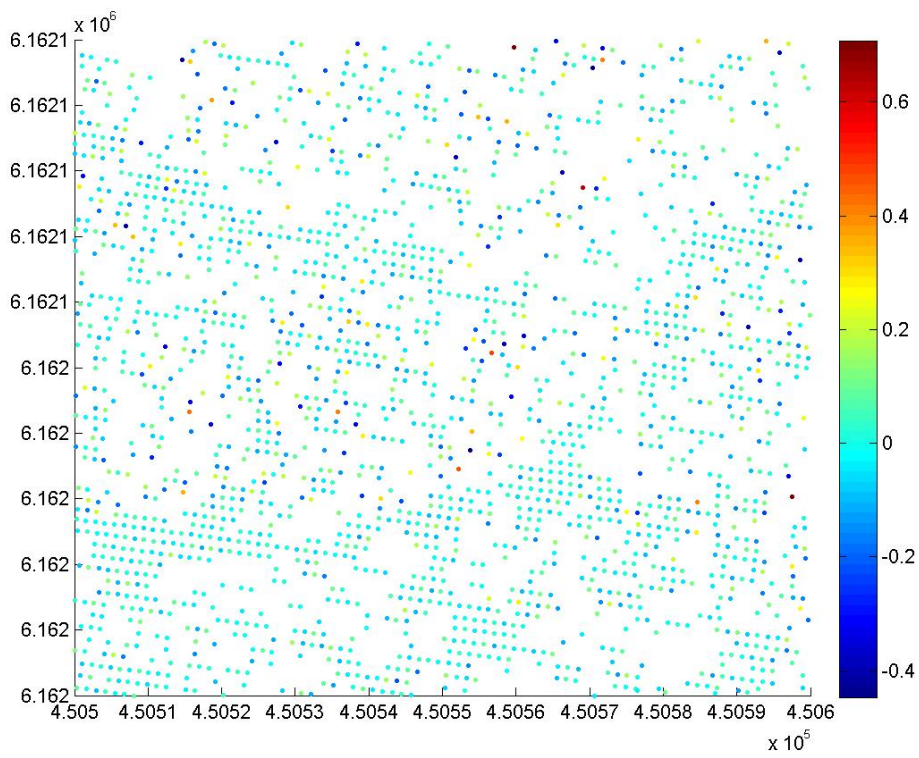
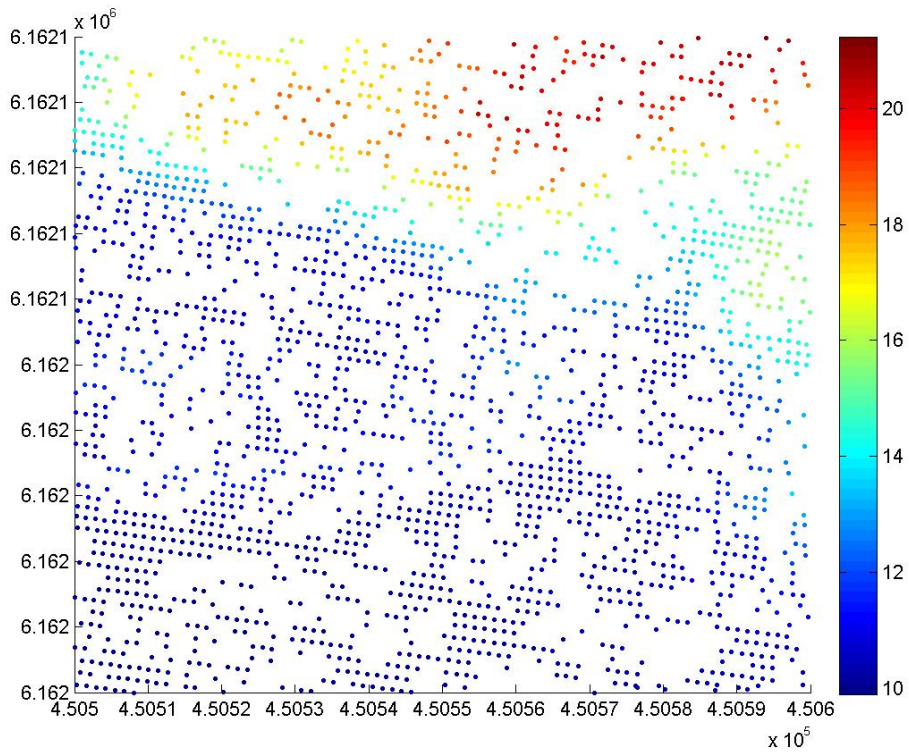


Bias for NN is 0.0006564 m, and for TIN -0.0008640 m. RMSE for NN is 0.078595 m, and for TIN 0.05876 m.

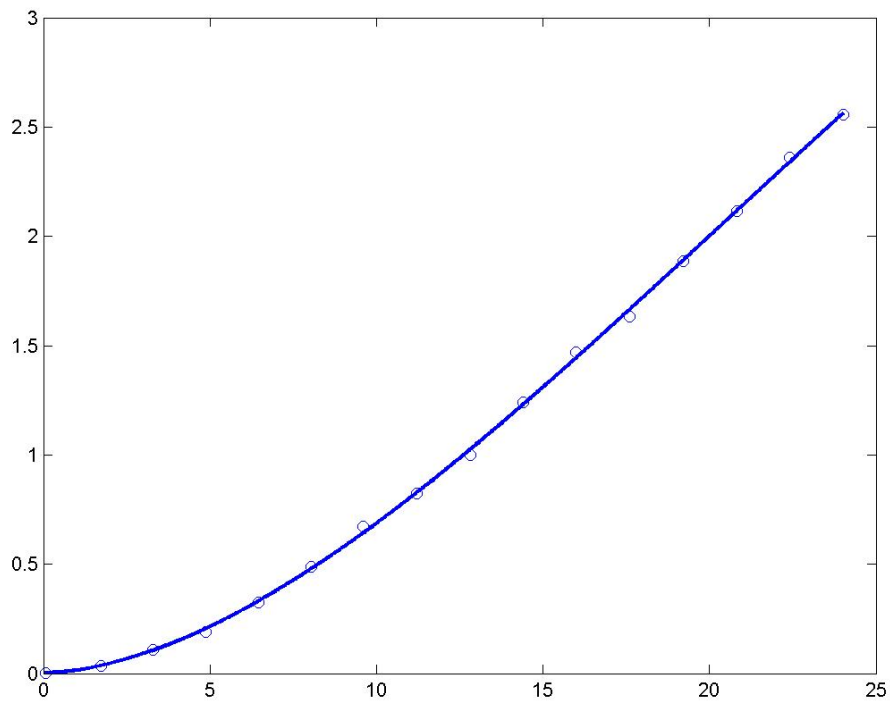
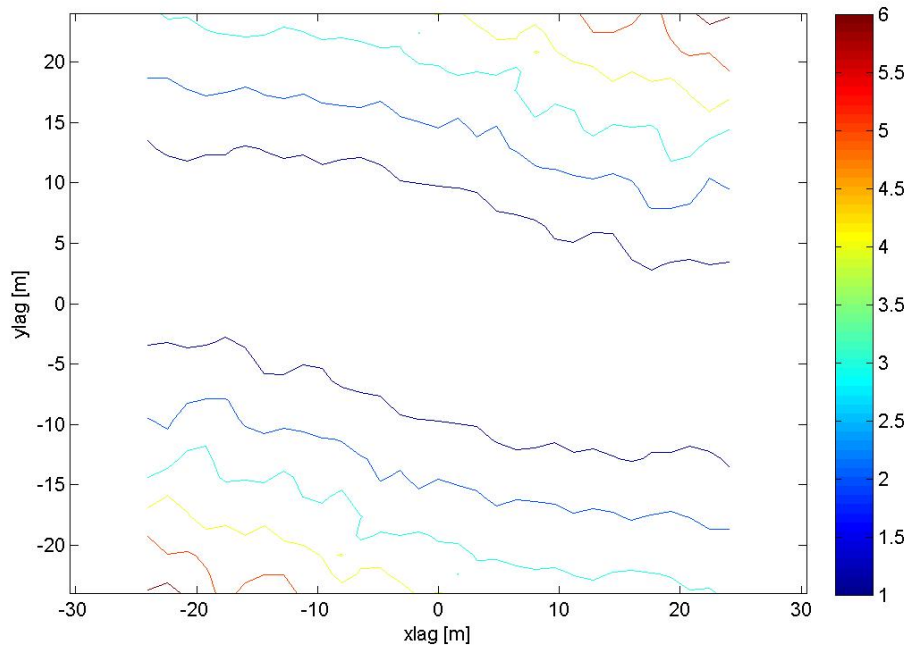
OKSh_2007



OKSx_2007

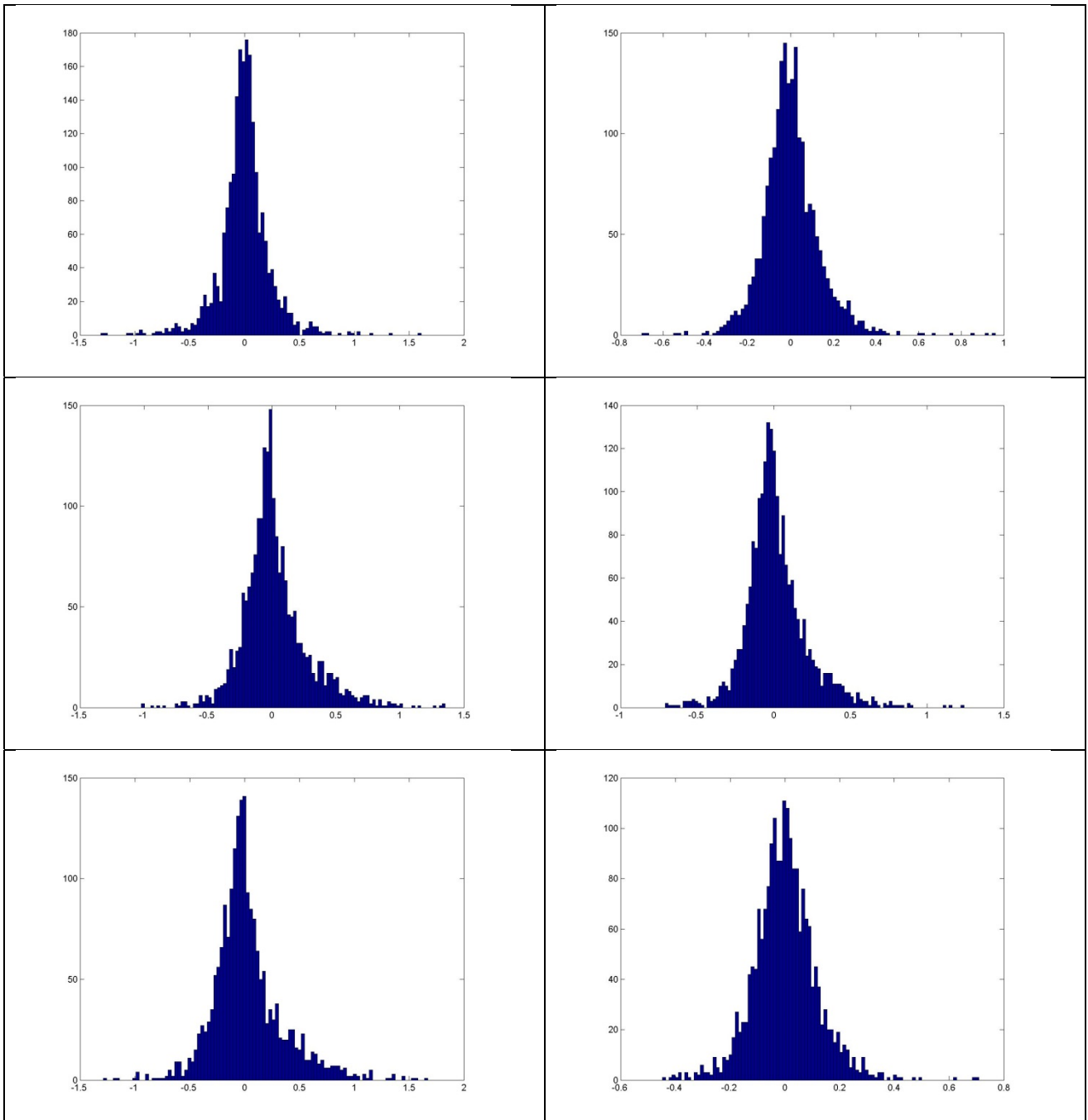


OKSx_2007

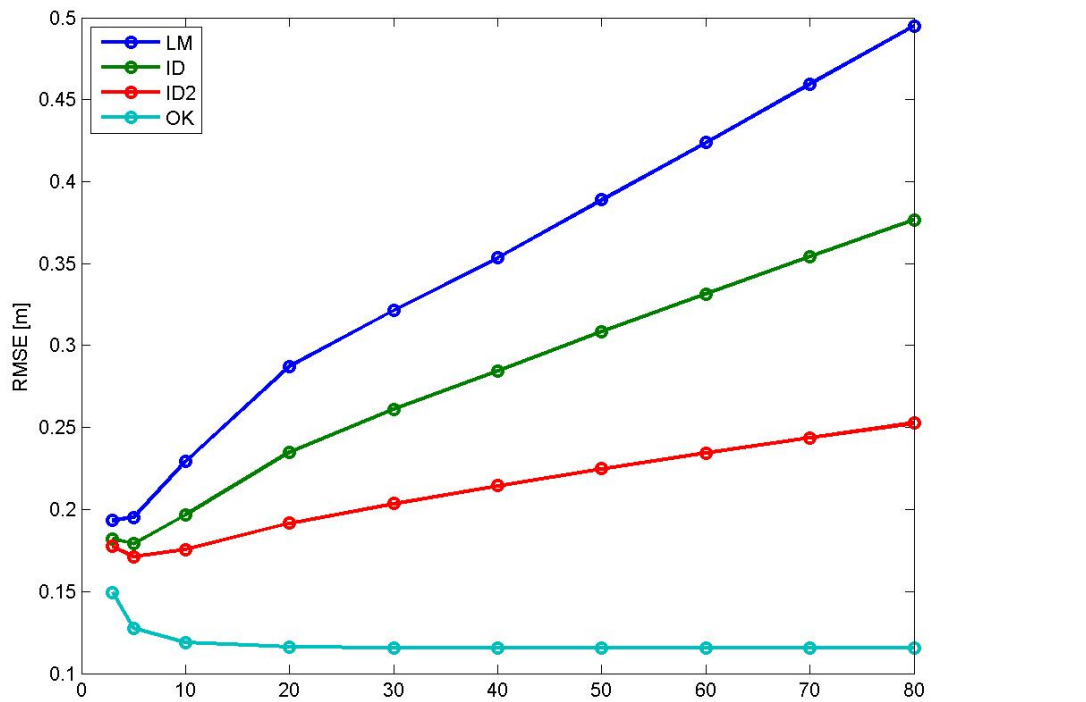
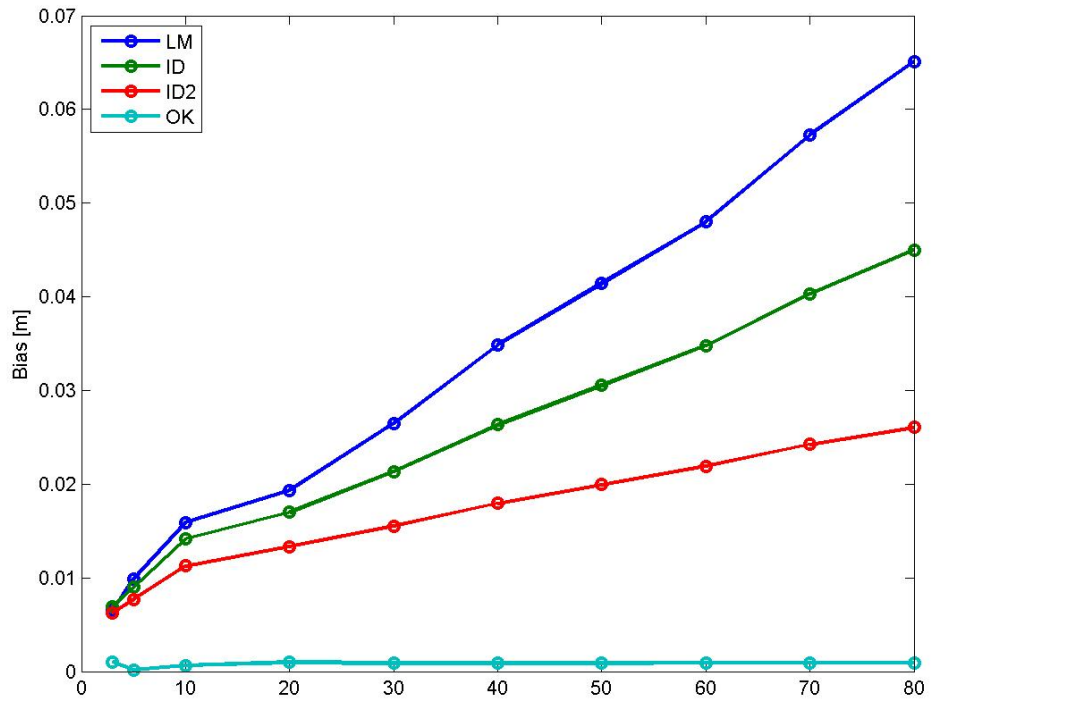


A modified Gaussian semivariogram model is used (power is 1.75 and not 2).

OKSx_2007

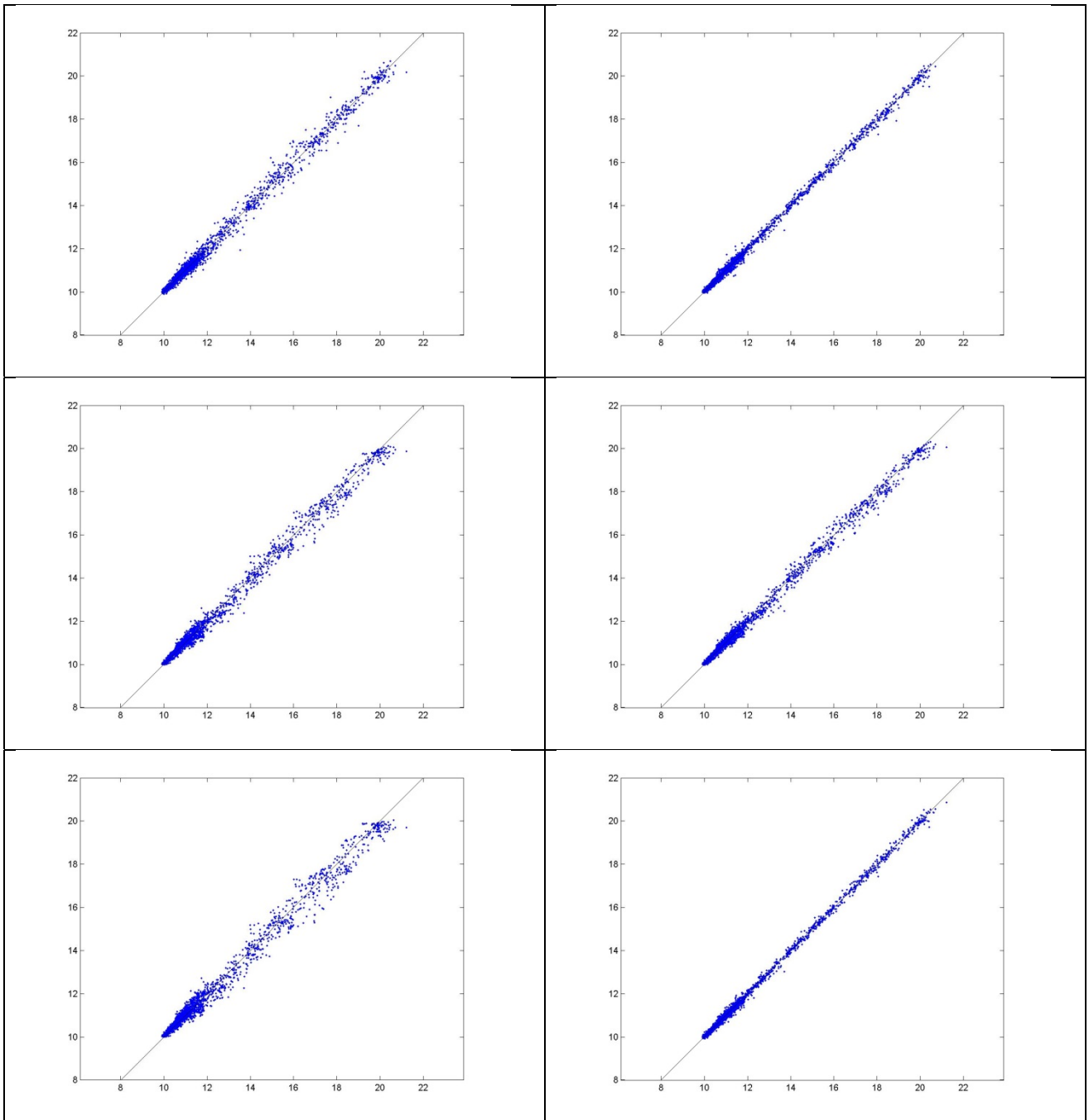


OXSx_2007

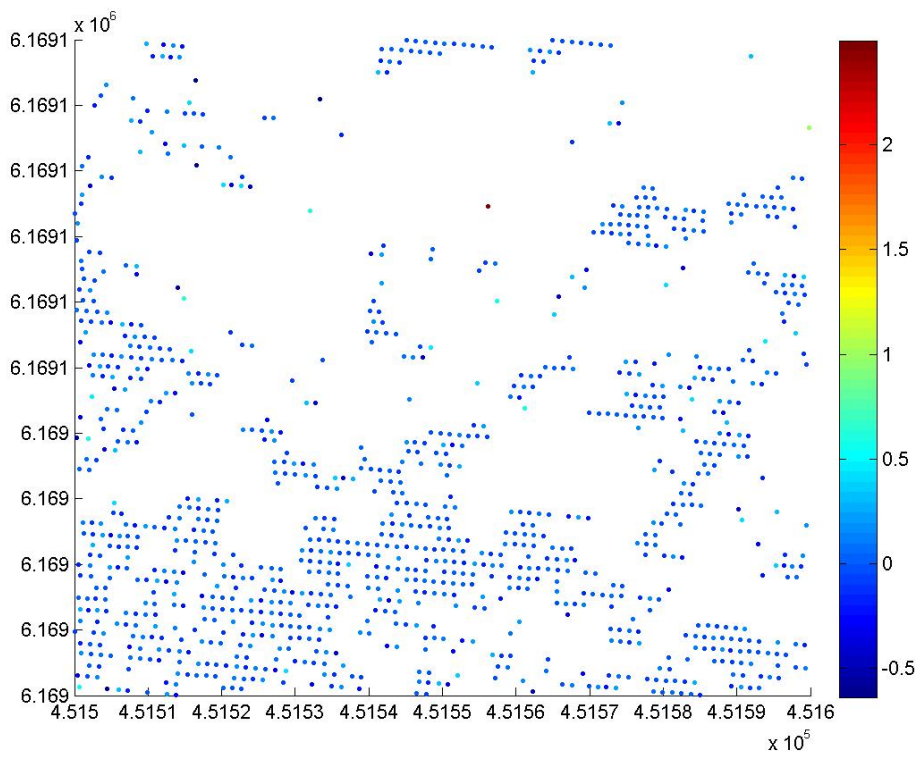
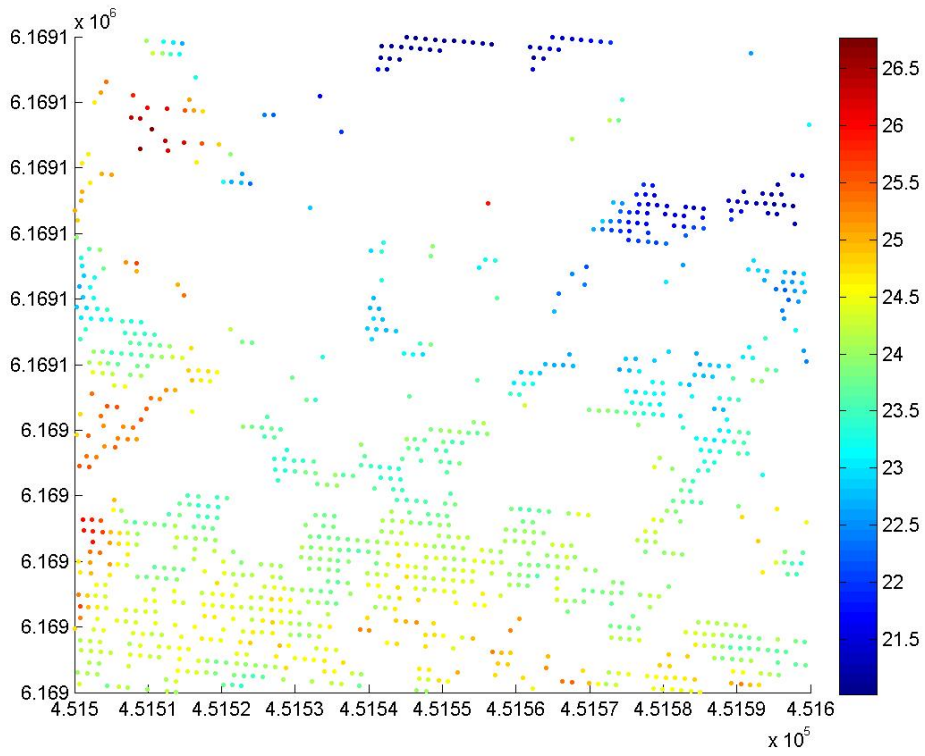


Bias for NN is 0.0014258 m, and for TIN 0.00068106 m. RMSE for NN is 0.2273 m, and for TIN 0.13433 m.

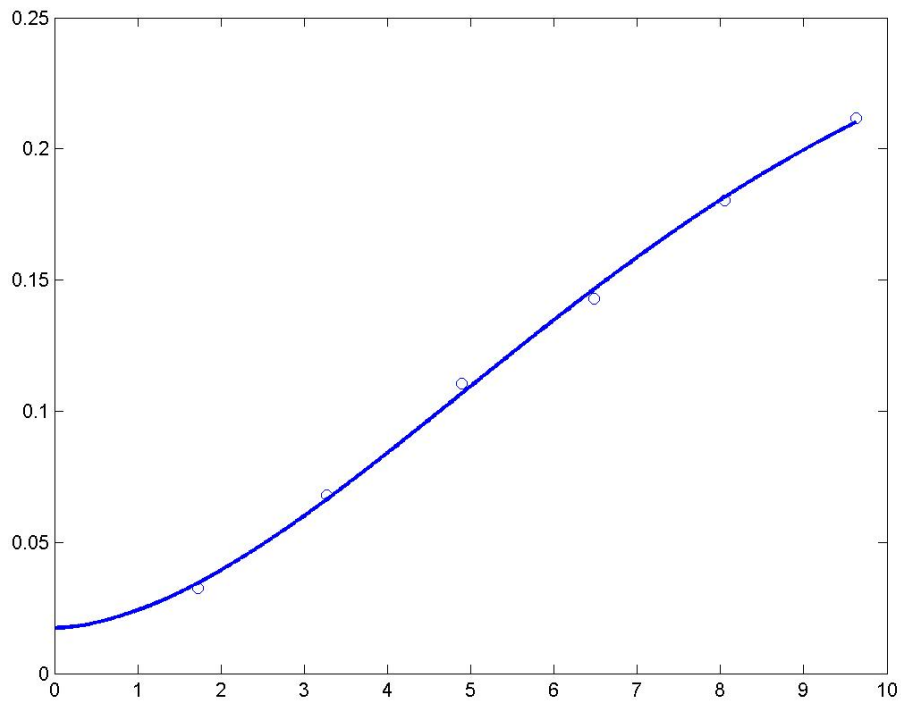
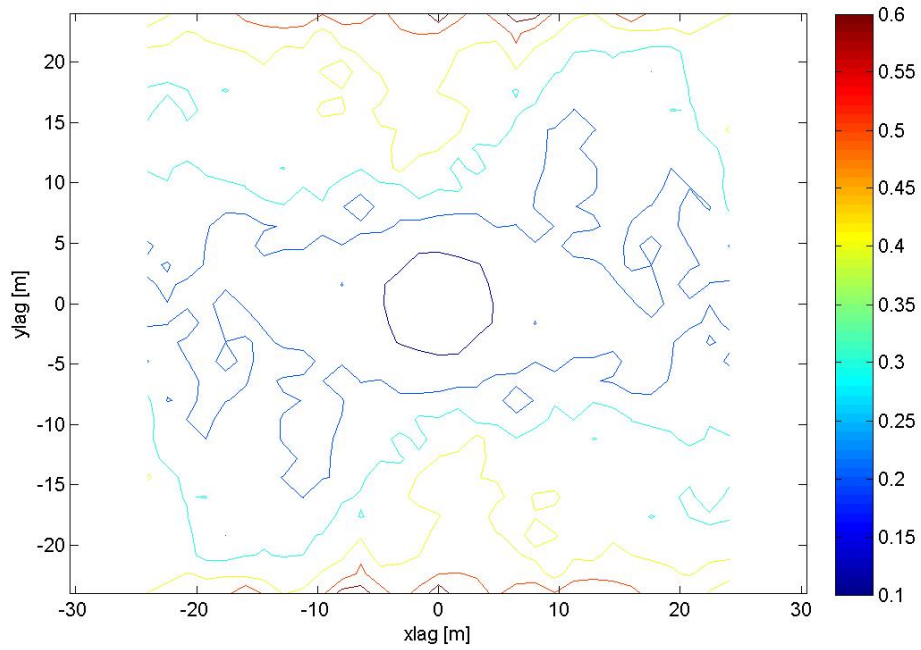
OKSx_2007



OKSy_2007

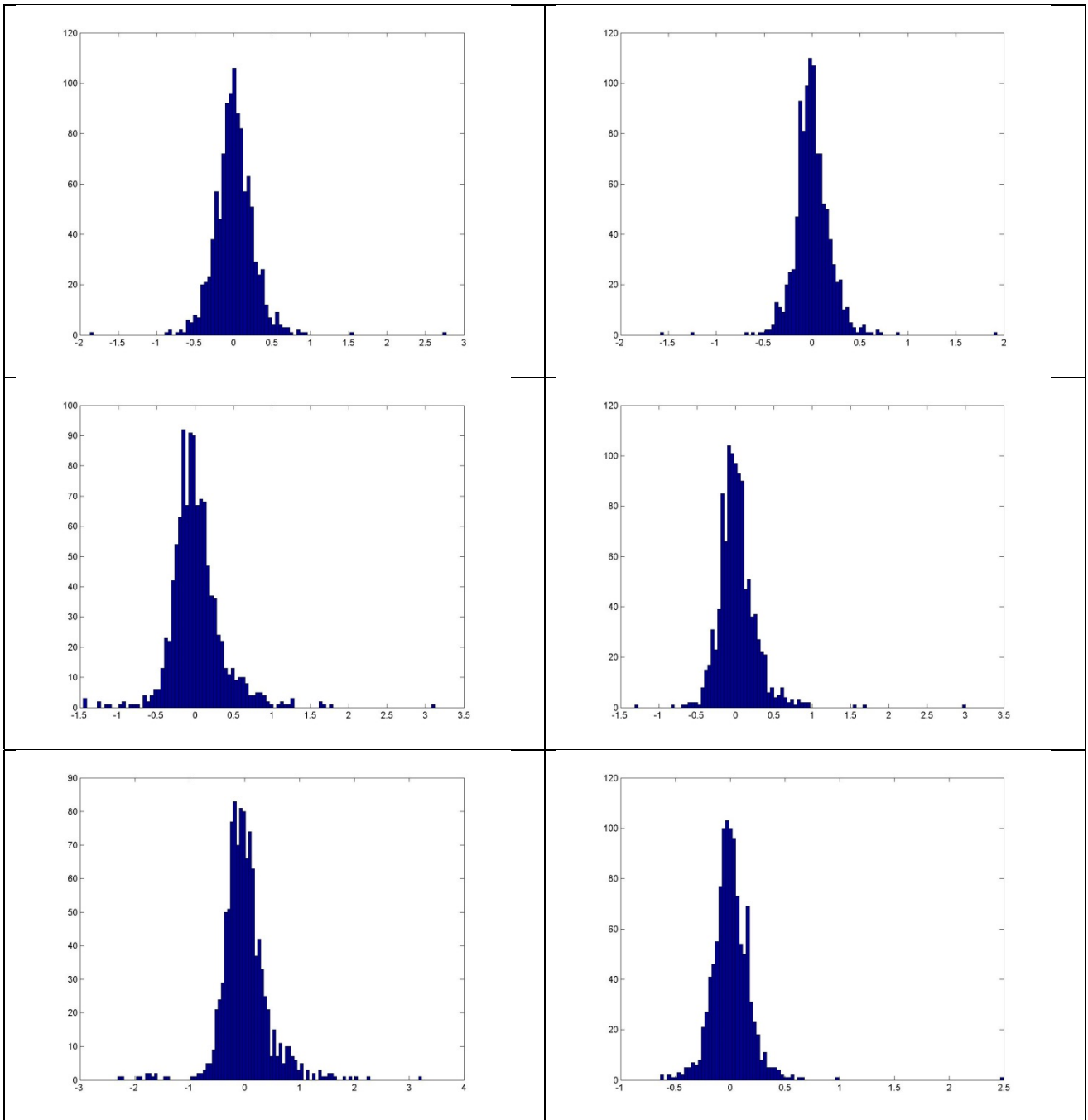


OKSy_2007

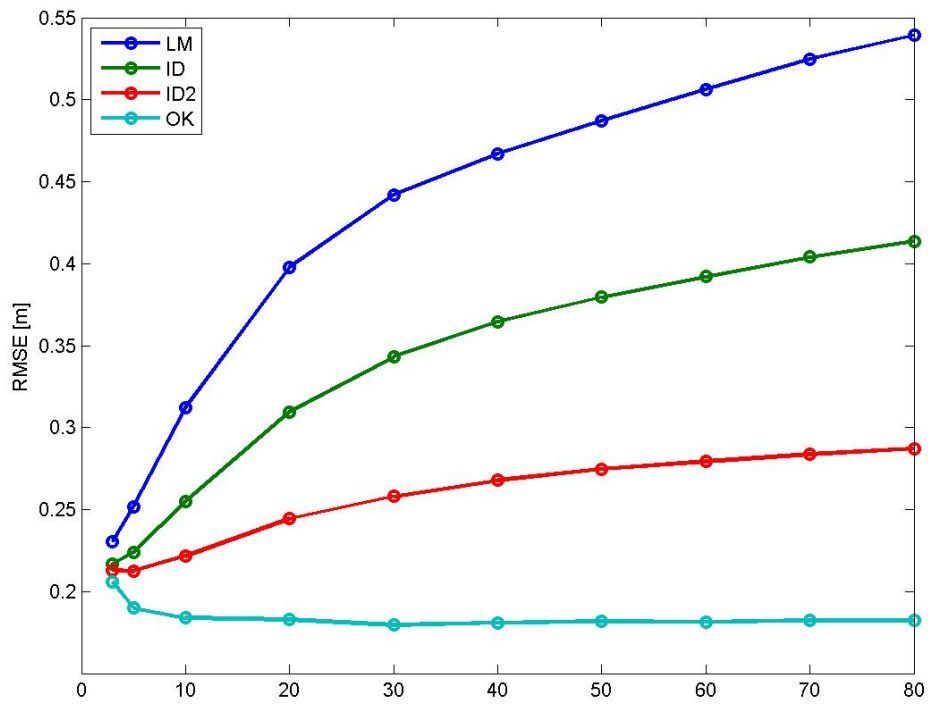
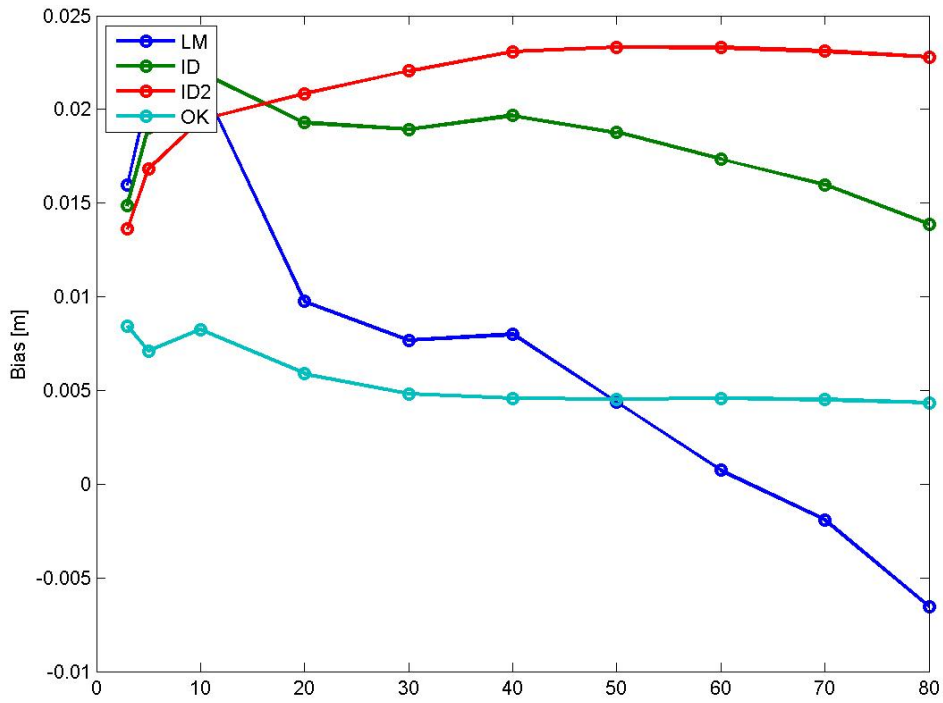


A modified Gaussian semivariogram model is used (power is 1.75 and not 2).

OKSy_2007

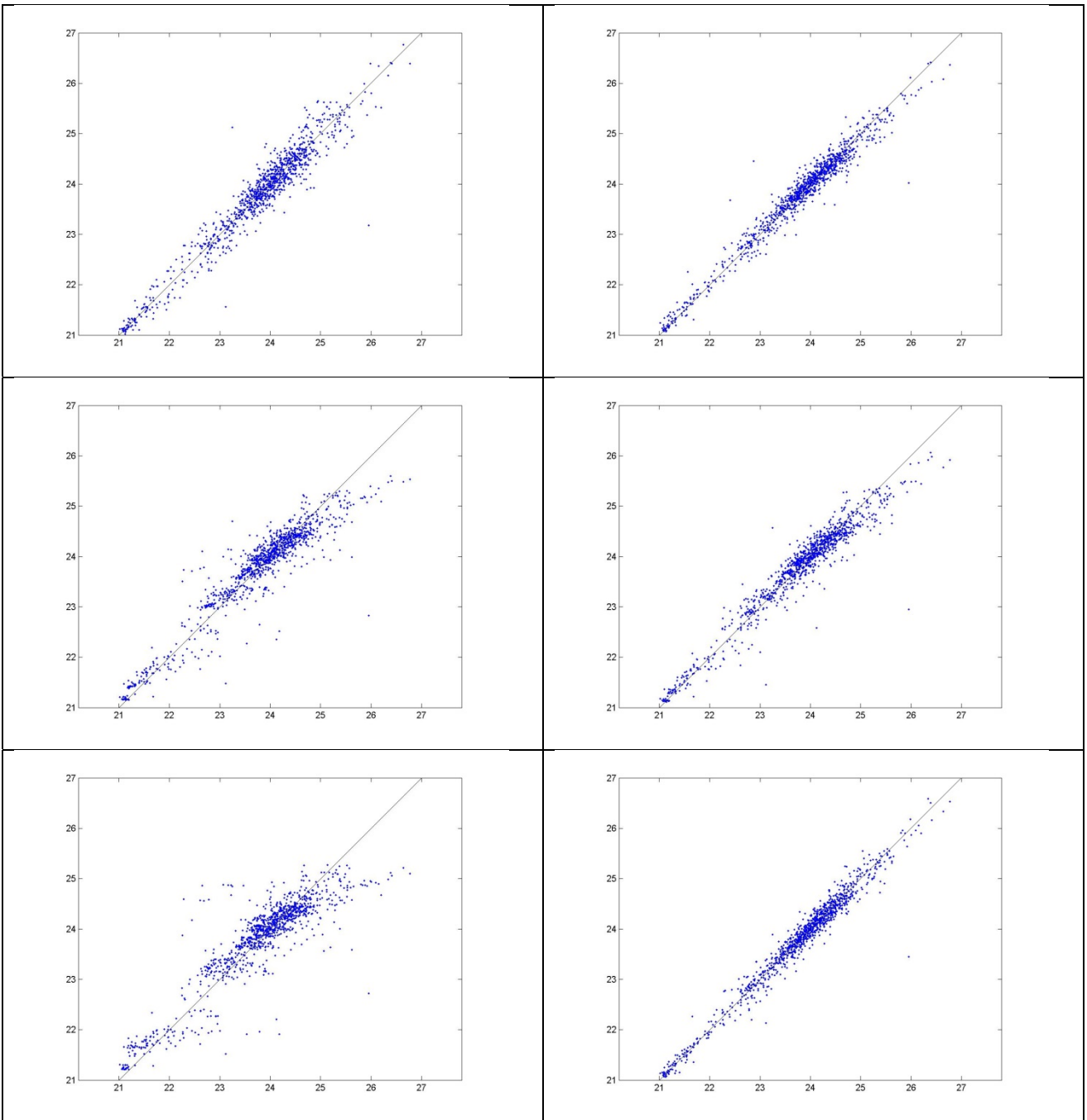


OKSy_2007

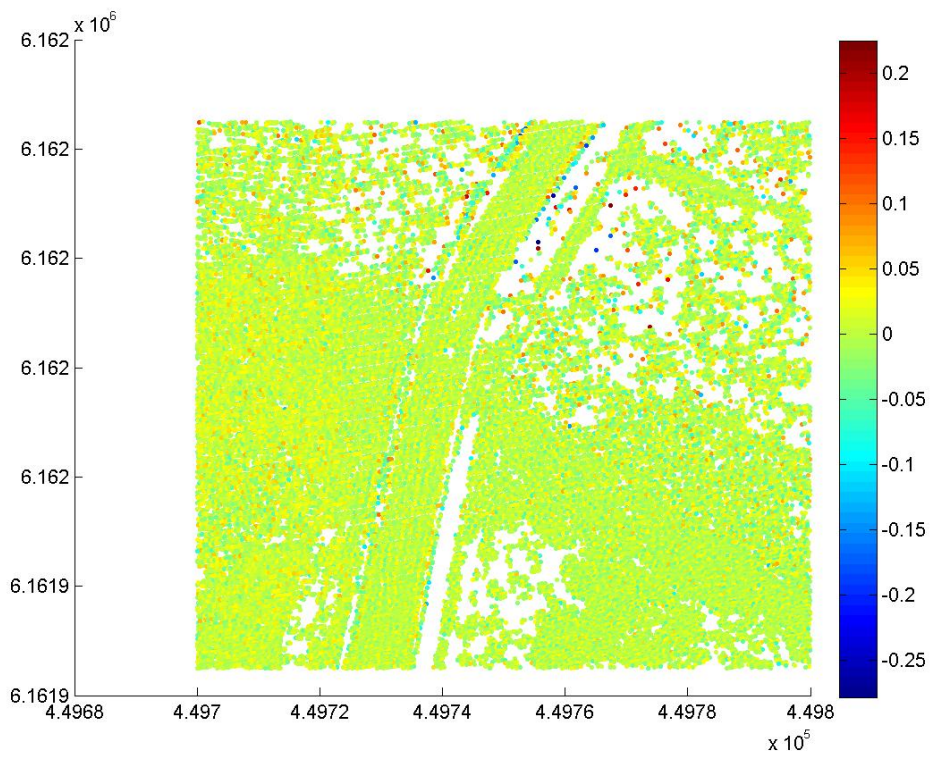
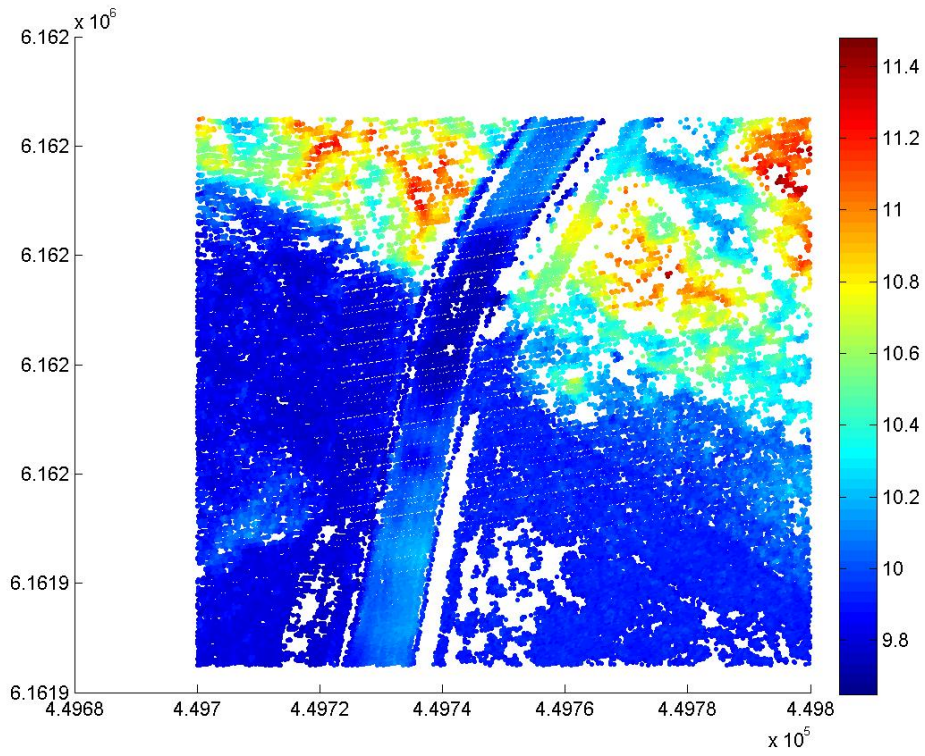


Bias for NN is 0.0069553 m, and for TIN 0.0019139 m. RMSE for NN is 0.26282 m, and for TIN 0.19465 m.

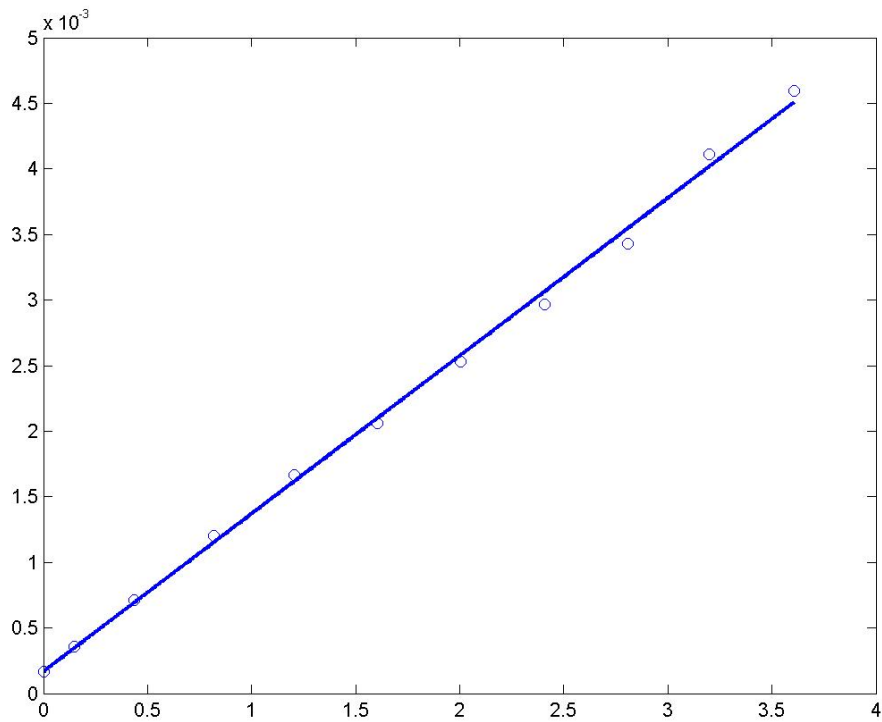
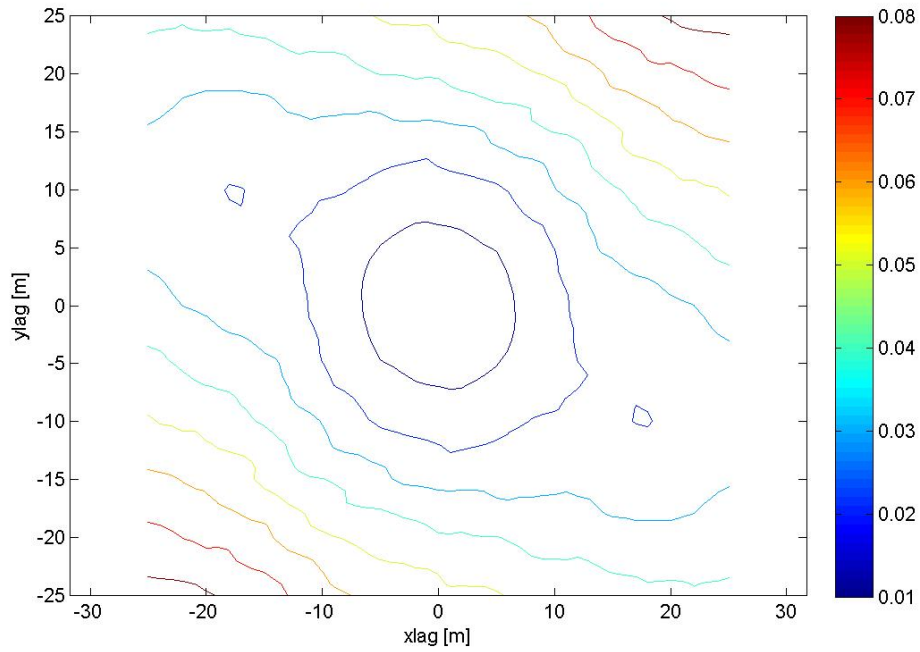
OKSy_2007



OKSa_2013

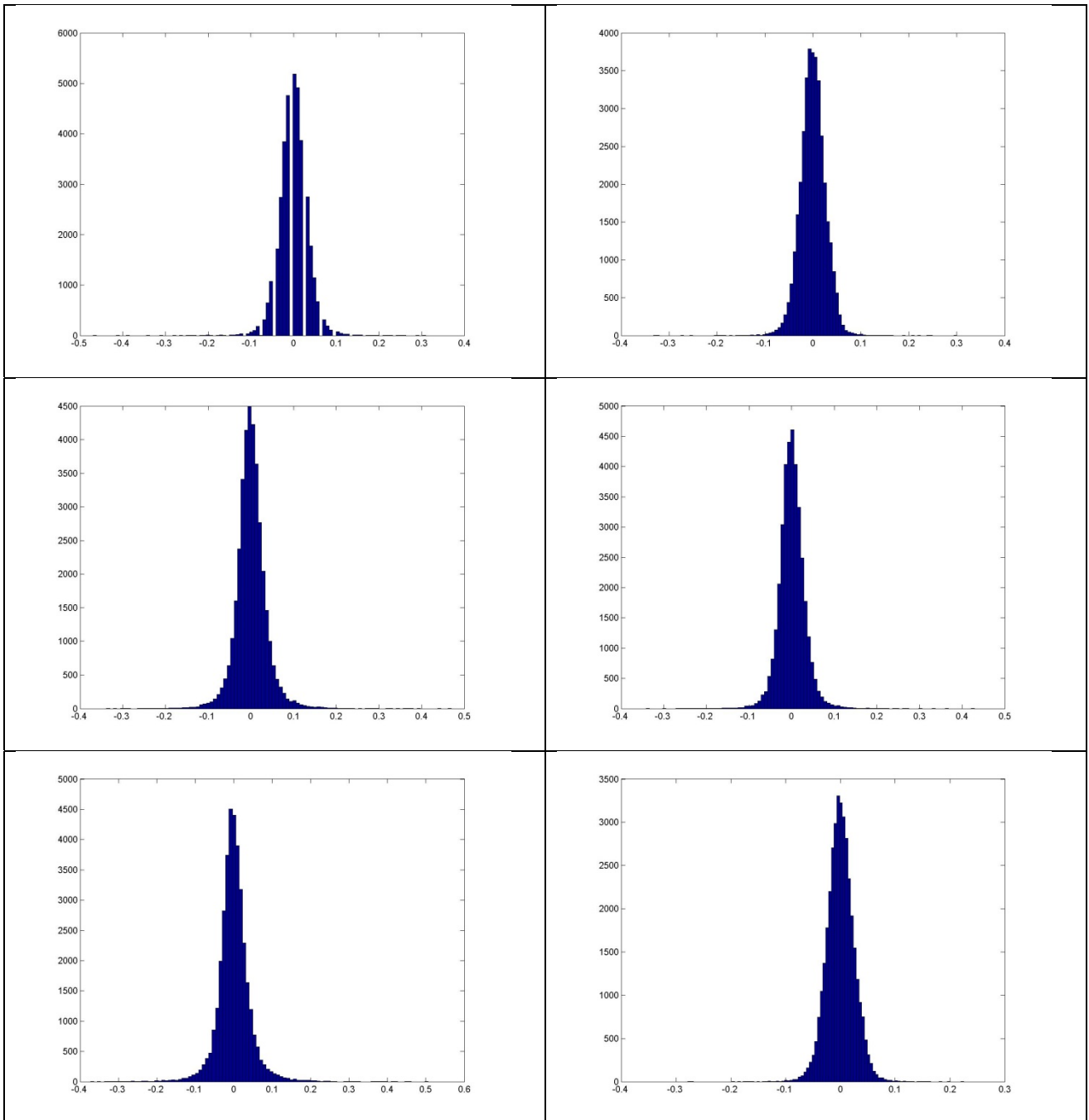


OKSa_2013

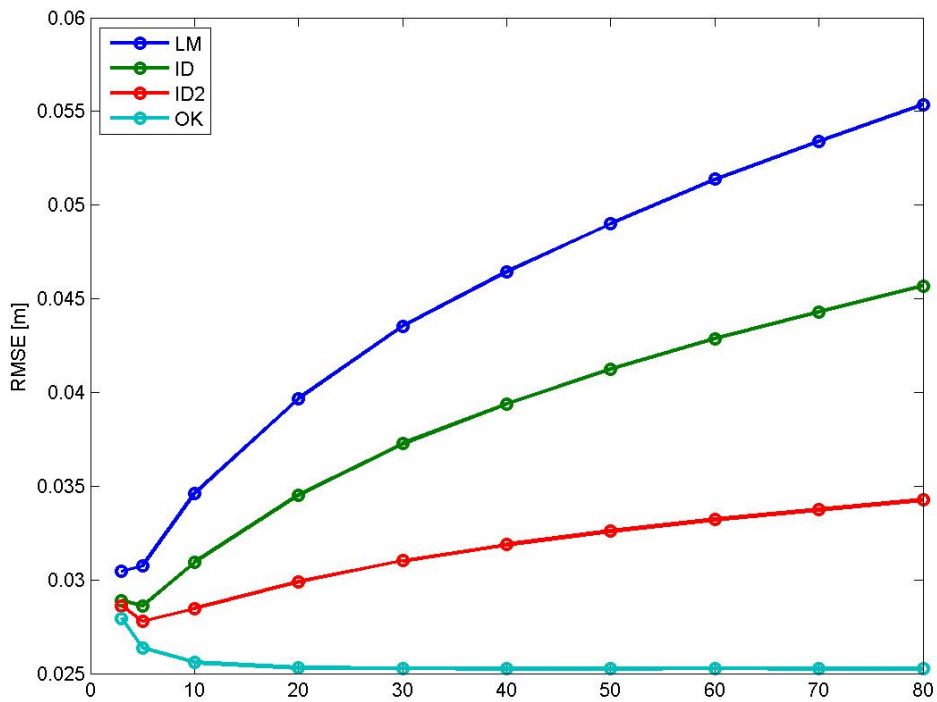
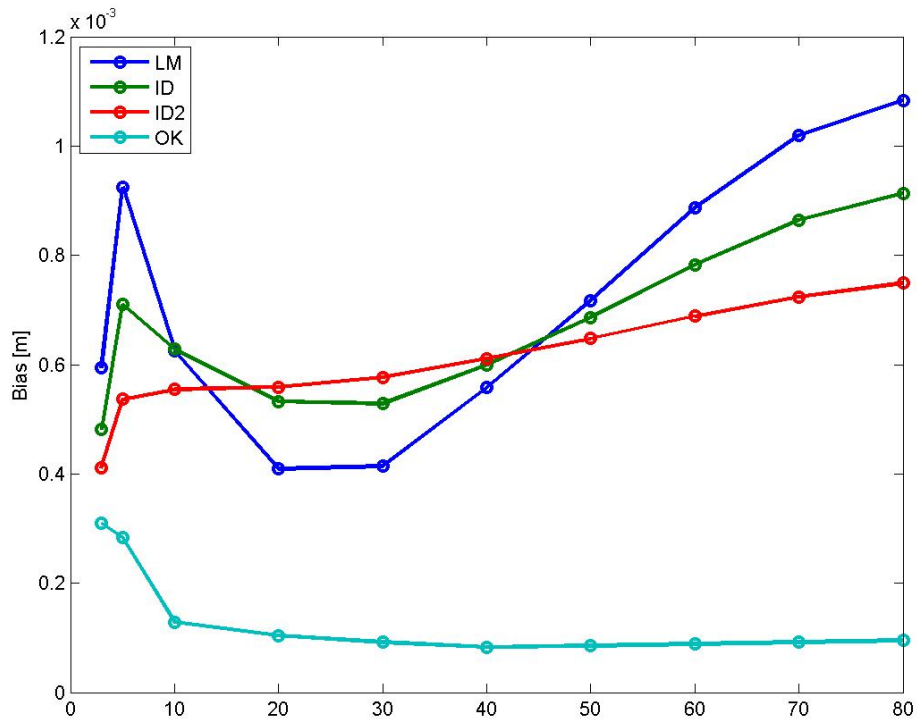


A spherical (linear) semivariogram model is used.

OKSa_2013

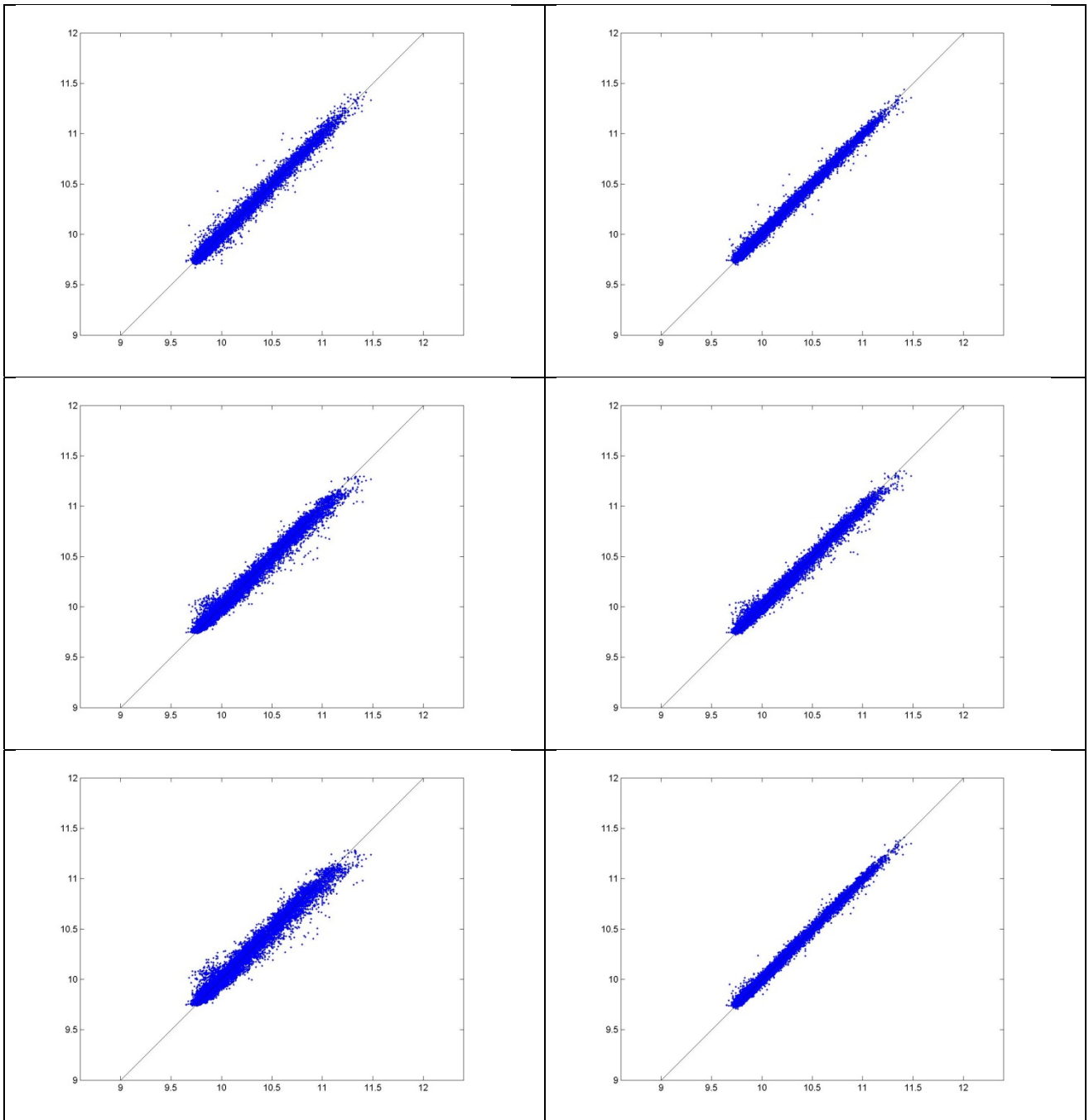


OKSa_2013

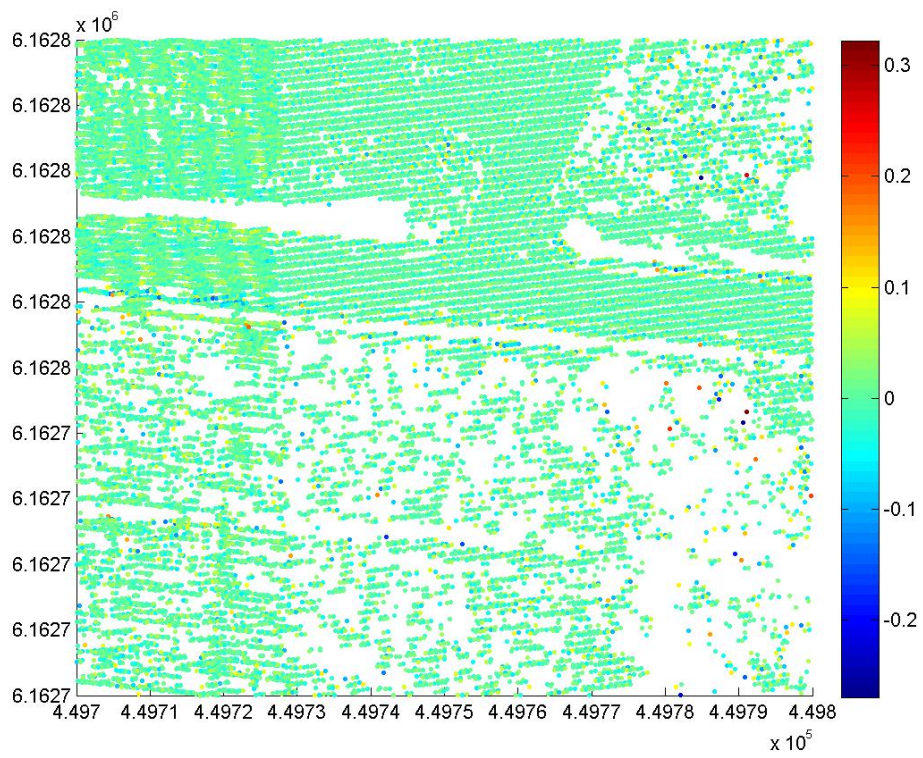
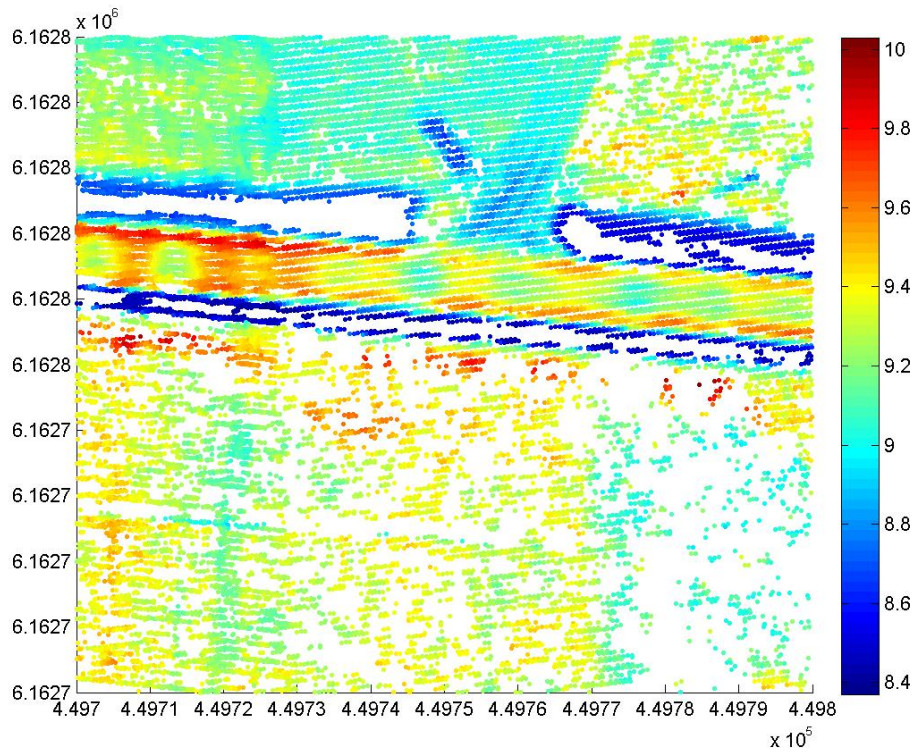


Bias for NN is 0.0003076 m, and for TIN -0.0004442 m. RMSE for NN is 0.03398 m, and for TIN 0.026897 m.

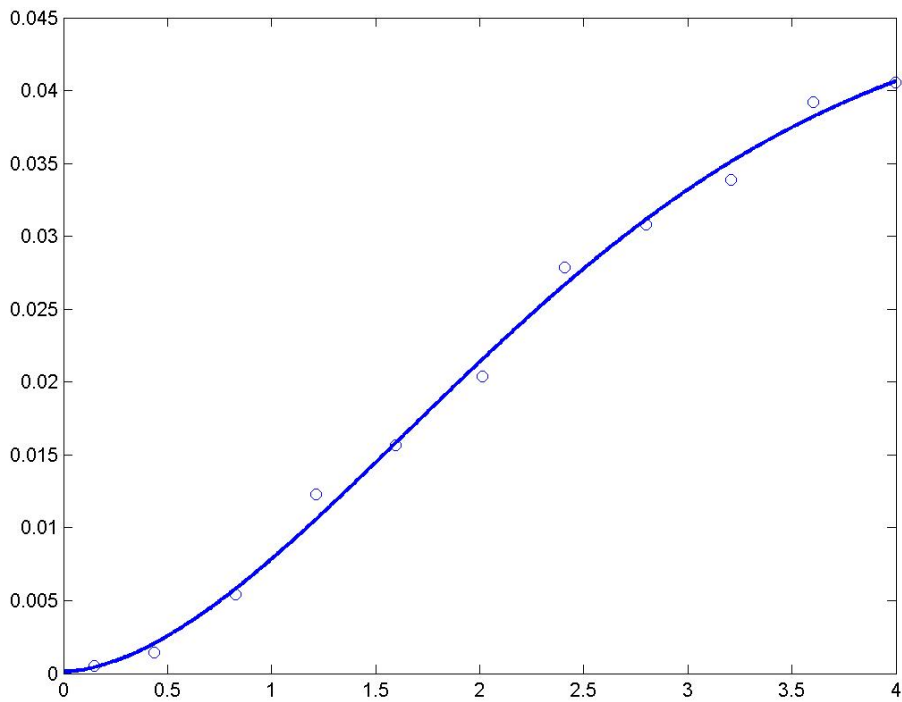
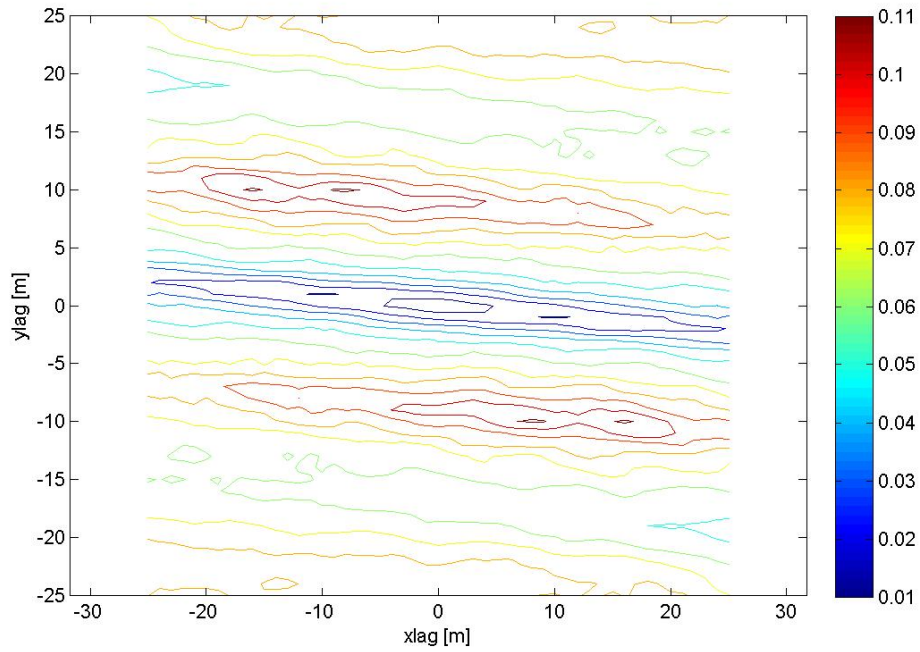
OKSa_2013



OKSb_2013

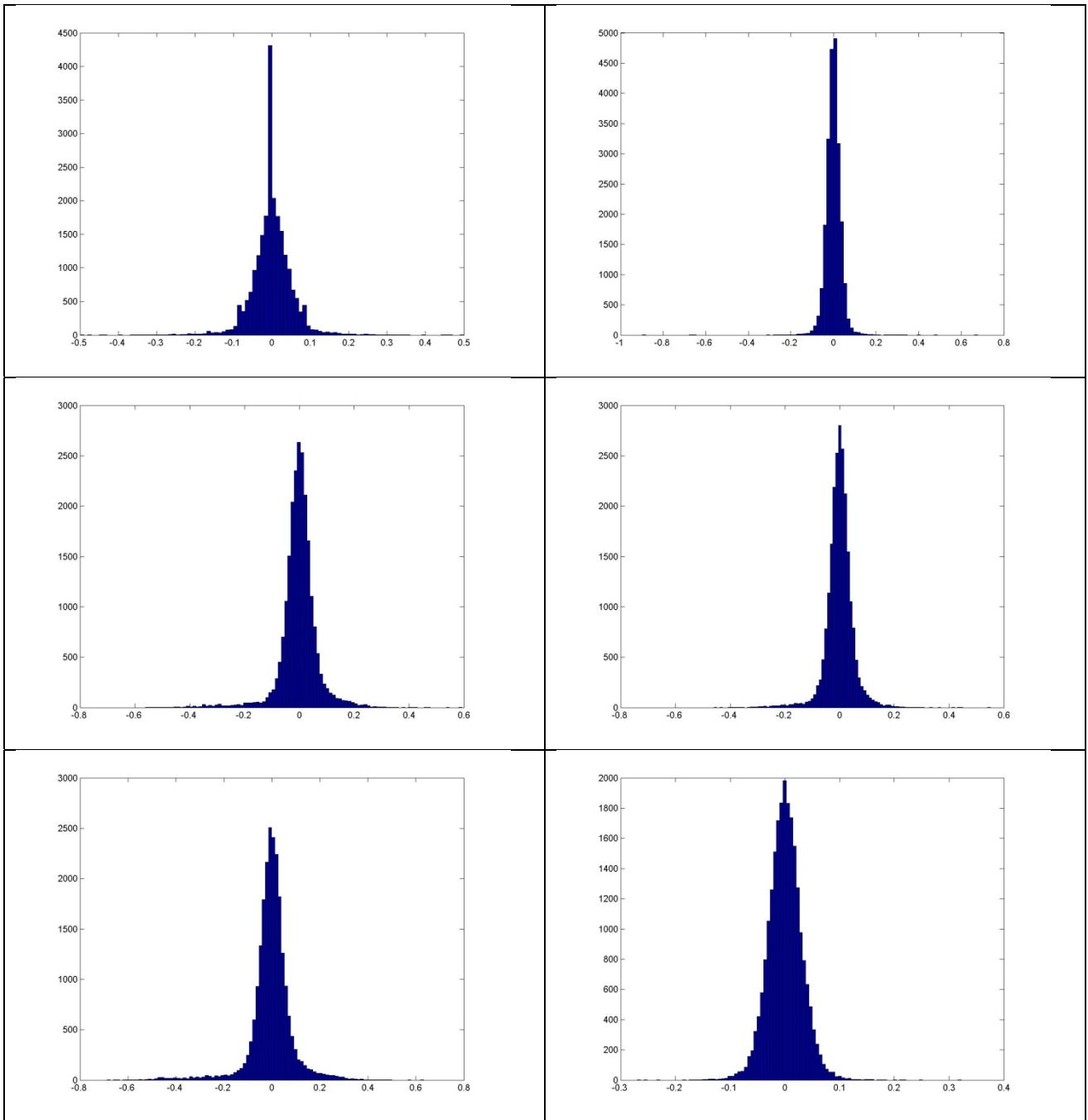


OKSb_2013

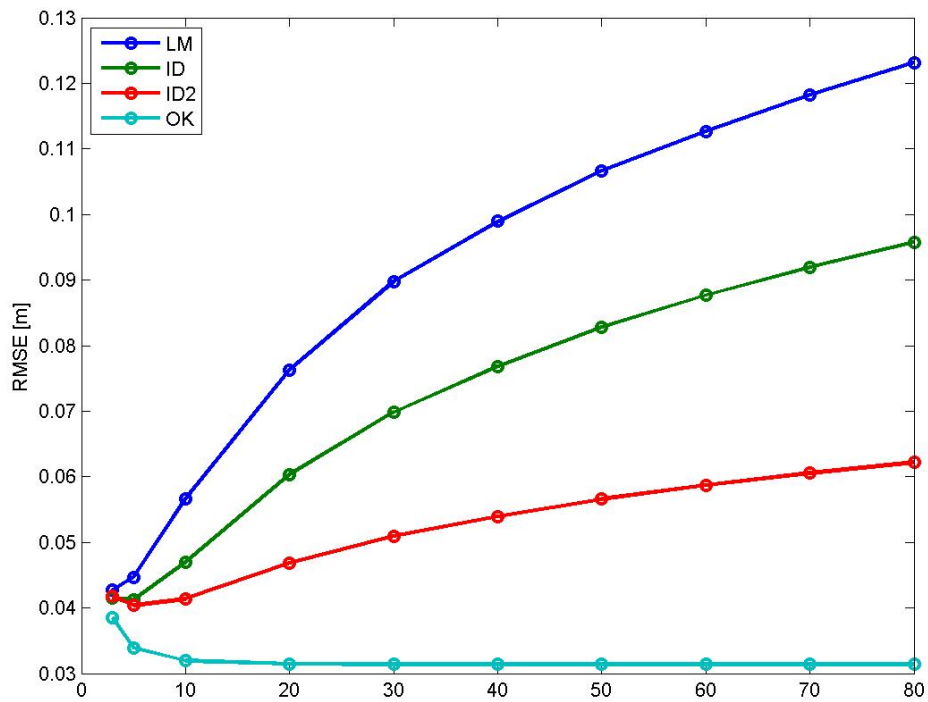
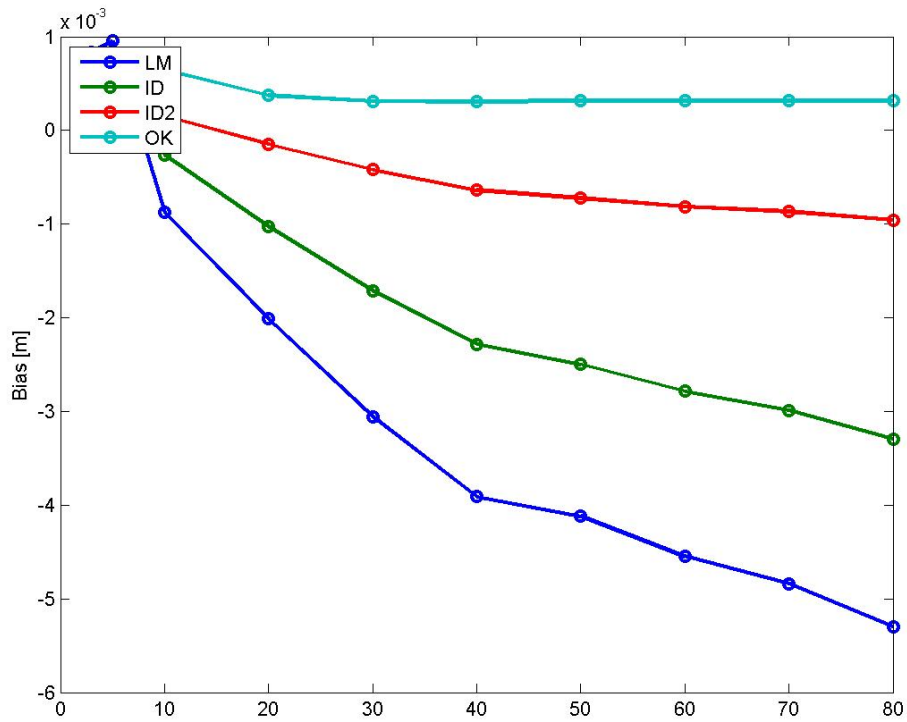


A modified Gaussian semivariogram model is used (power is 1.75 and not 2).

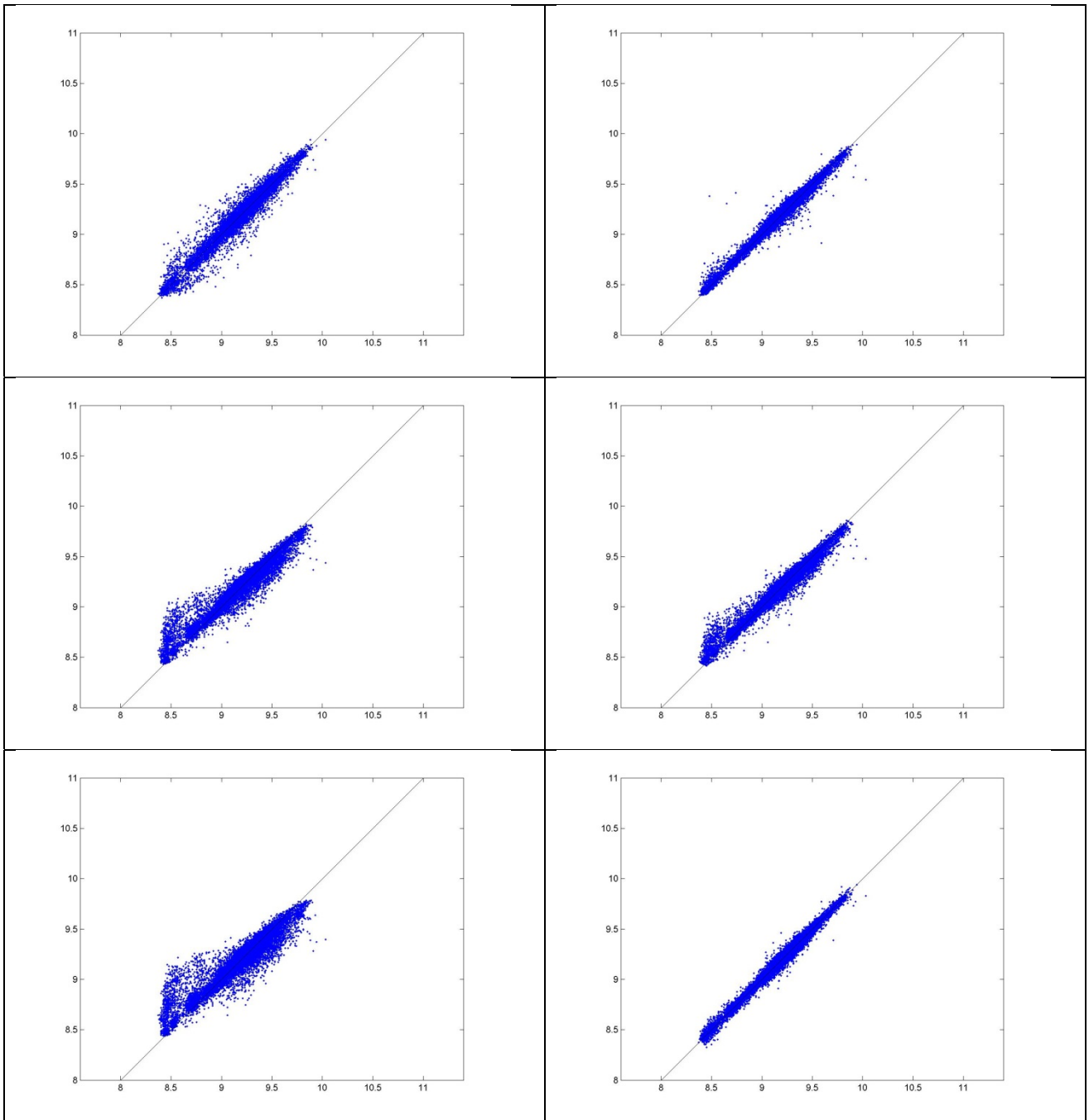
OKSb_2013



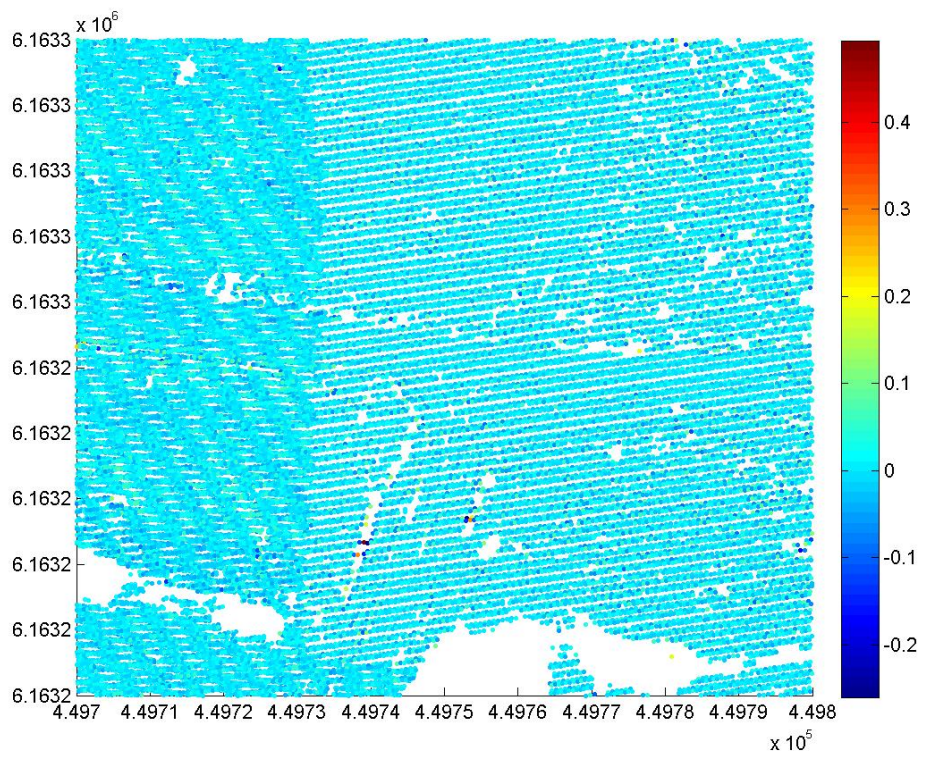
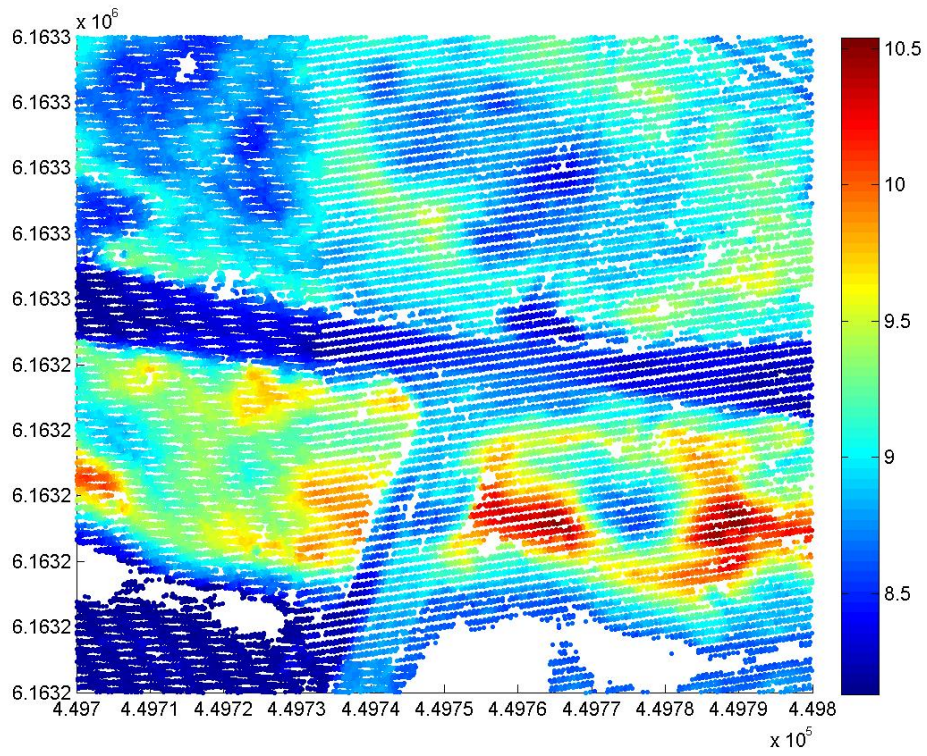
OKSb_2013



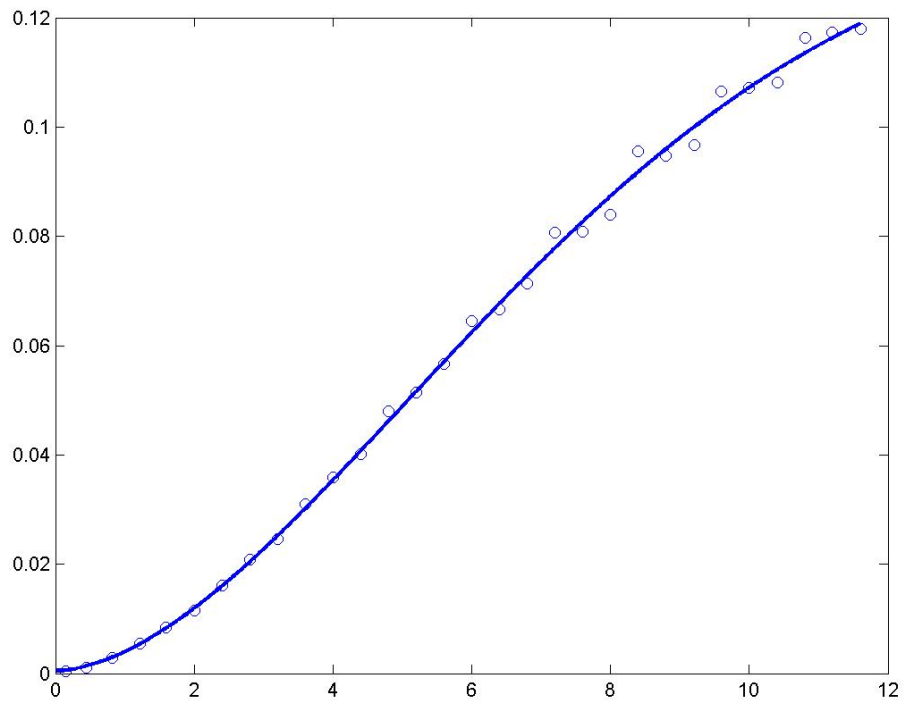
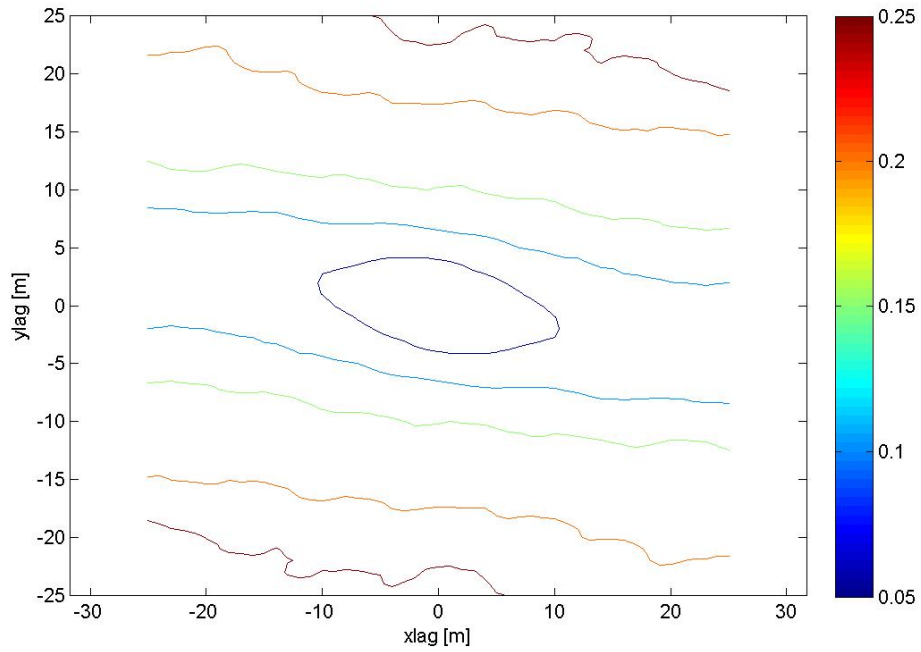
Bias for NN is 0.00029805 m, and for TIN -0.0003113 m. RMSE for NN is 0.05375 m, and for TIN 0.03583 m.



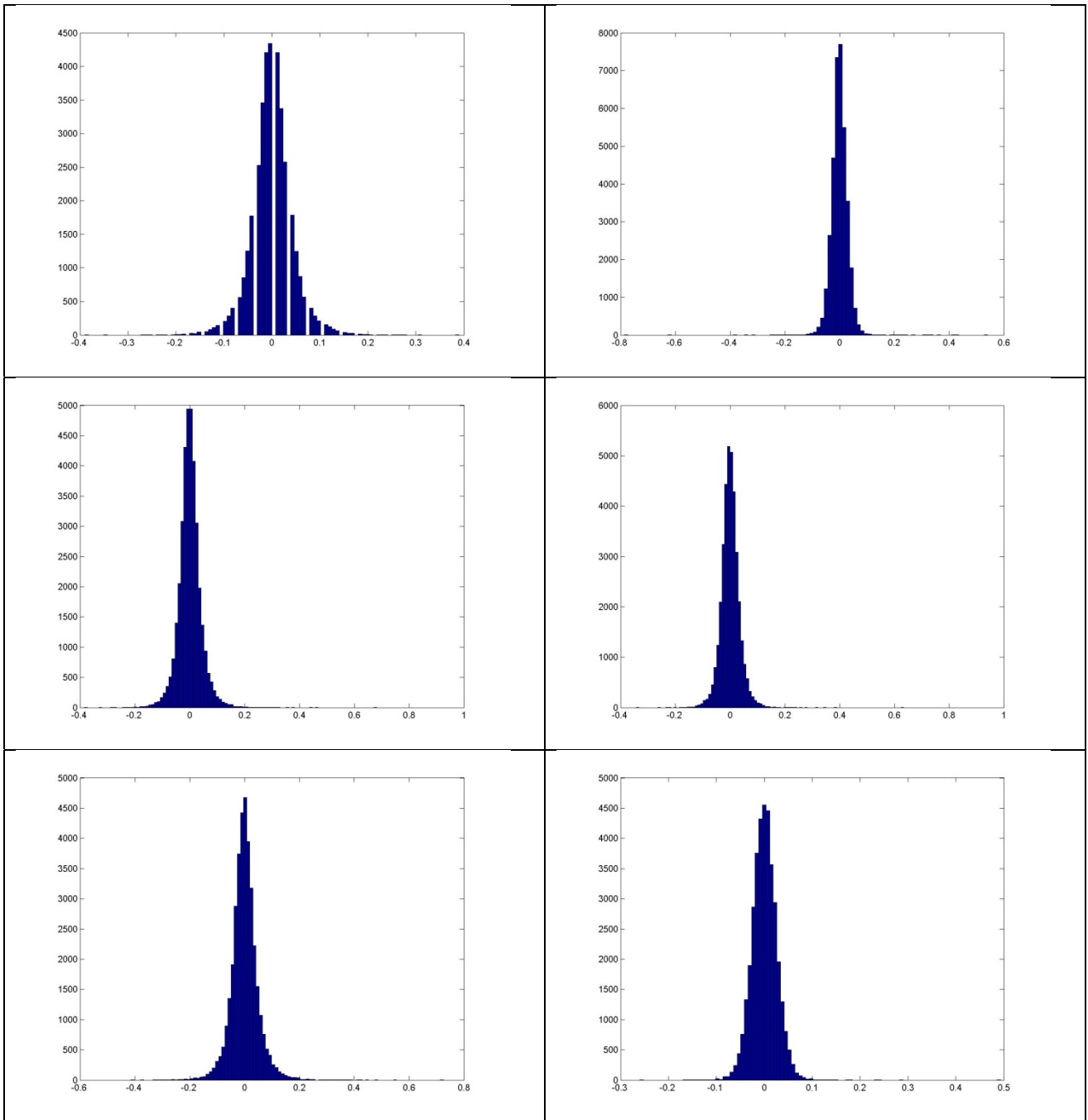
OKSc_2013



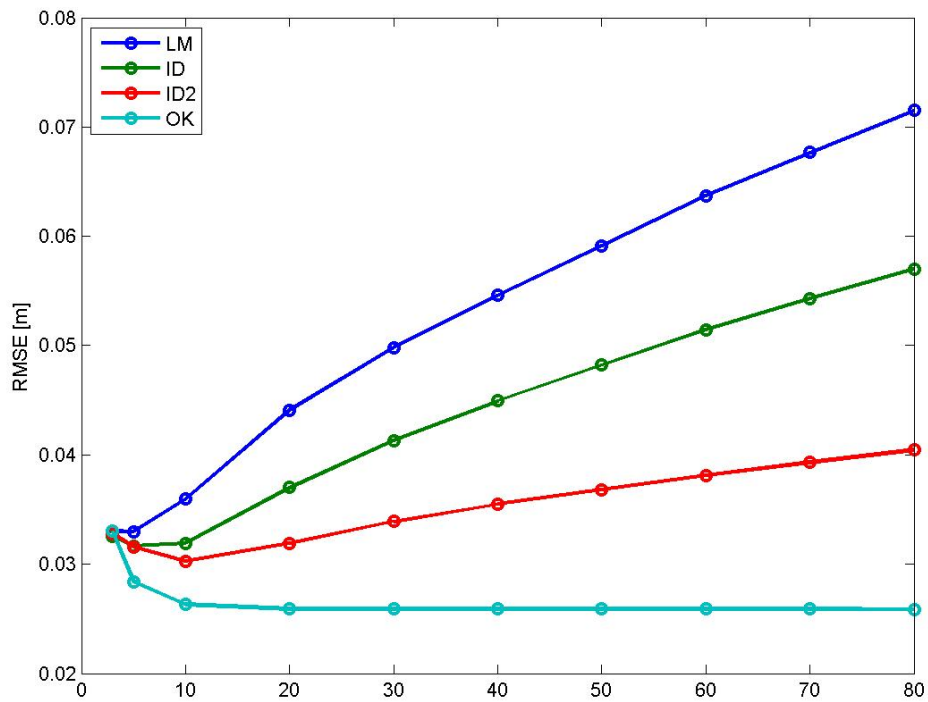
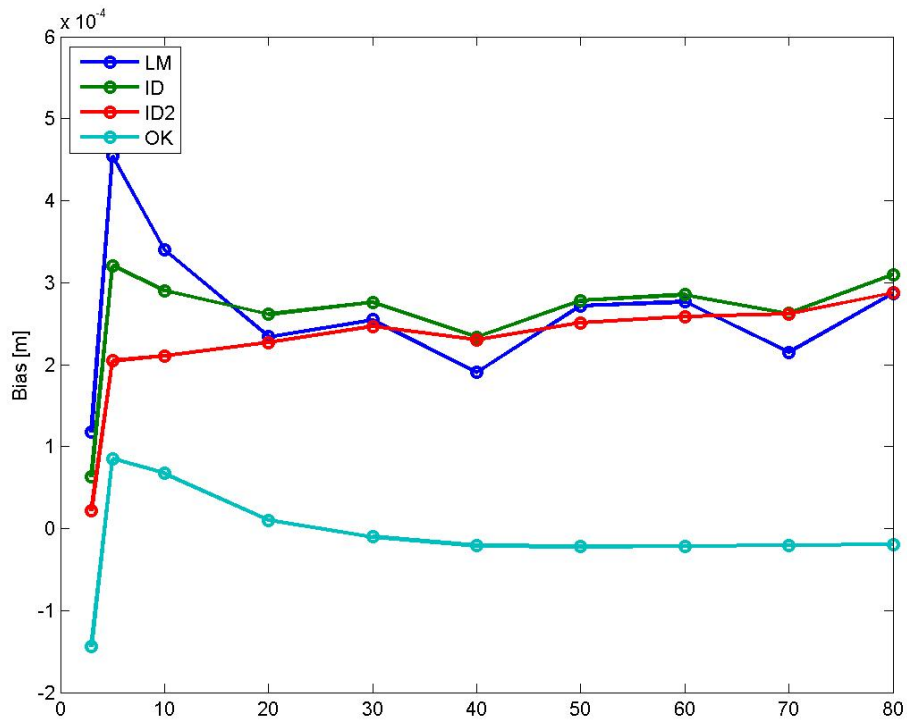
OKSc_2013



A modified Gaussian semivariogram model is used (power is 1.75 and not 2).

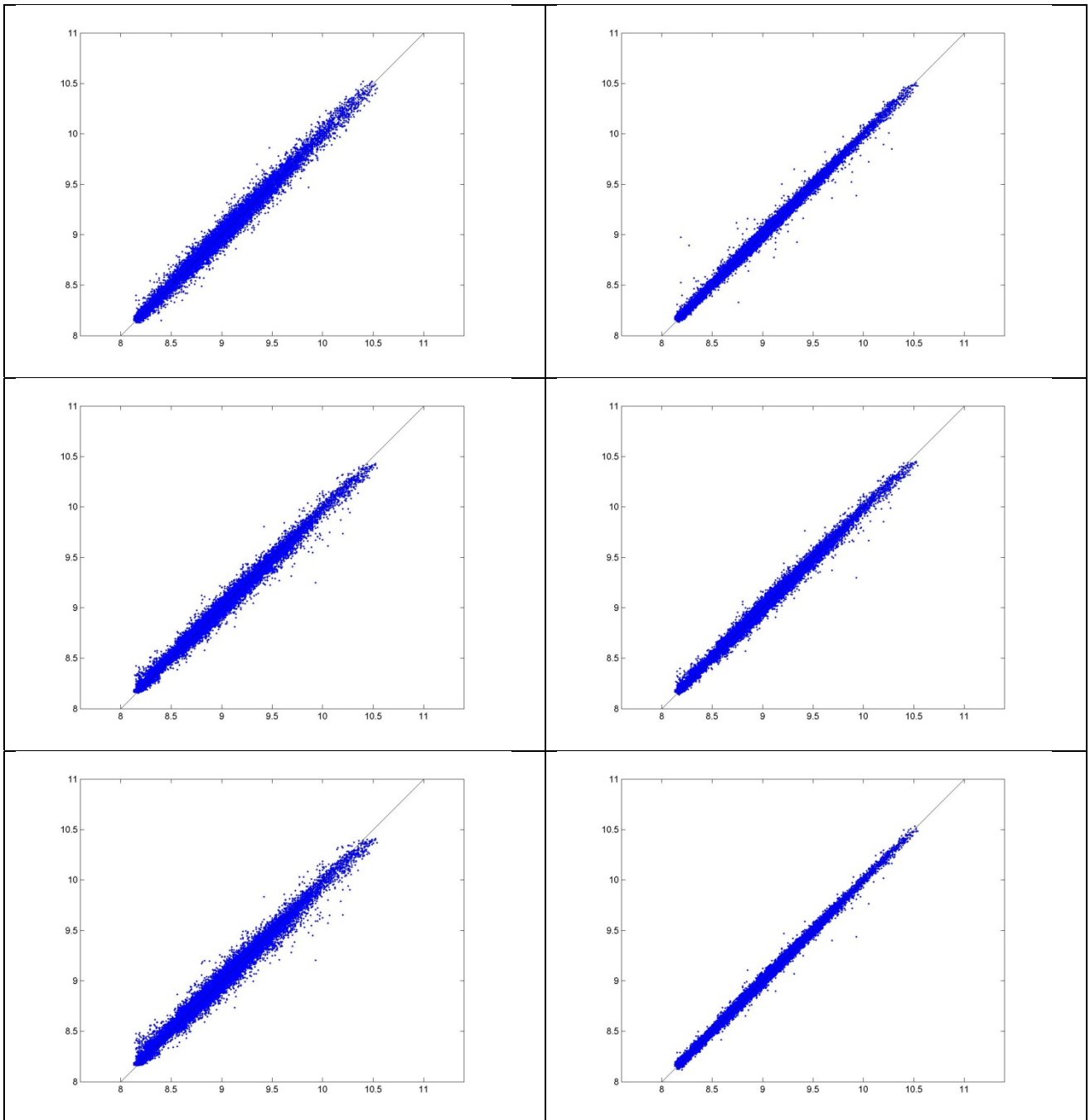


OKSc_2013

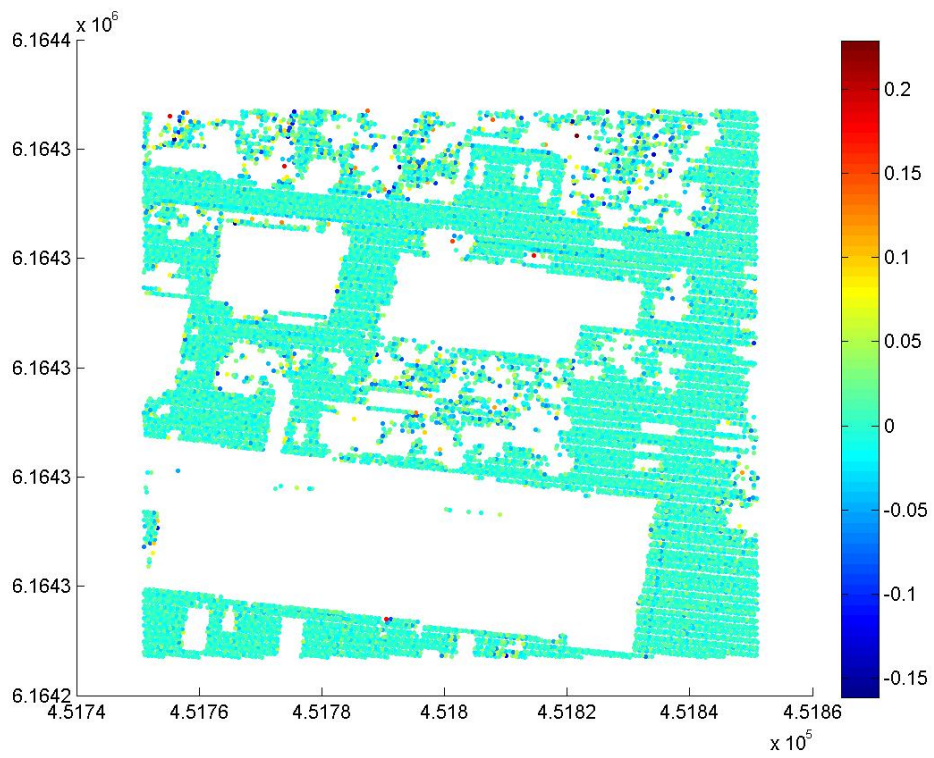
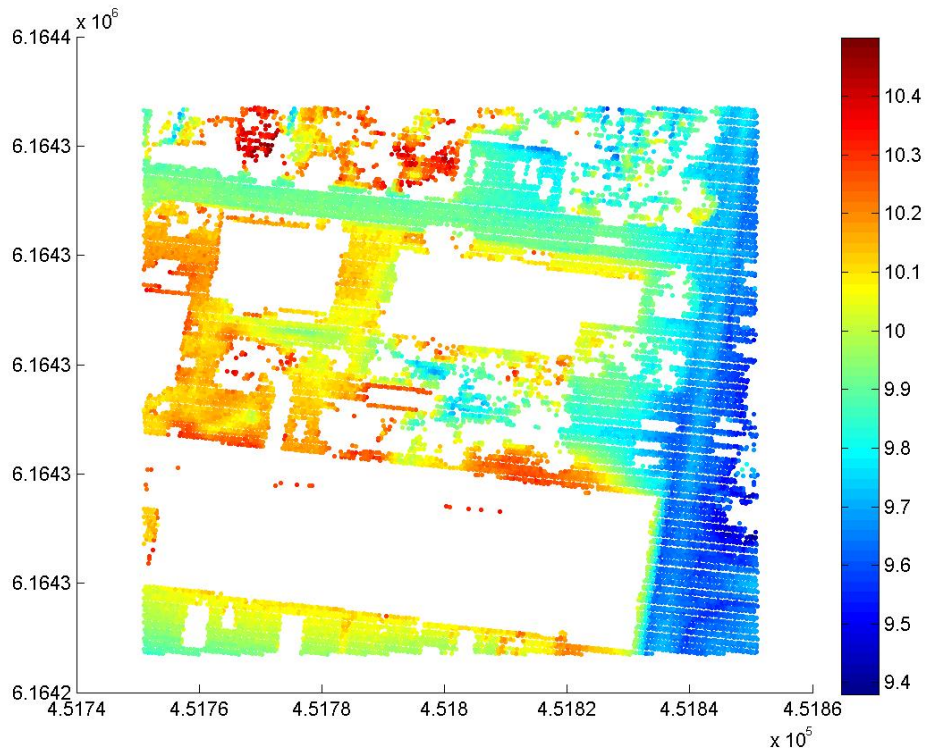


Bias for NN is 0.000143 m, and for TIN -0.0003364 m. RMSE for NN is 0.043426 m, and for TIN 0.029595 m.

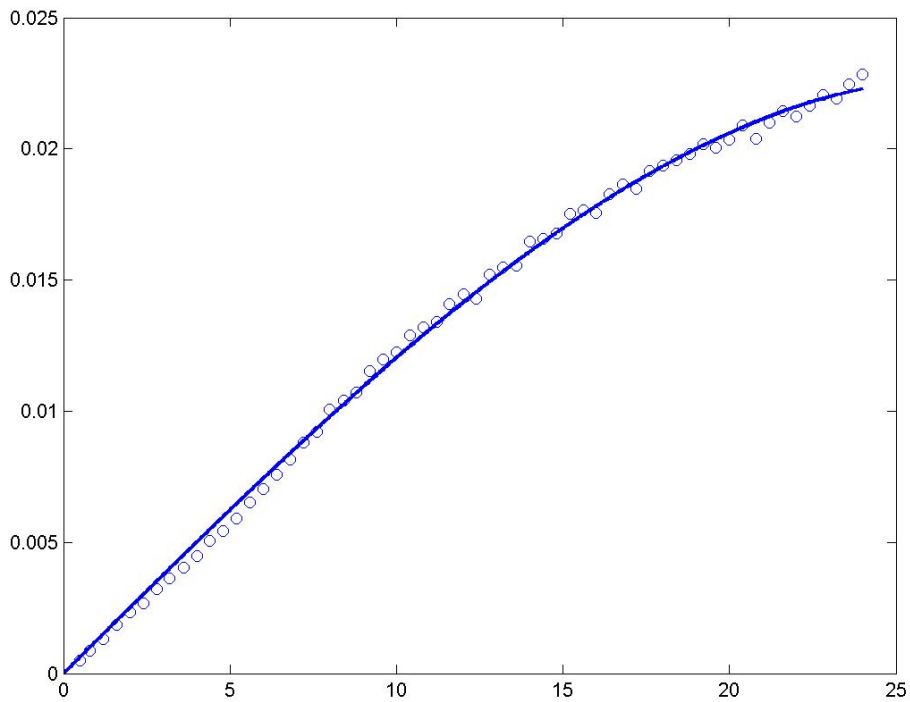
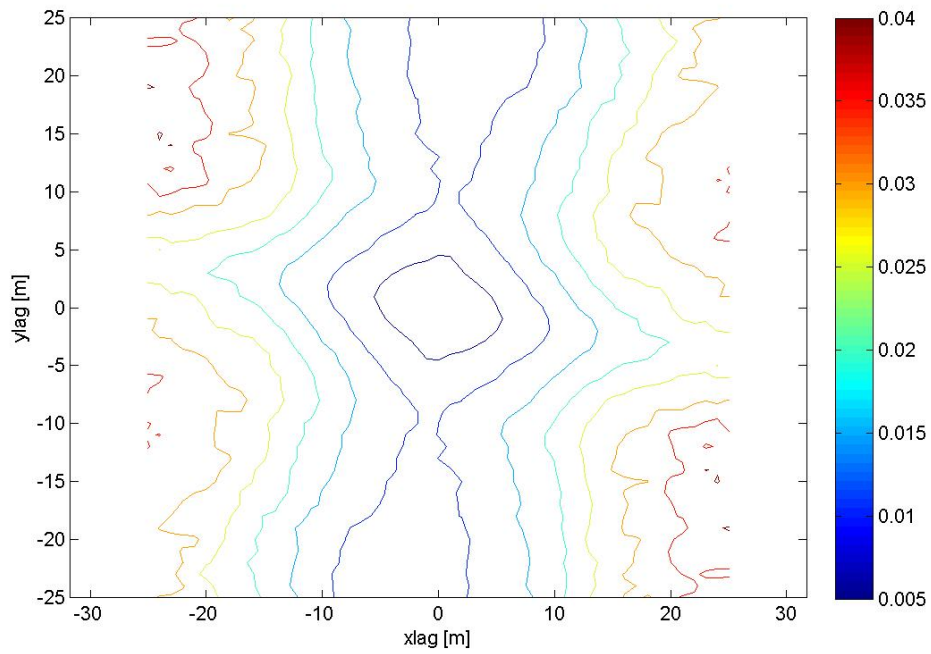
OKSc_2013



OKSd_2013

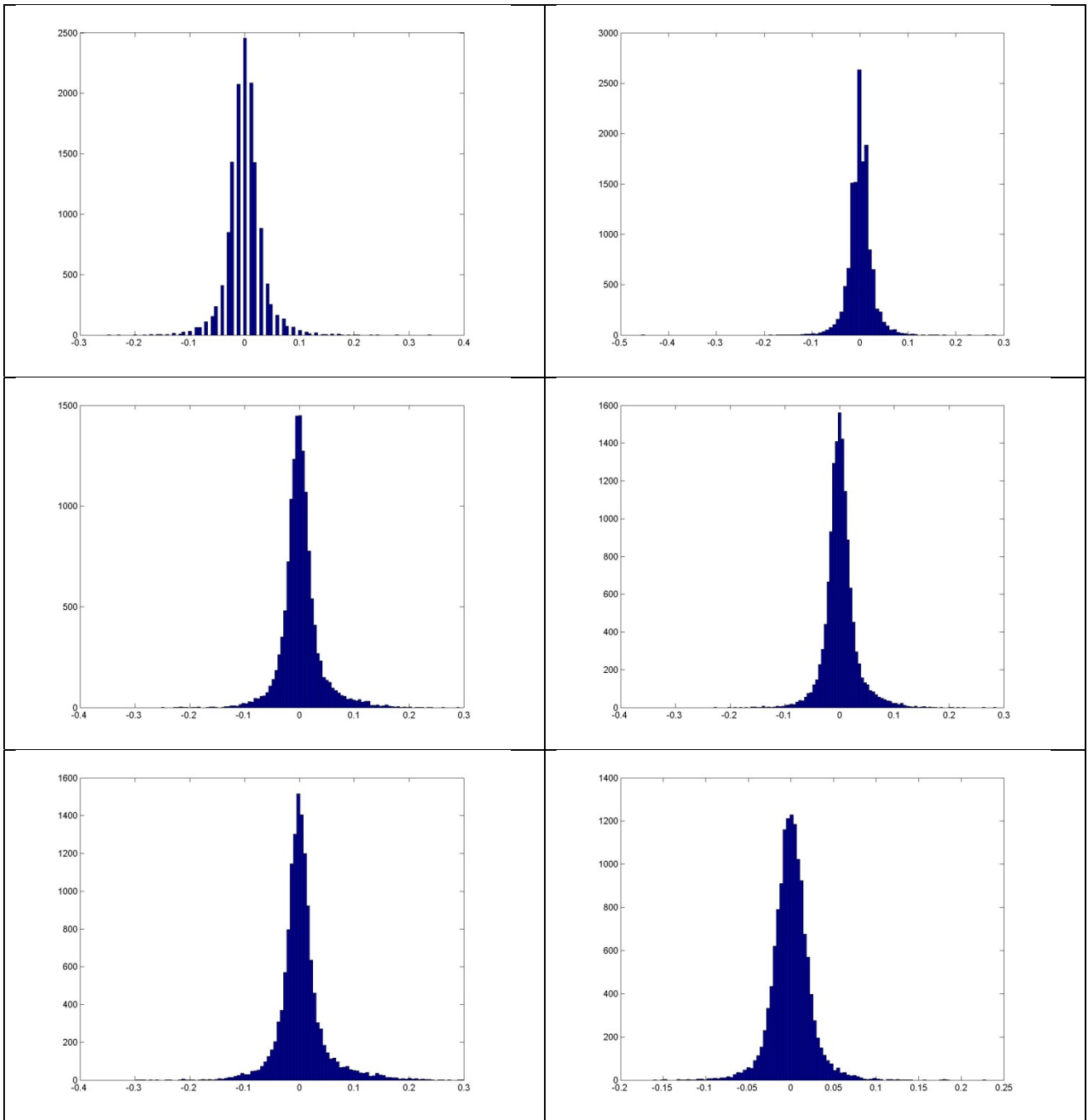


OKSd_2013

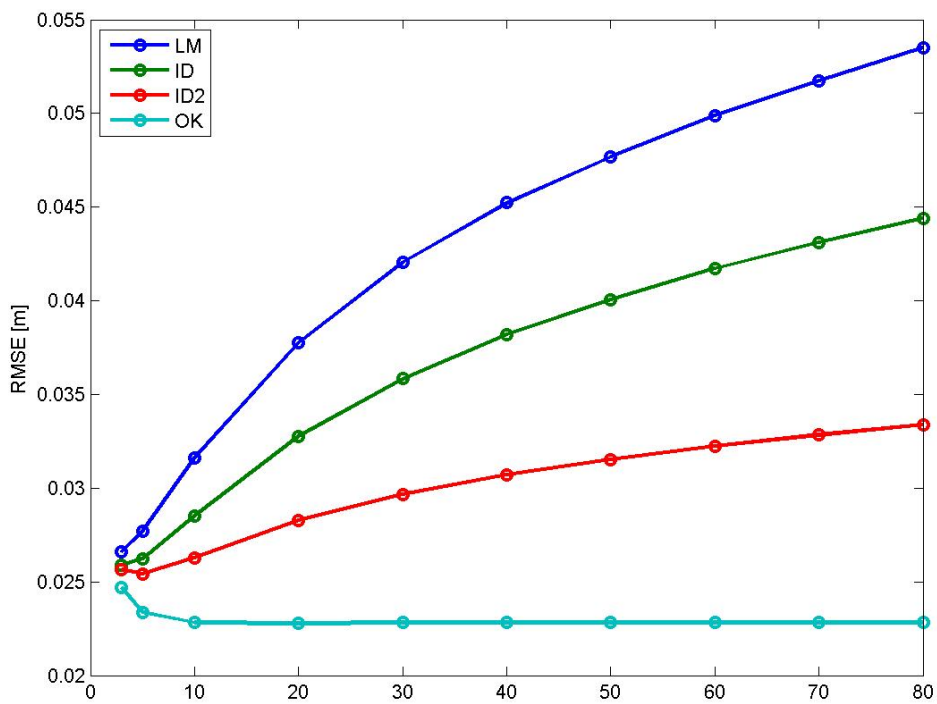
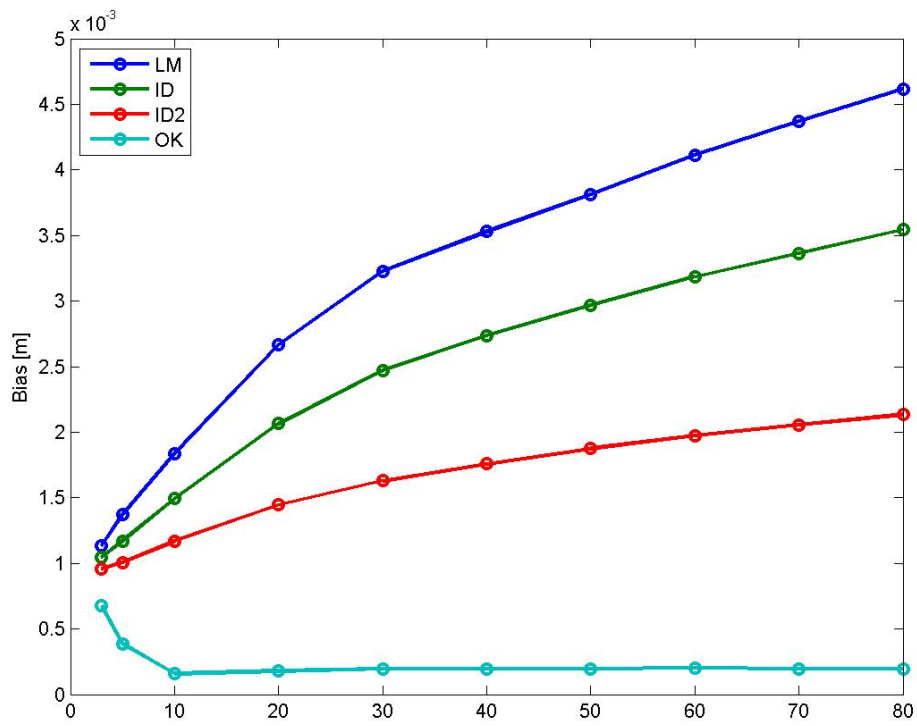


A spherical semivariogram model is used, no nugget effect.

OKSd_2013

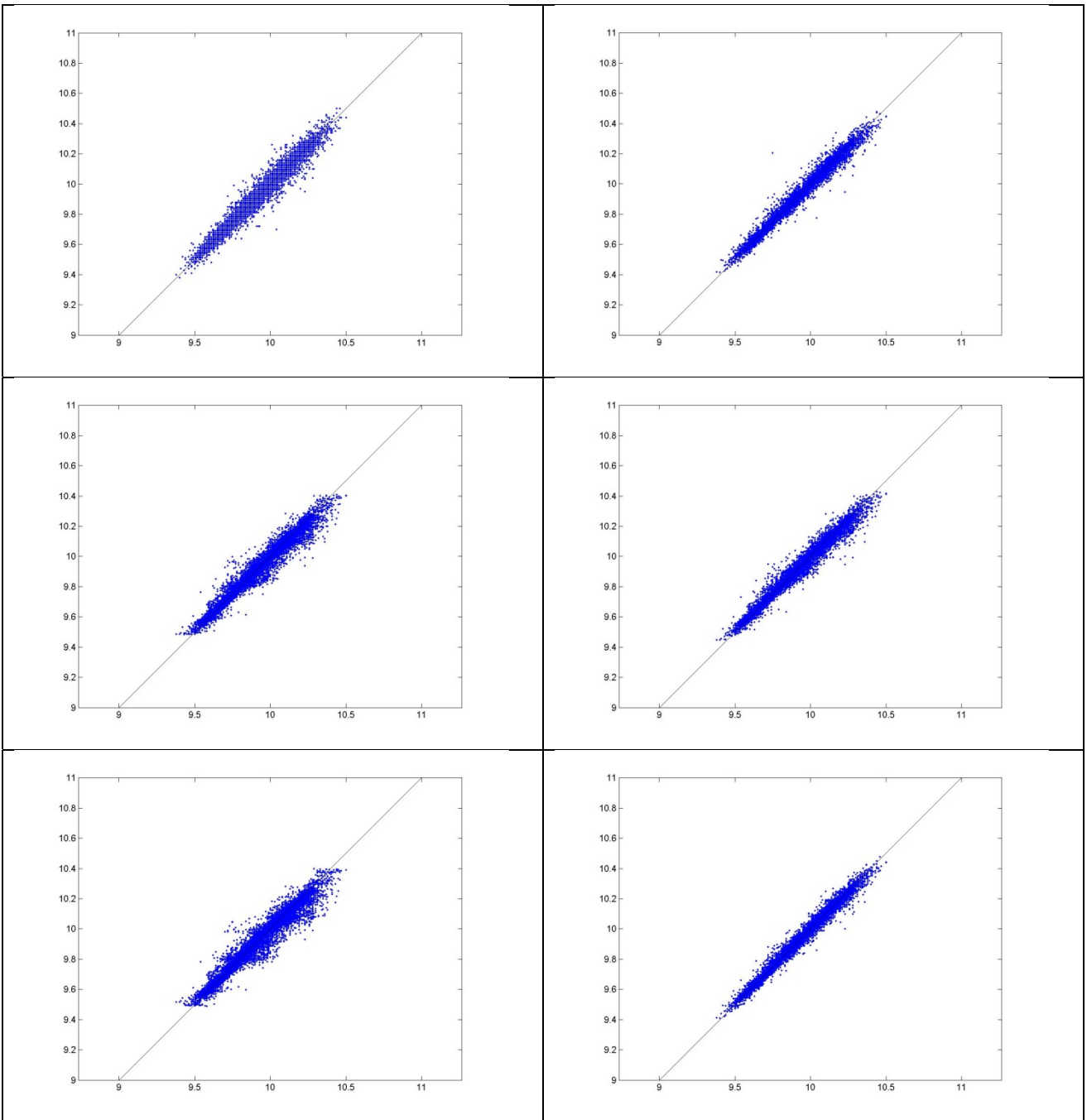


OKSd_2013

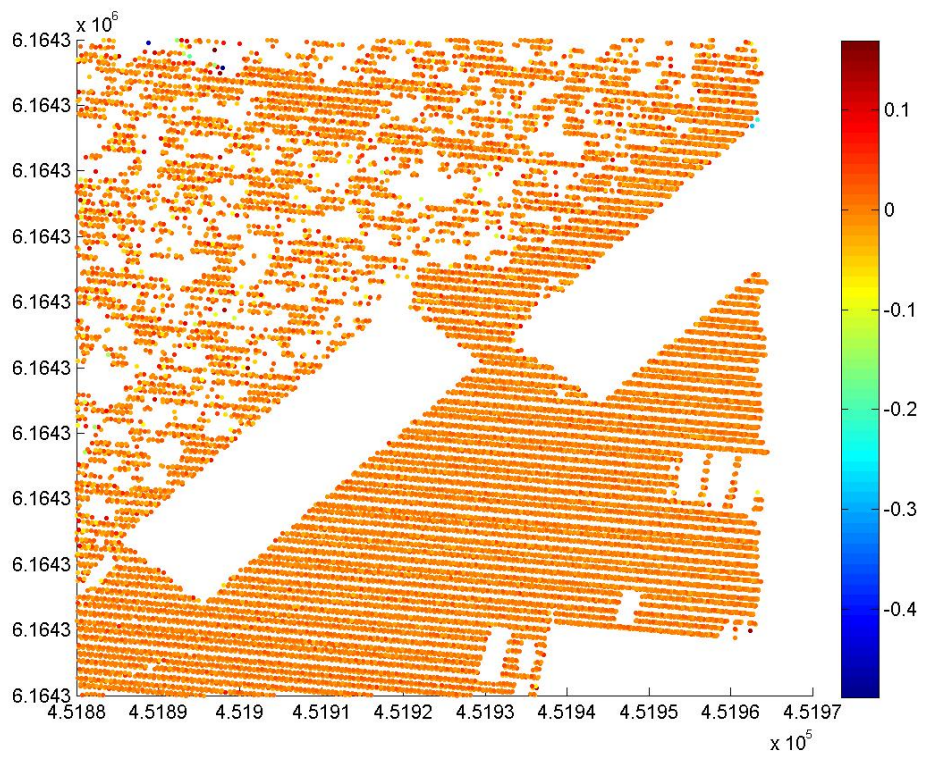
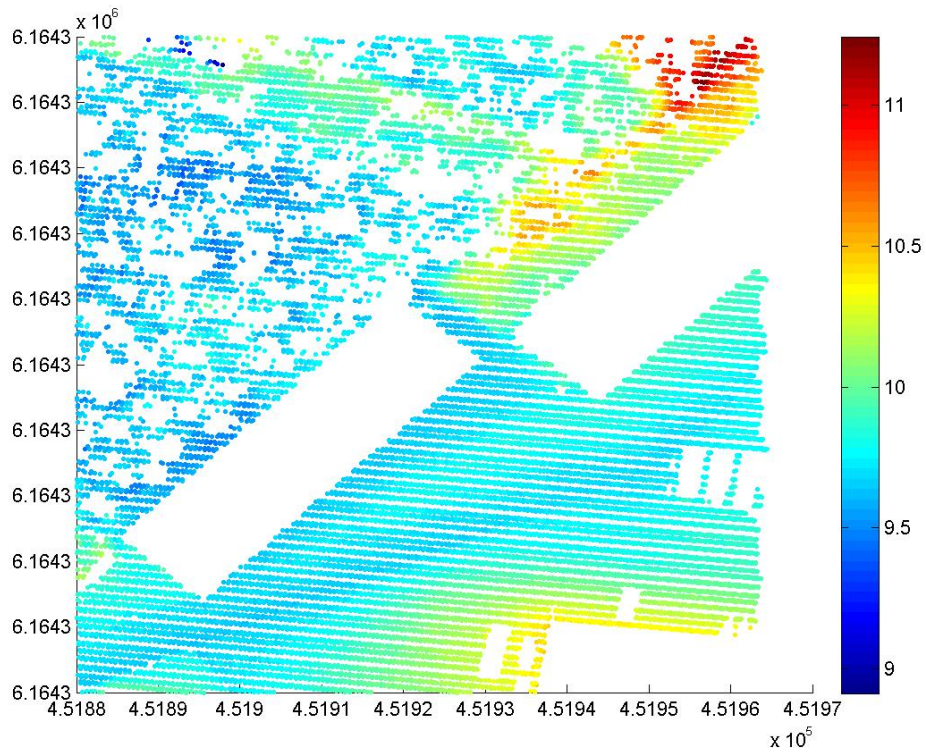


Bias for NN is 0.0007297 m, and for TIN -0.0002848 m. RMSE for NN is 0.03236 m, and for TIN 0.025345 m.

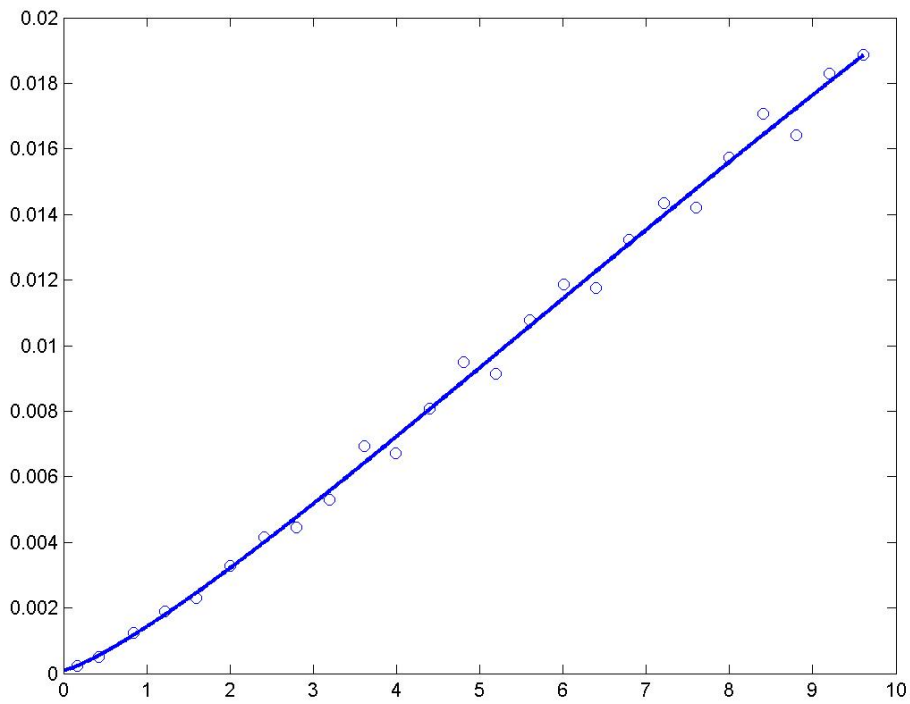
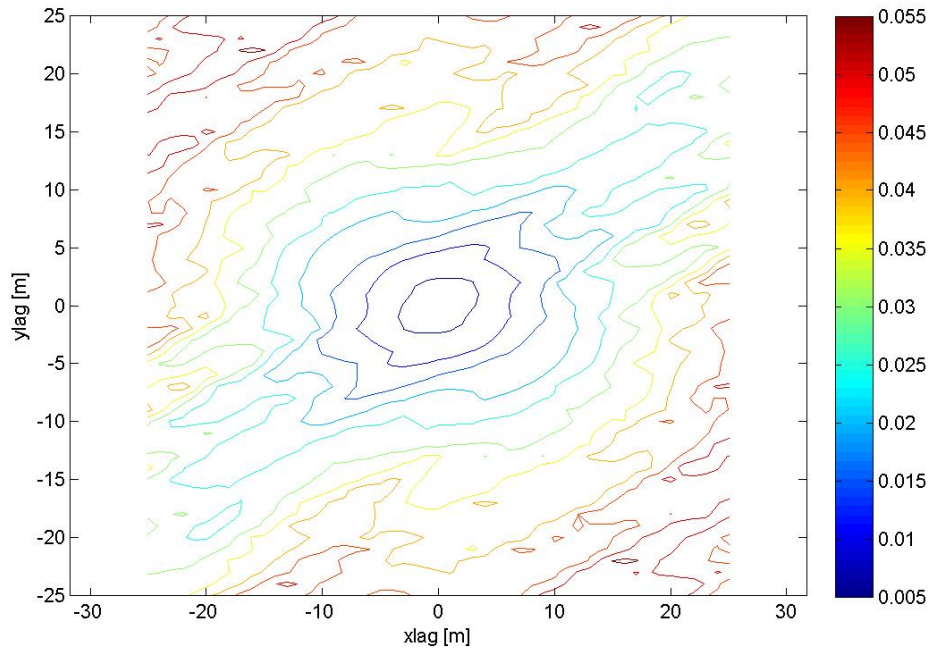
OKSd_2013



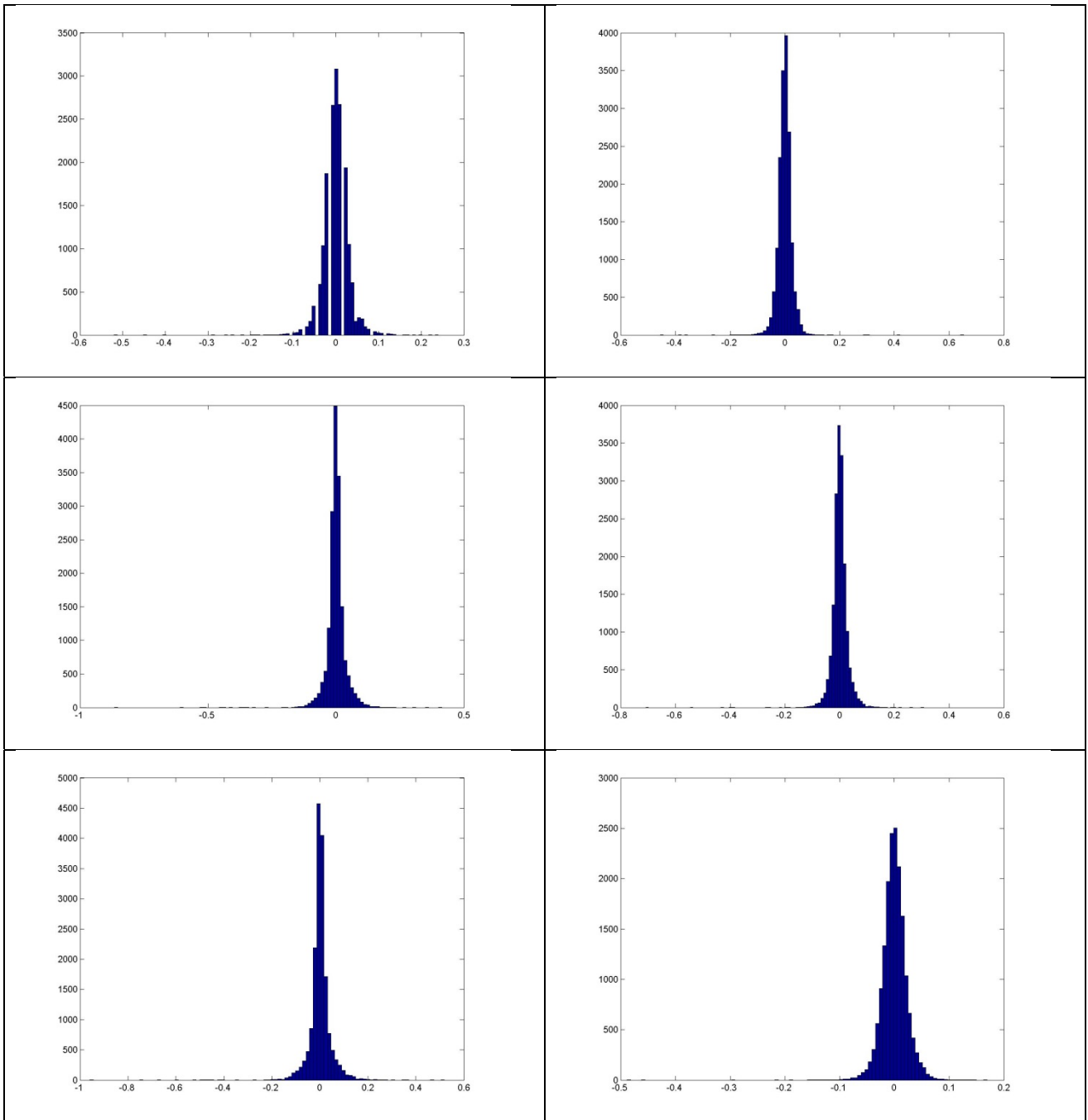
OKSe_2013



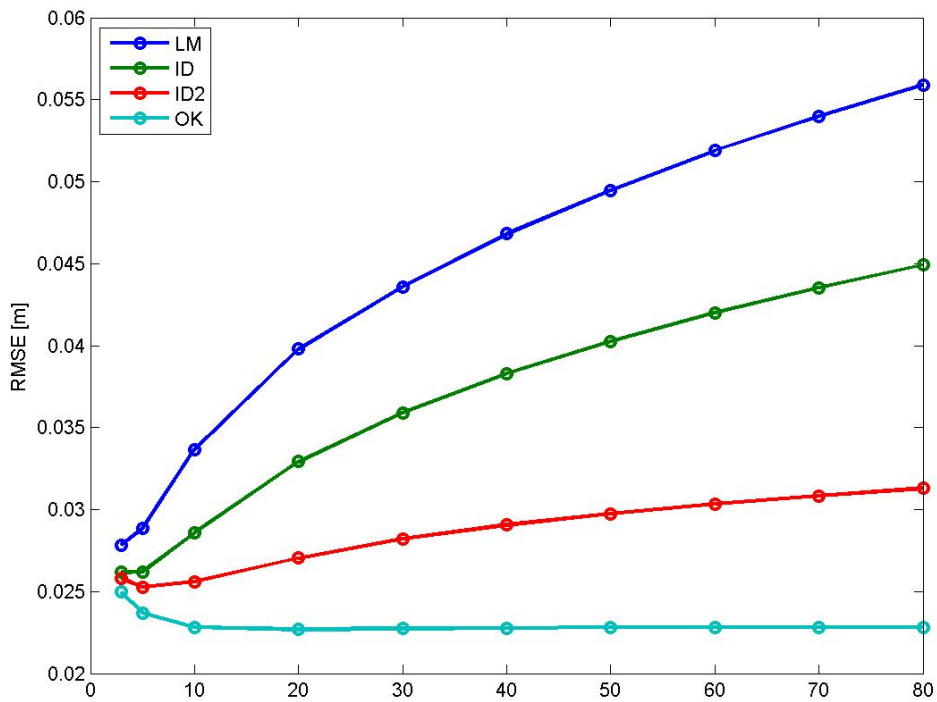
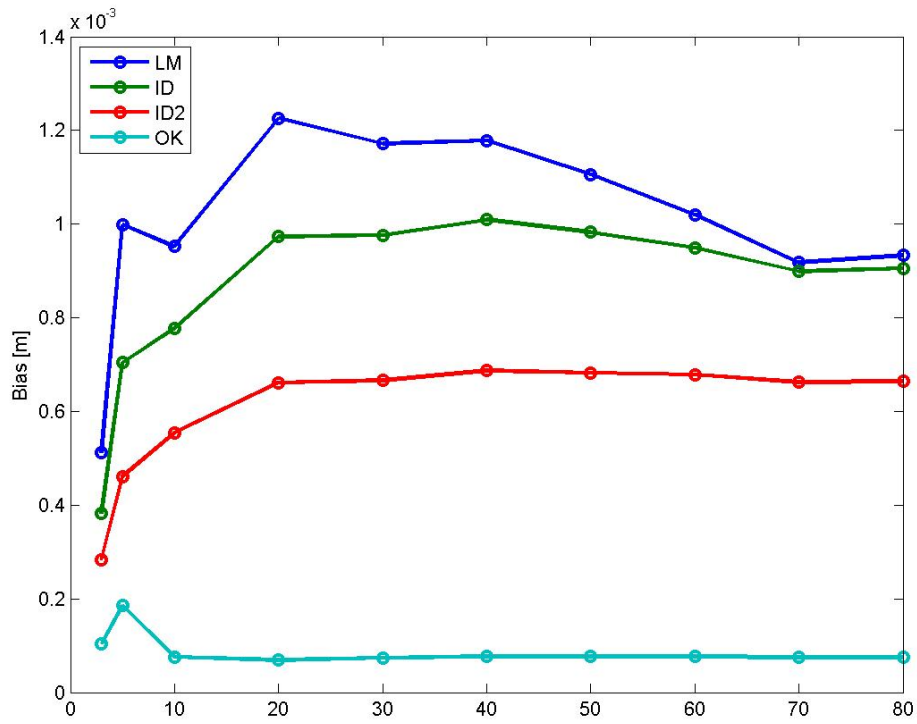
OKSe_2013



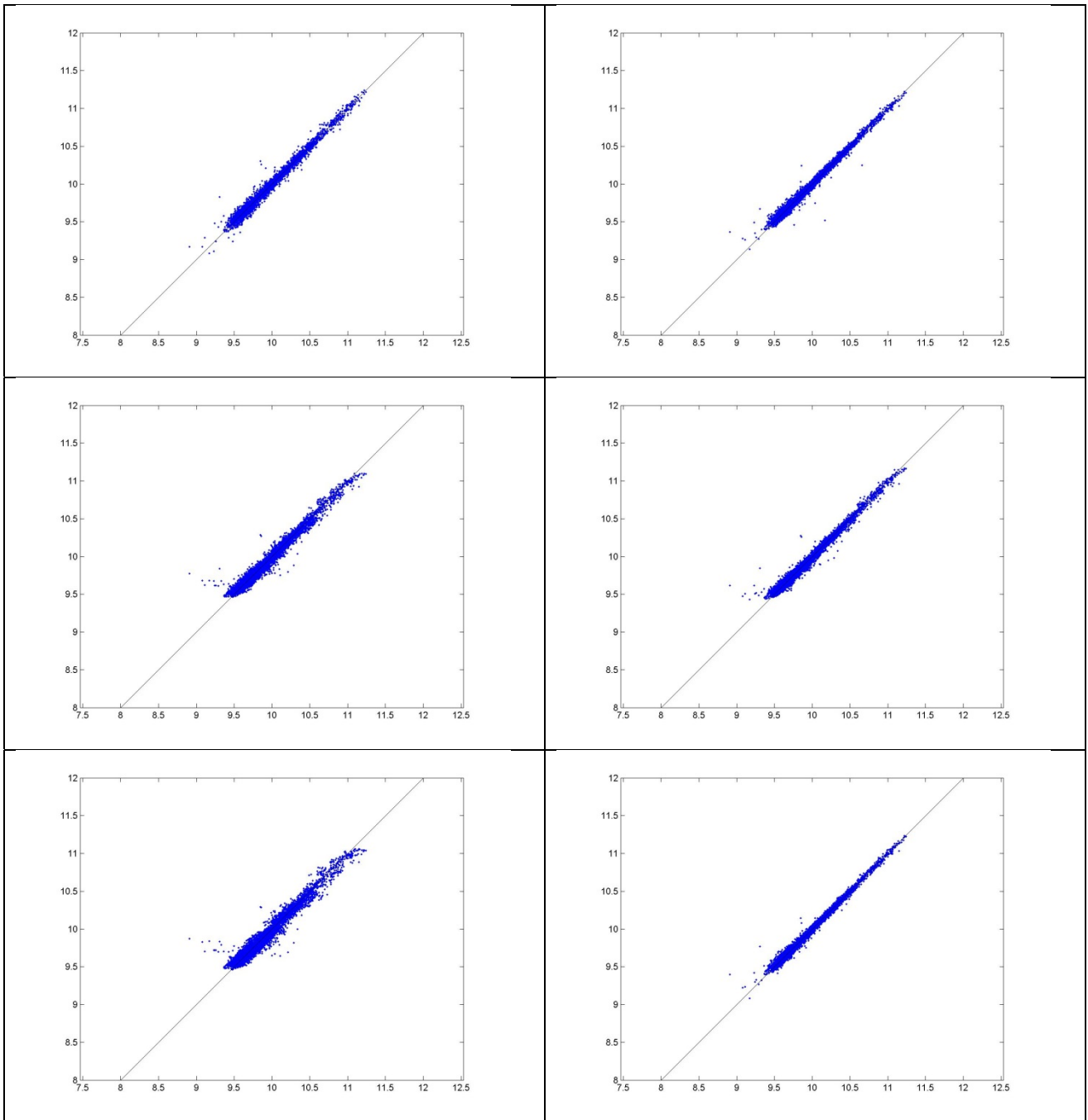
A modified Gaussian semivariogram model is used (power is 1.25 and not 2).



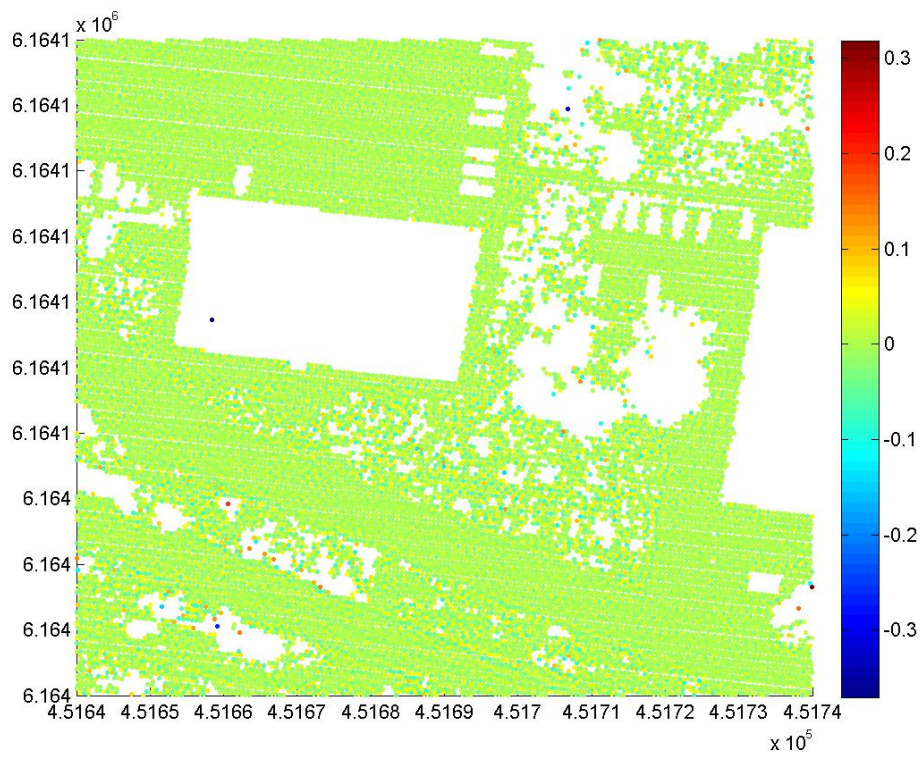
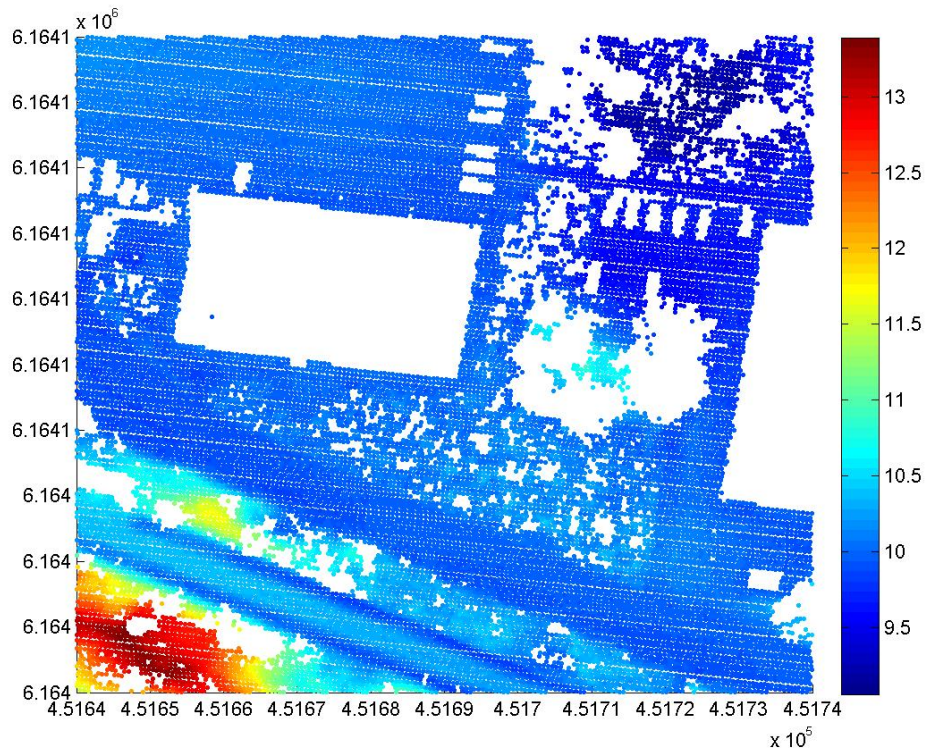
OKSe_2013



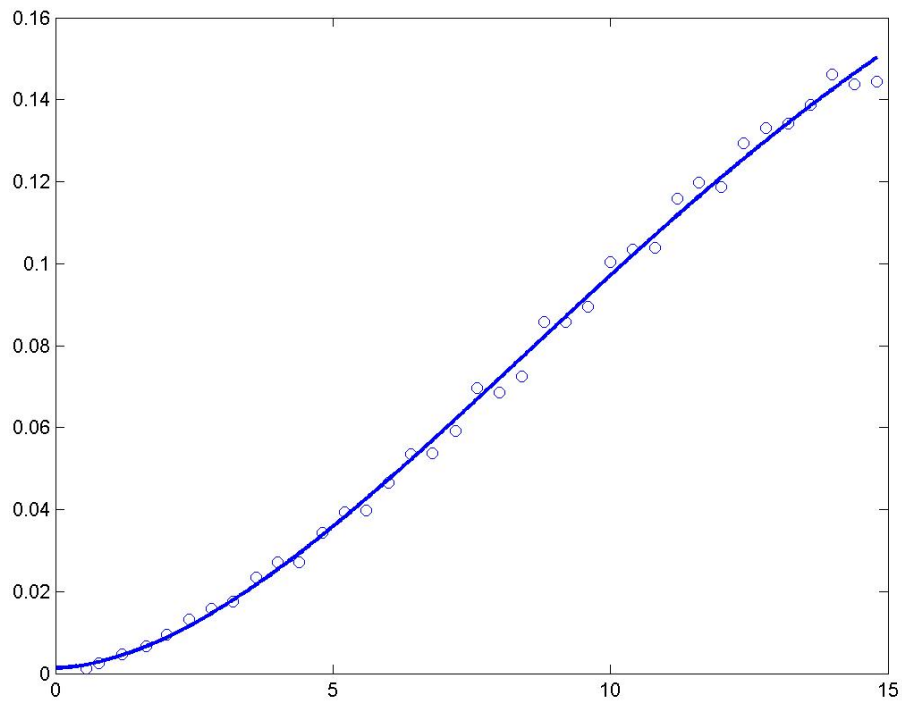
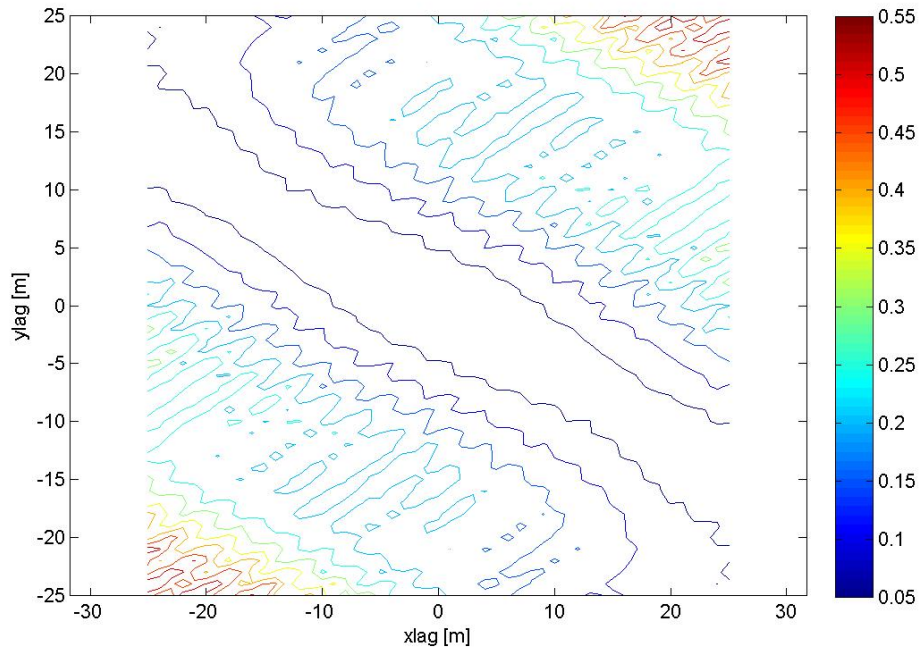
Bias for NN is 0.0002760 m, and for TIN -0.0005435 m. RMSE for NN is 0.02933 m, and for TIN 0.025123 m.



OKSf_2013

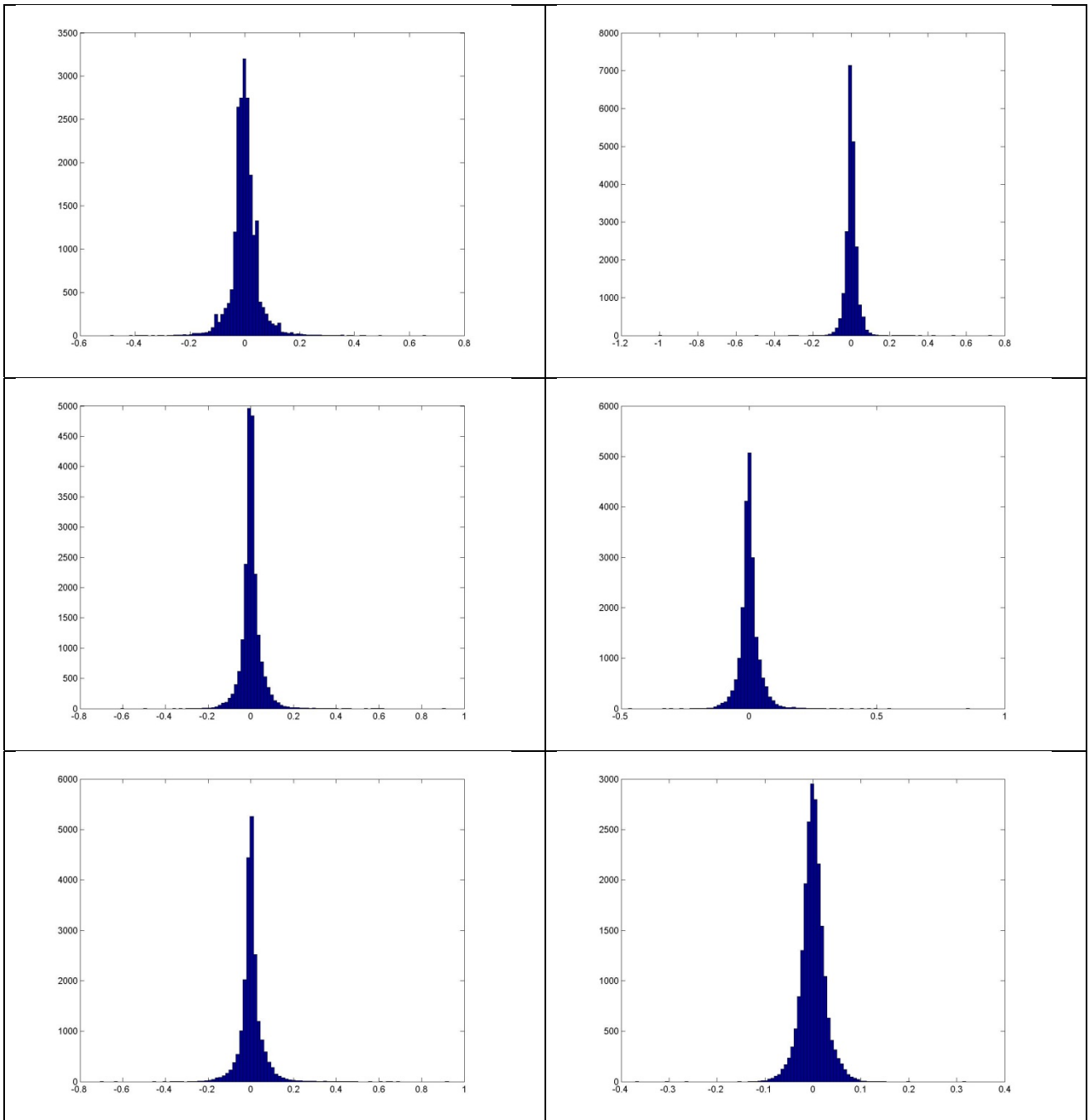


OKSf_2013

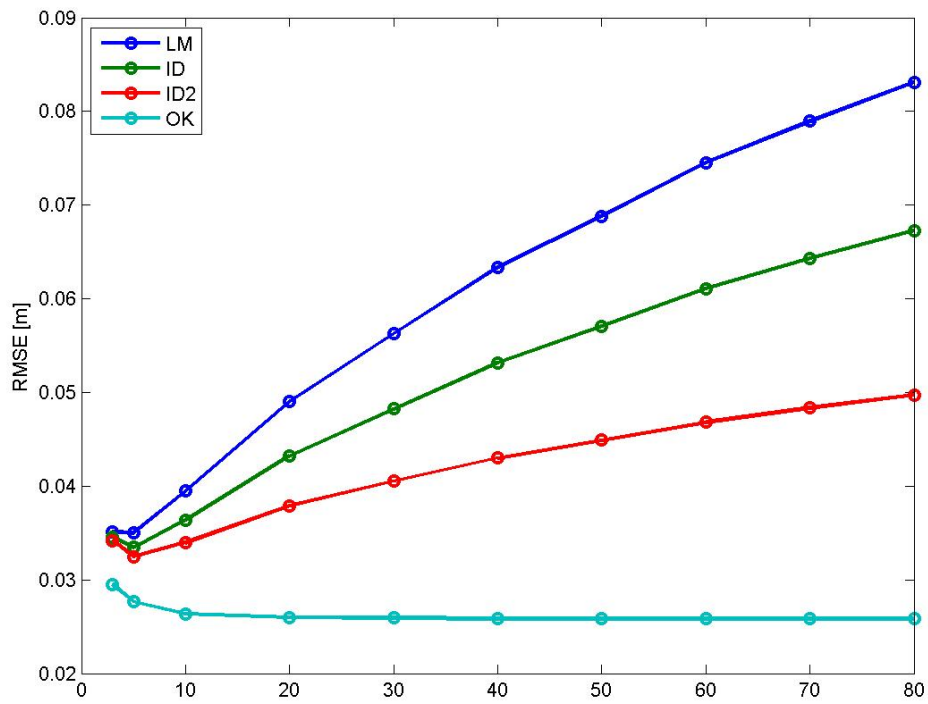
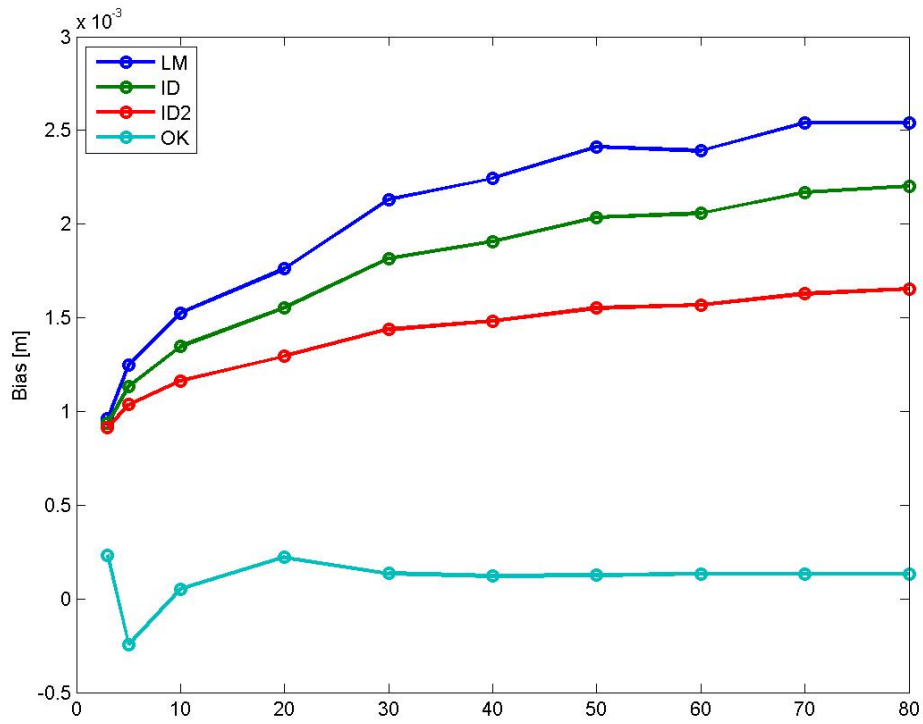


A modified Gaussian semivariogram model is used (power is 1.75 and not 2).

OKSf_2013

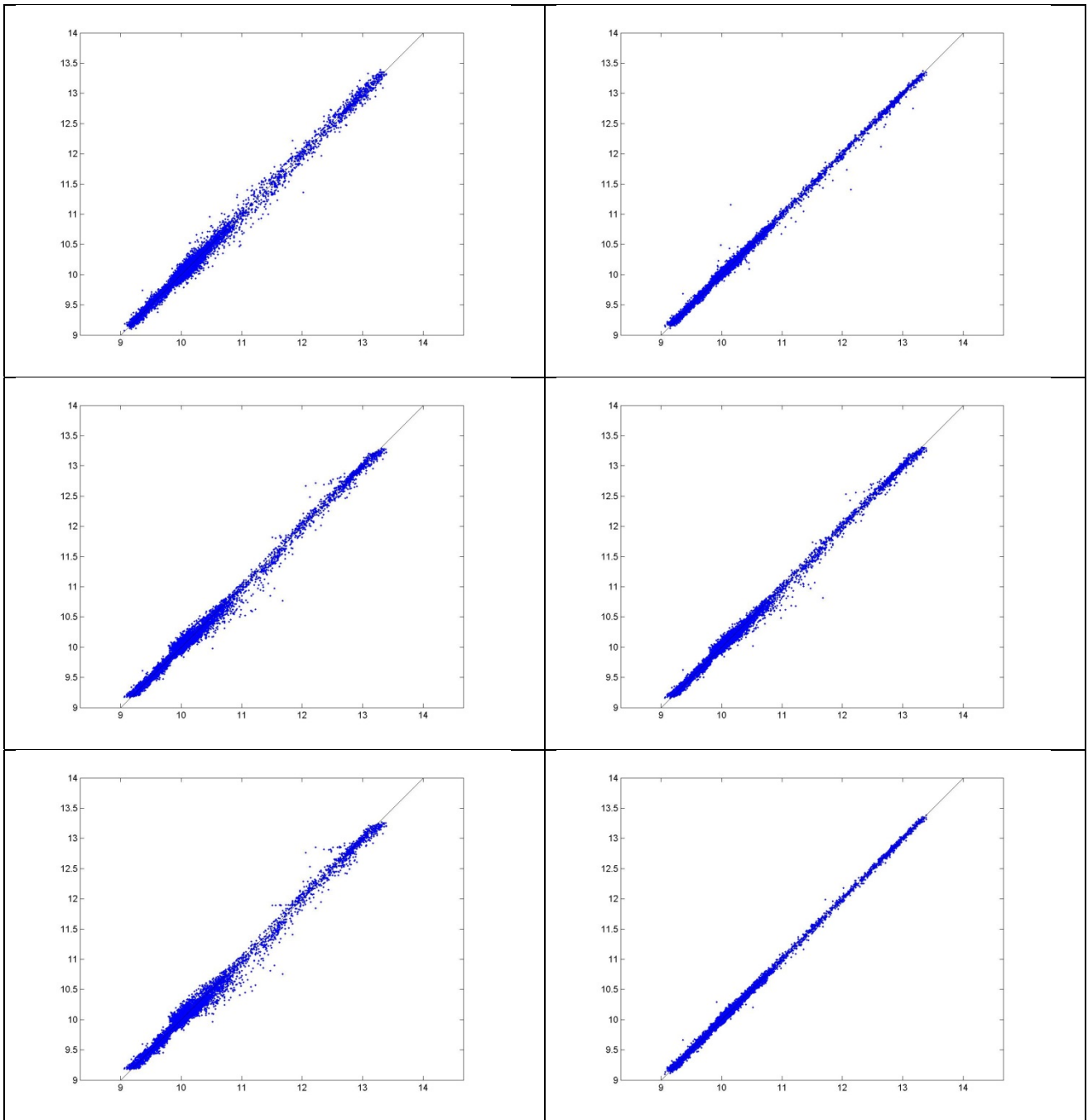


OKSf_2013

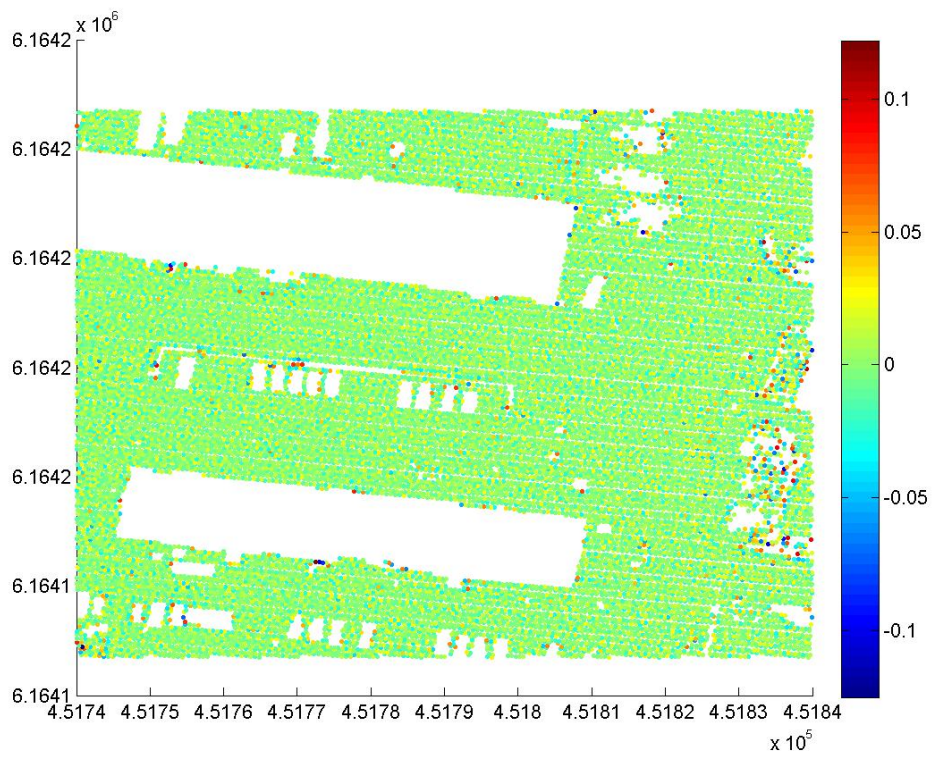
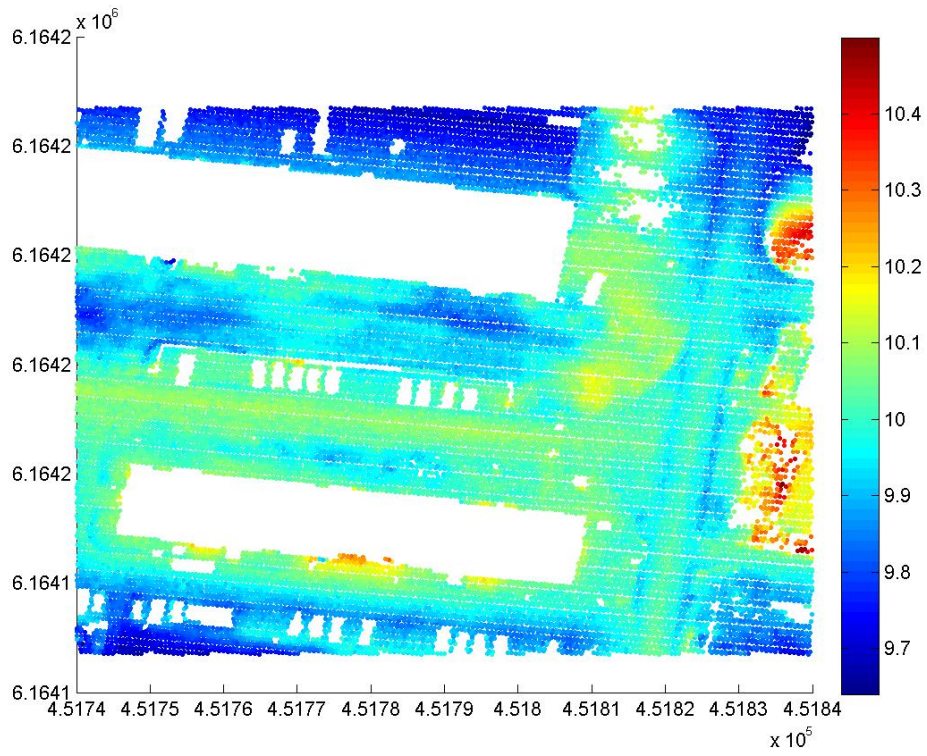


Bias for NN is 0.000717 m, and for TIN -0.000118 m. RMSE for NN is 0.049237 m, and for TIN 0.030853 m.

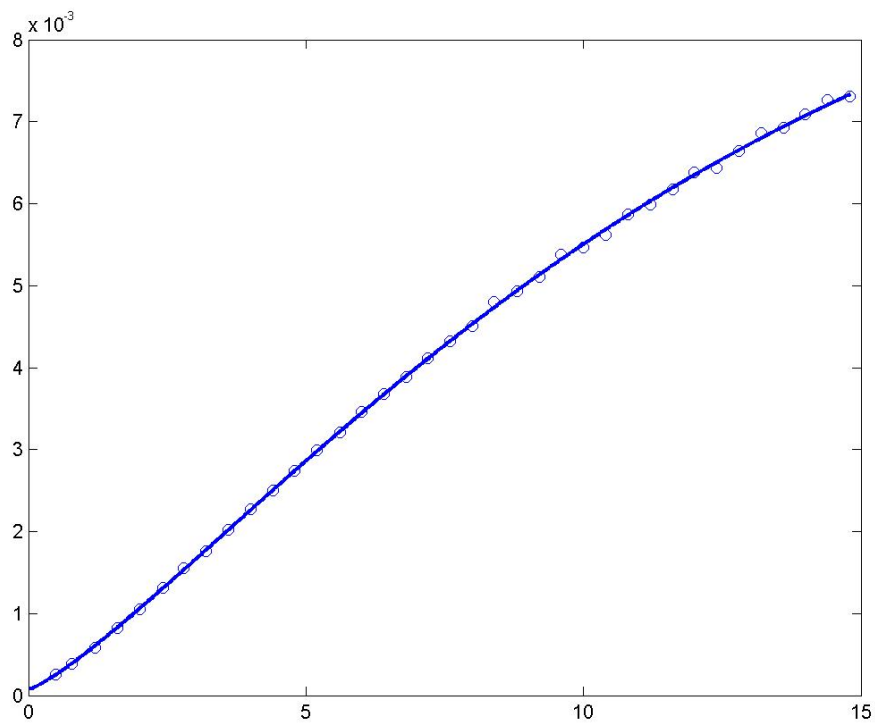
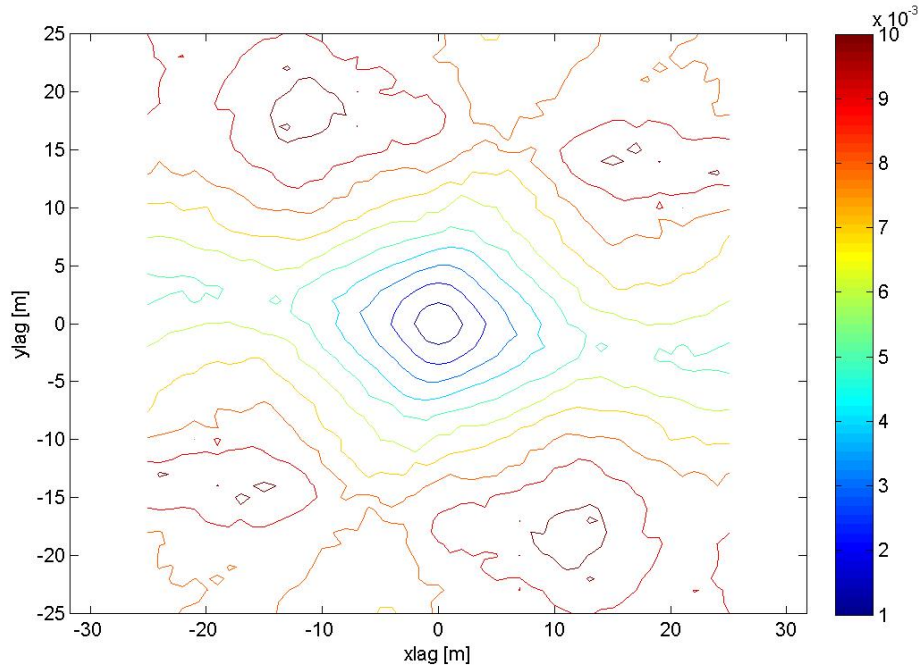
OKSf_2013



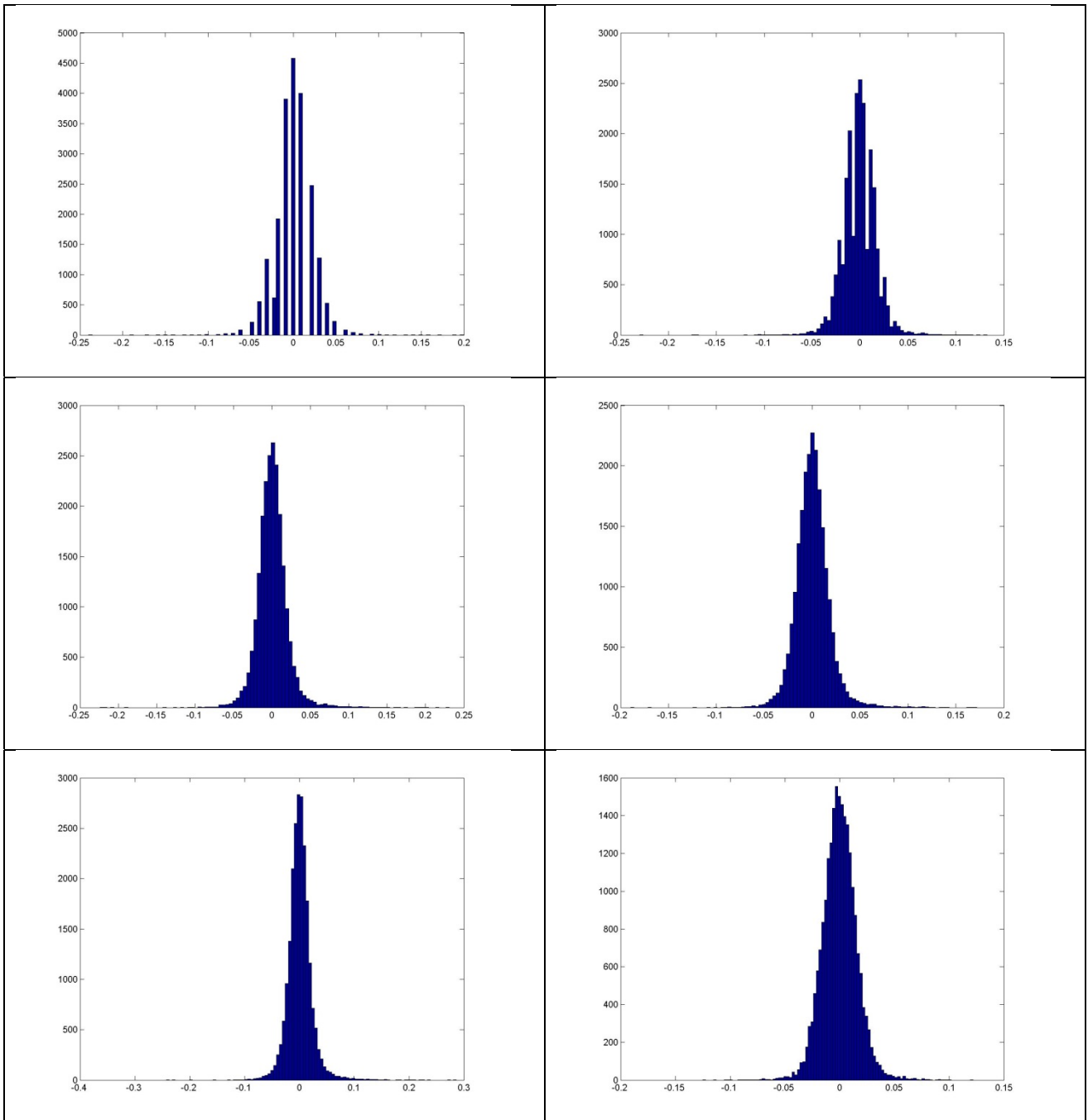
OKSg_2013



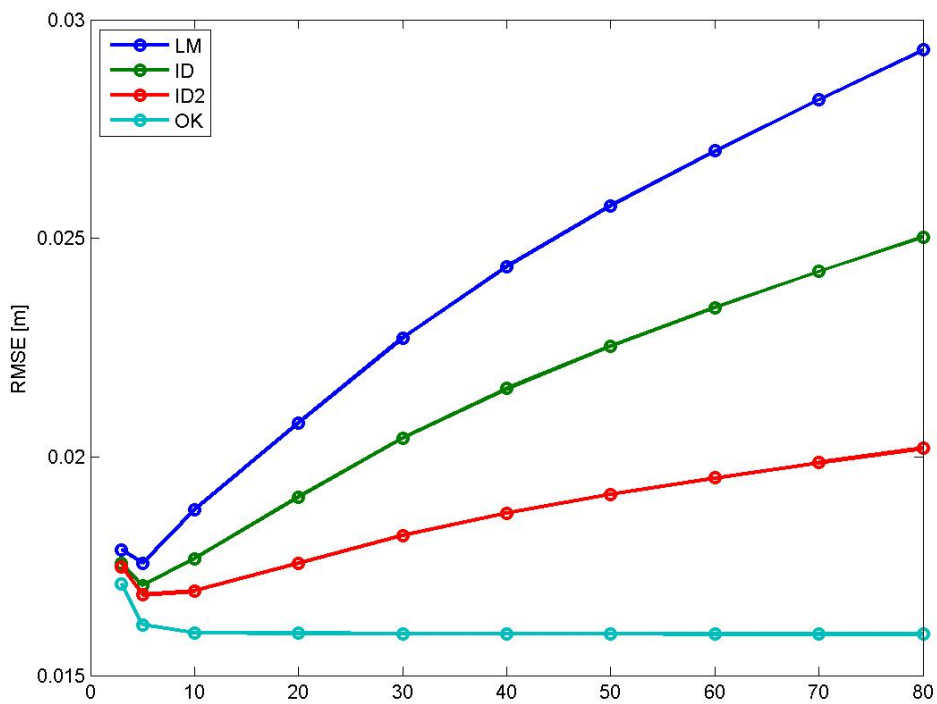
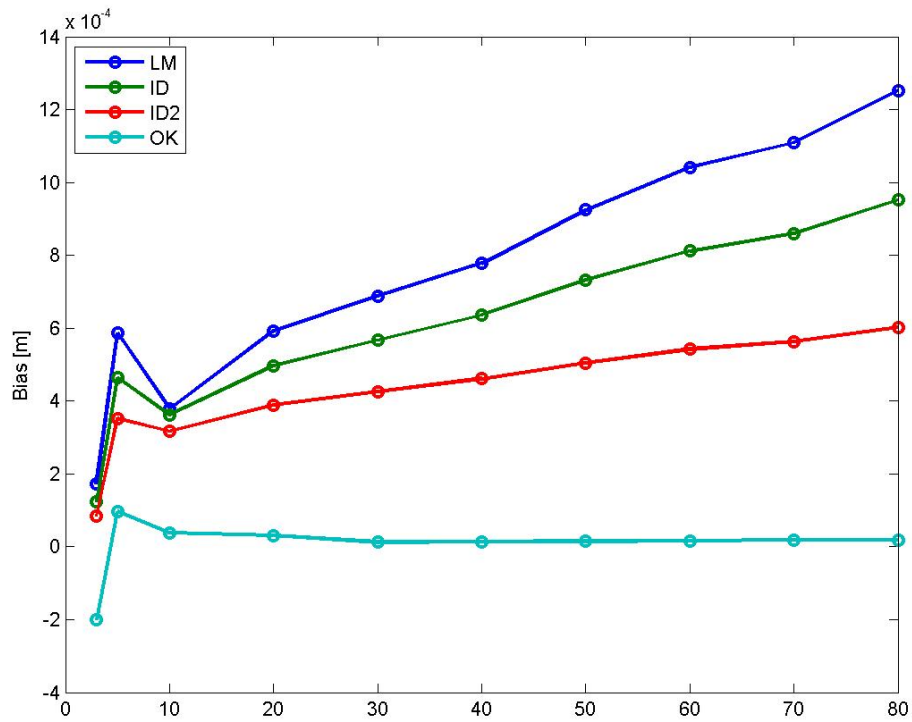
OKSg_2013



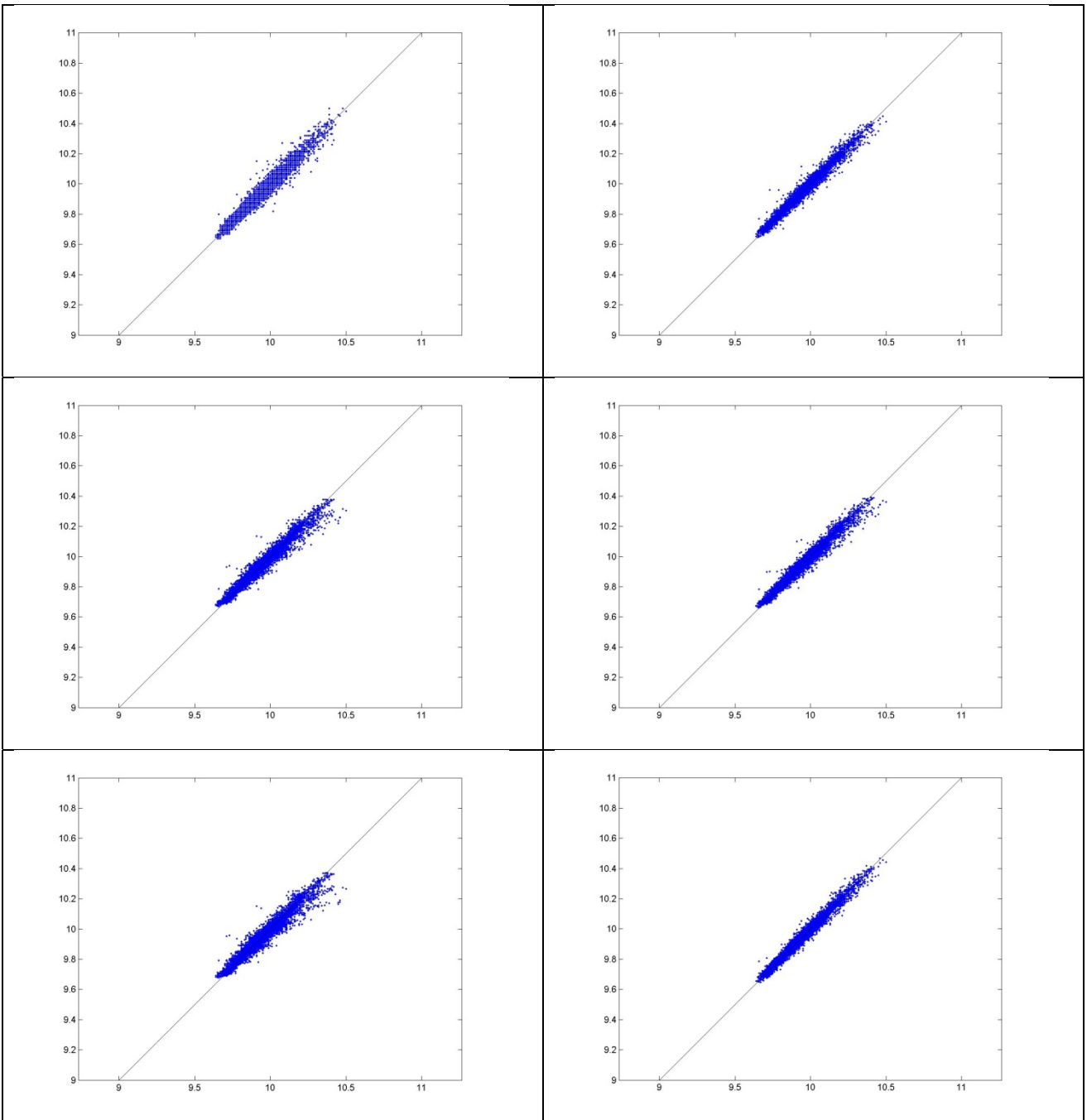
A modified Gaussian semivariogram model is used (power is 1.25 and not 2).



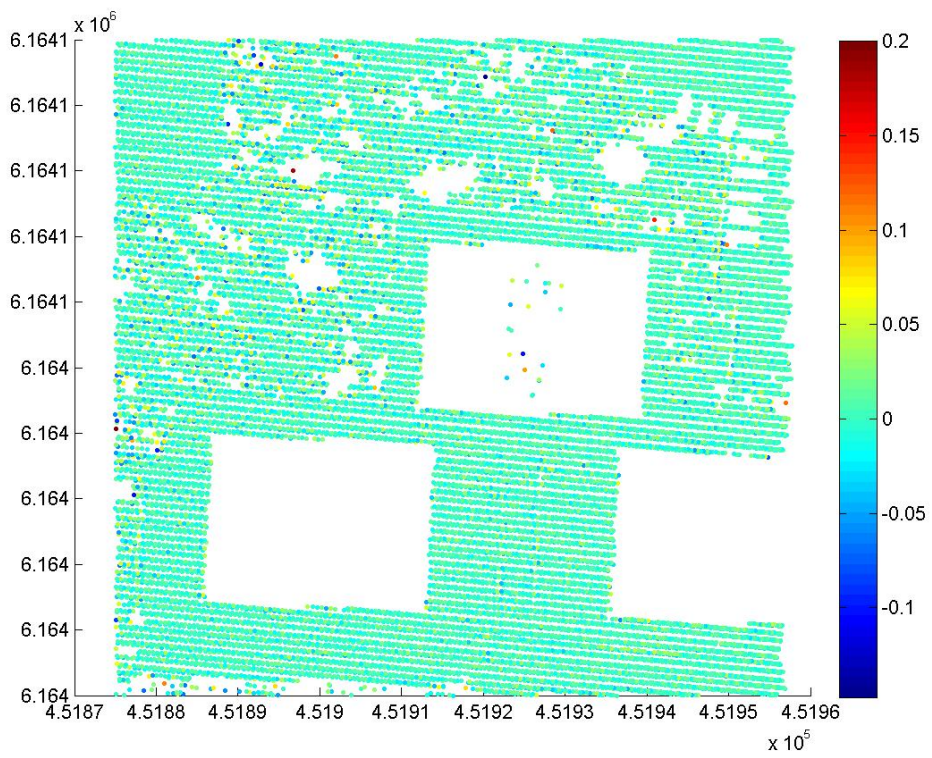
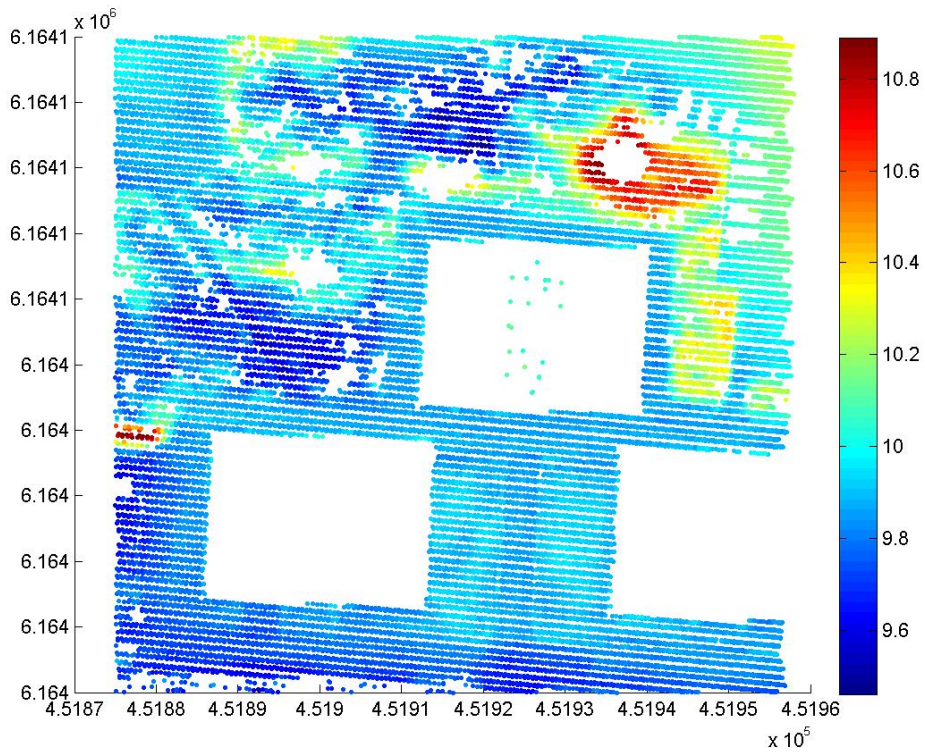
OKSg_2013



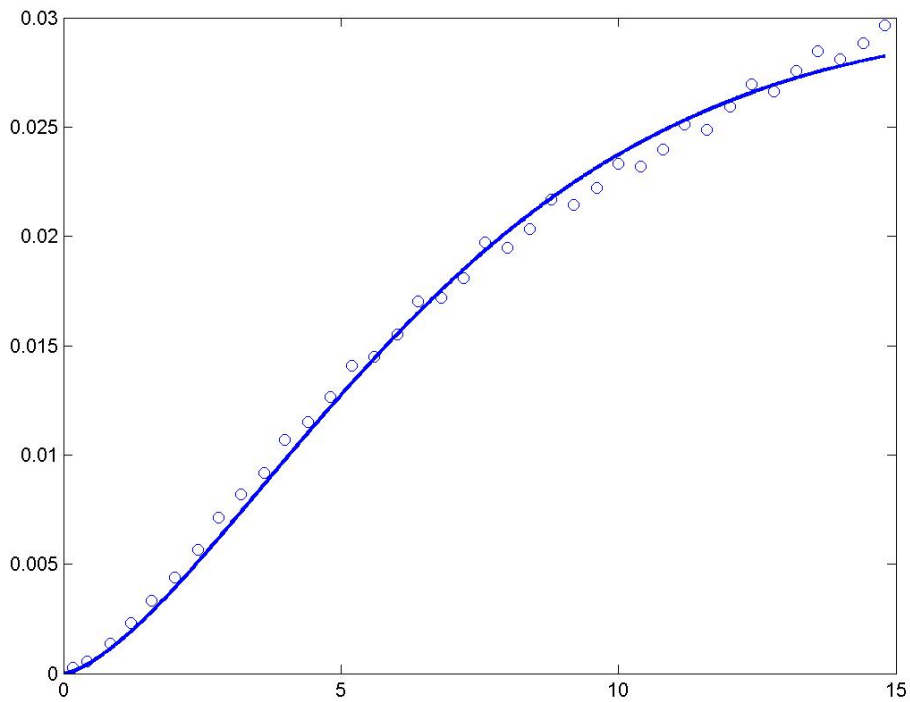
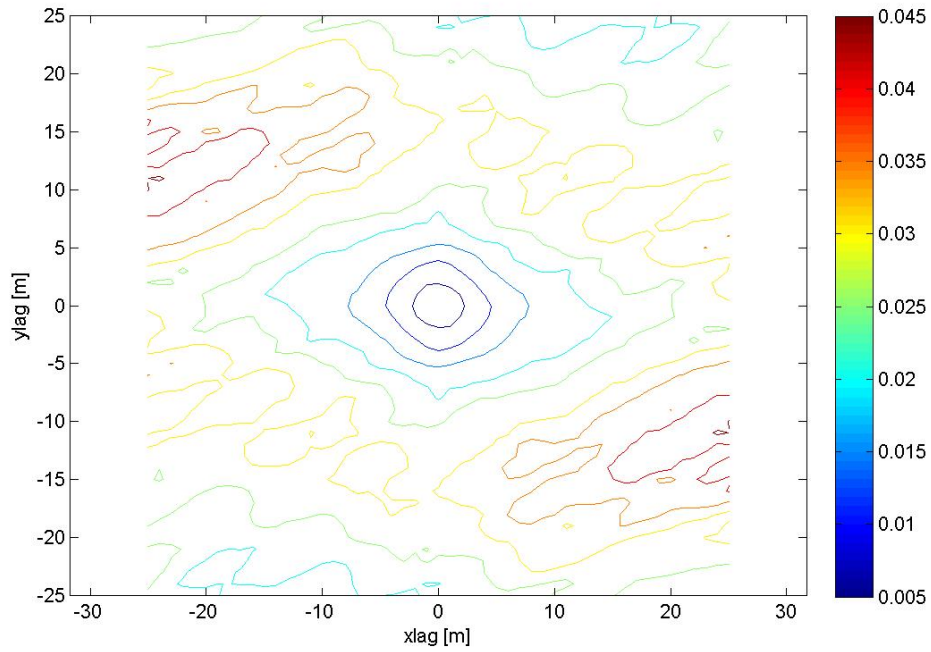
Bias for NN is 7.425e-05 m, and for TIN -0.0008217 m. RMSE for NN is 0.02177 m, and for TIN 0.01730 m.



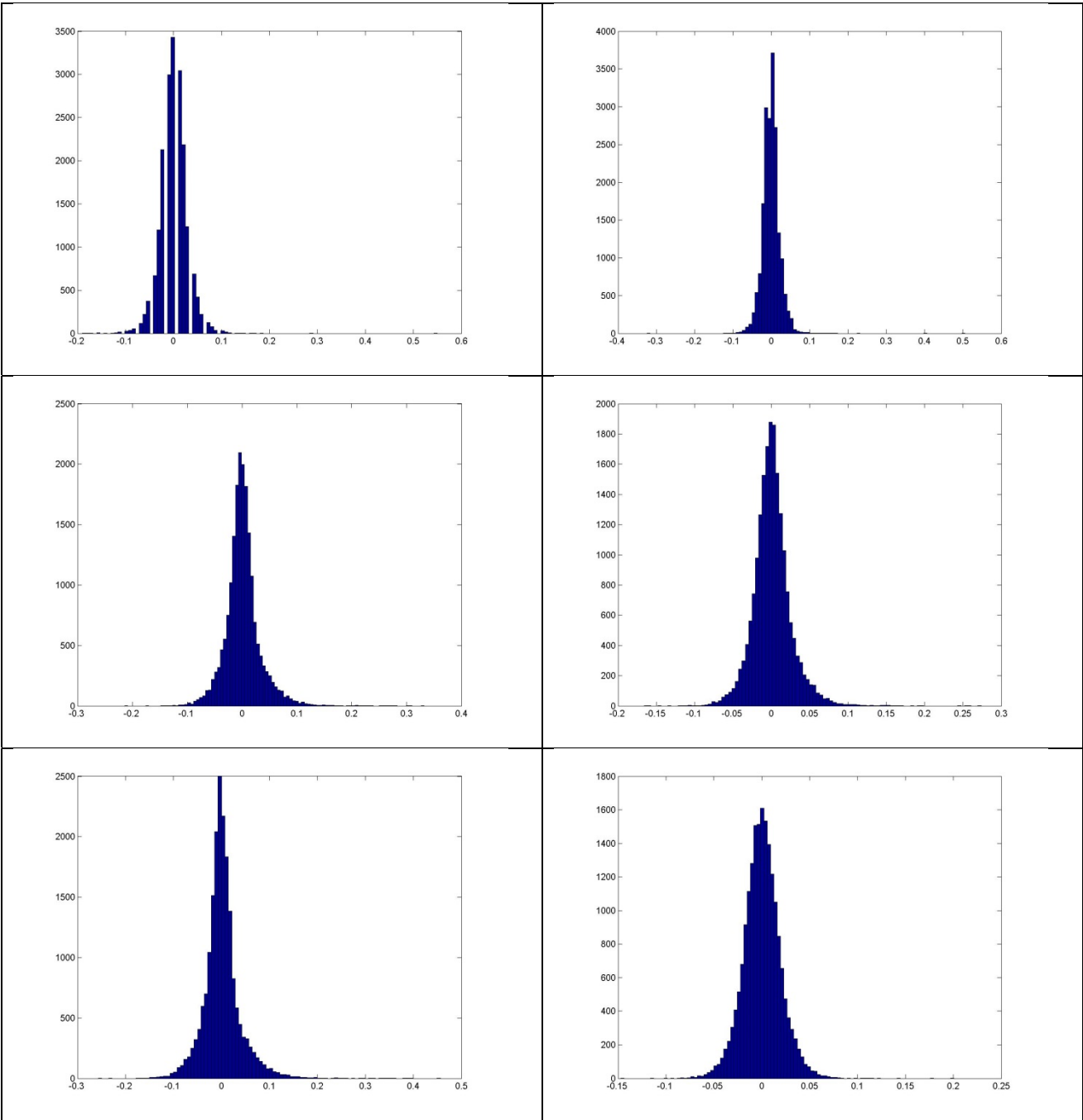
OKSh_2013



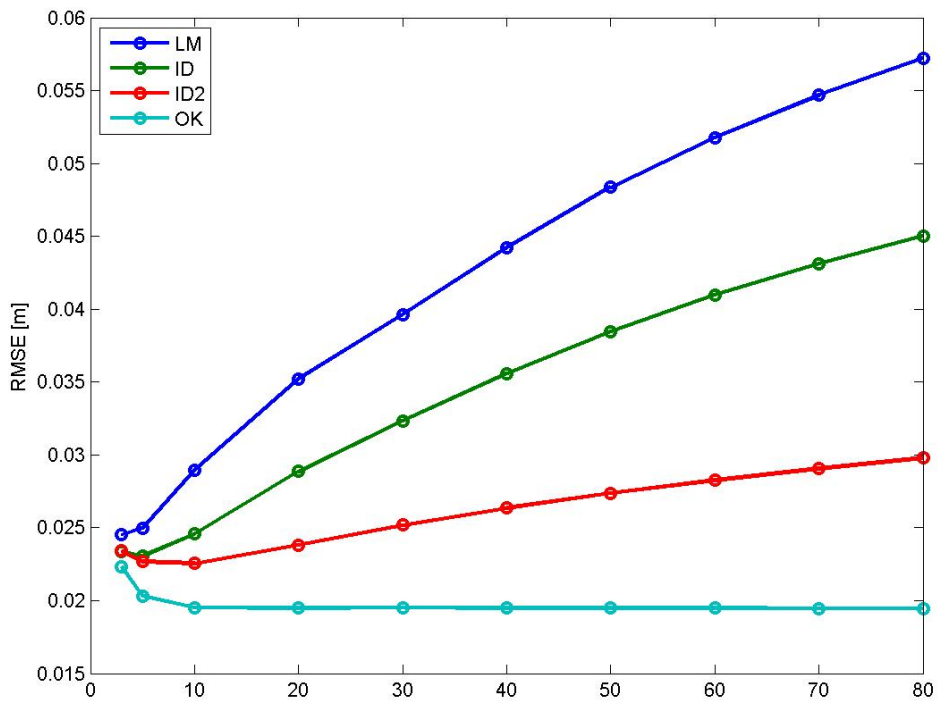
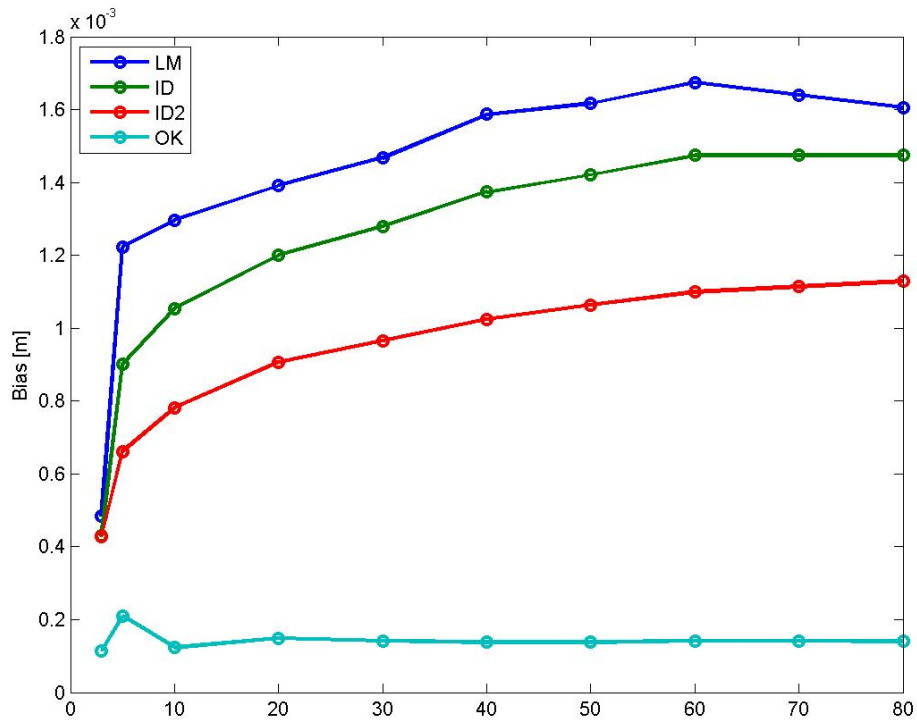
OKSh_2013



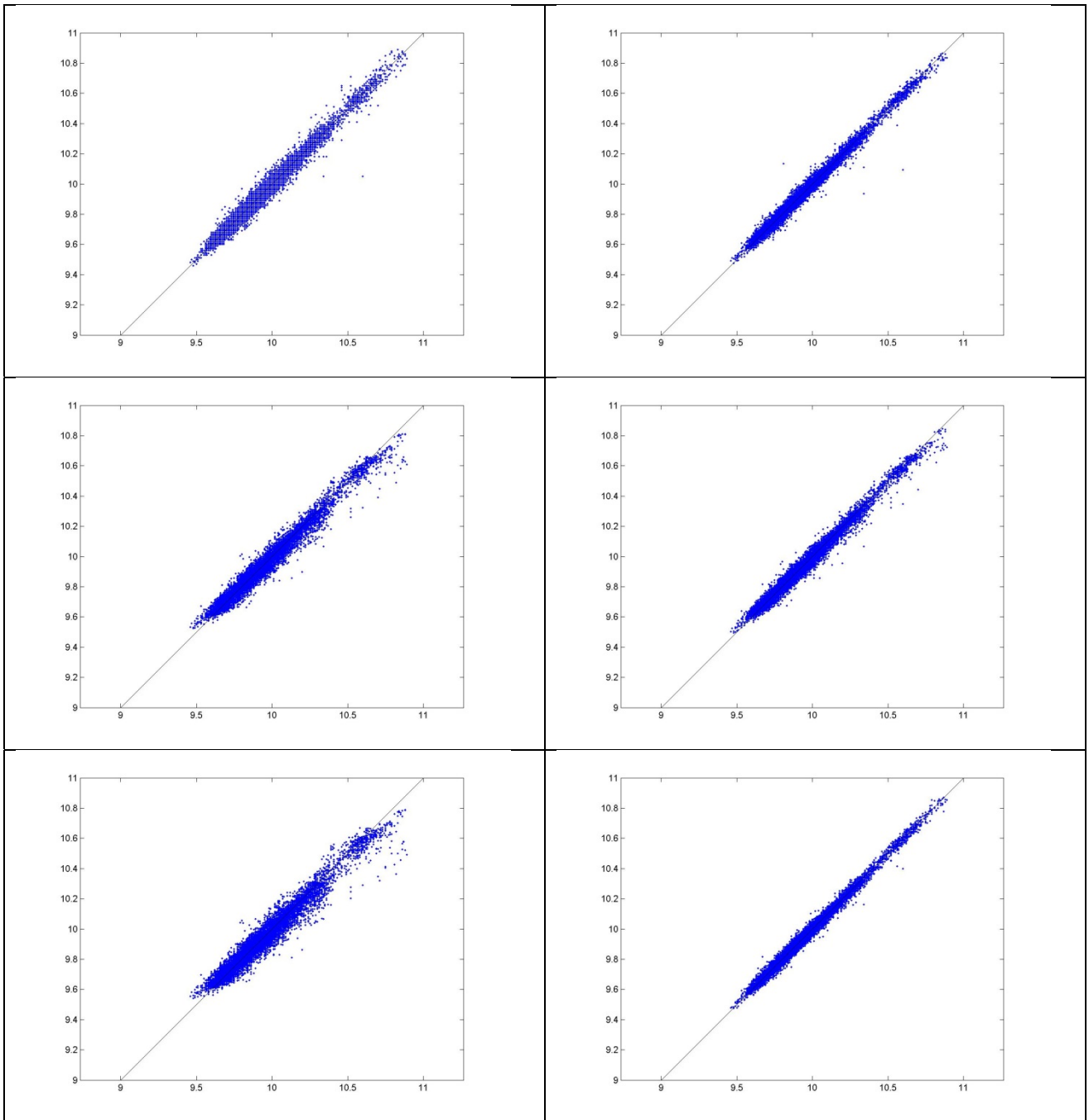
A modified Gaussian semivariogram model is used (power is 1.25 and not 2), no nugget effect.



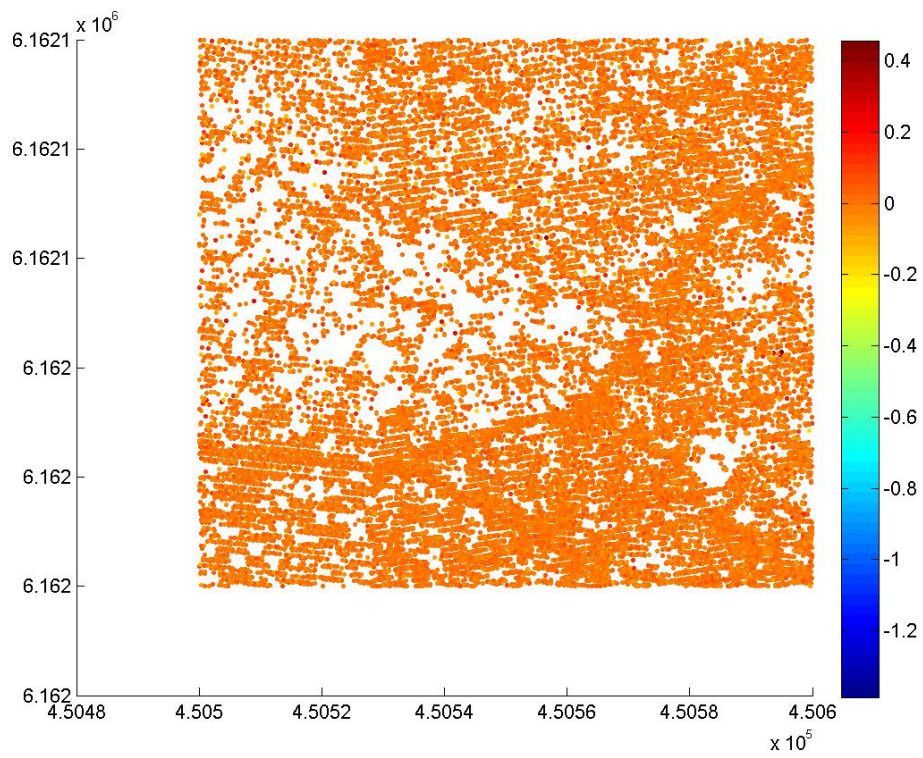
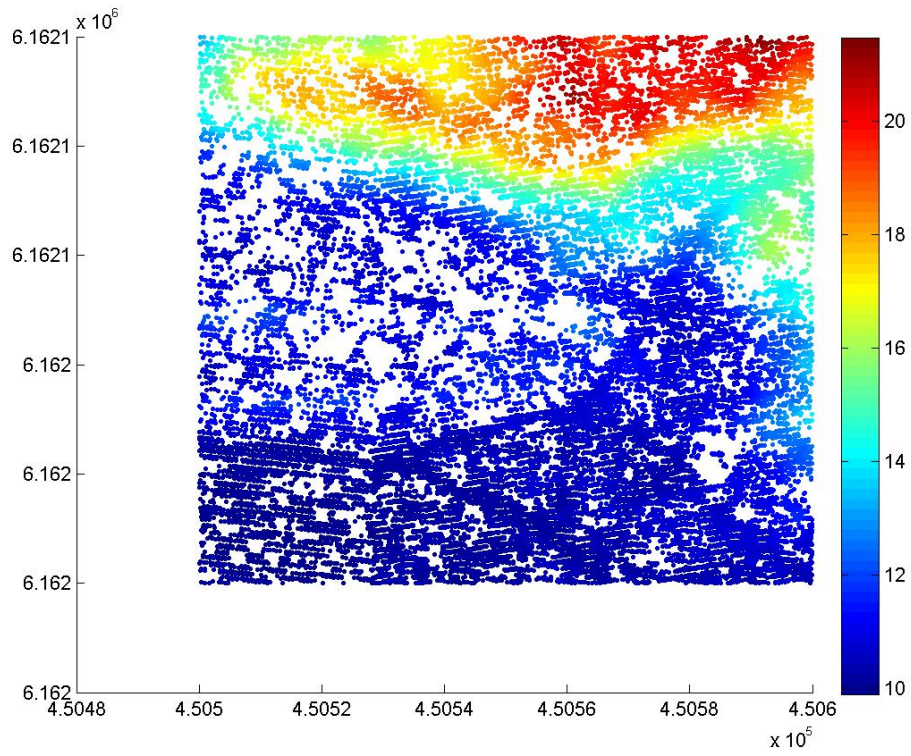
OKSh_2013



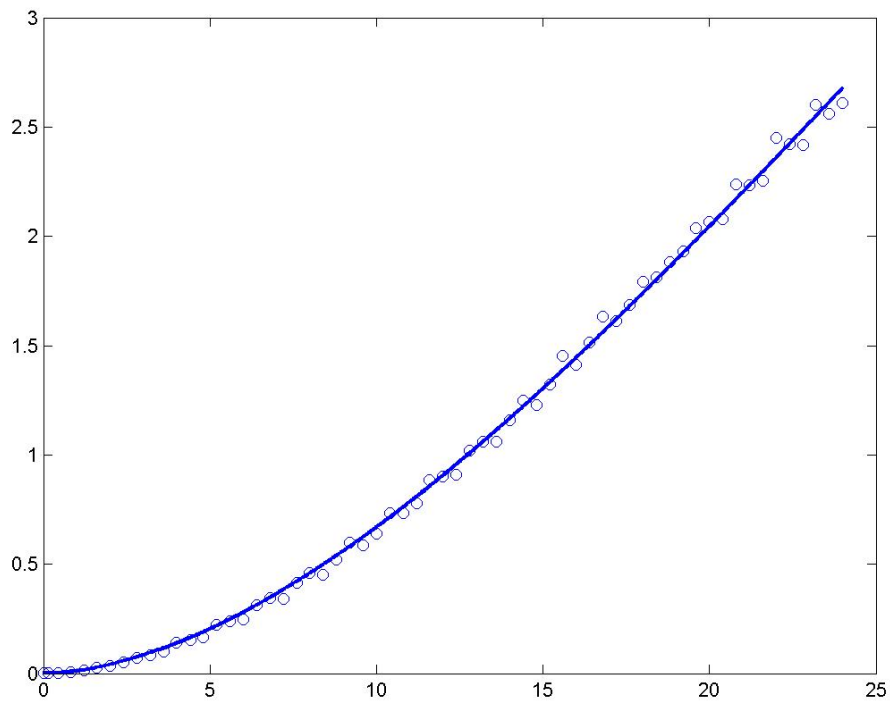
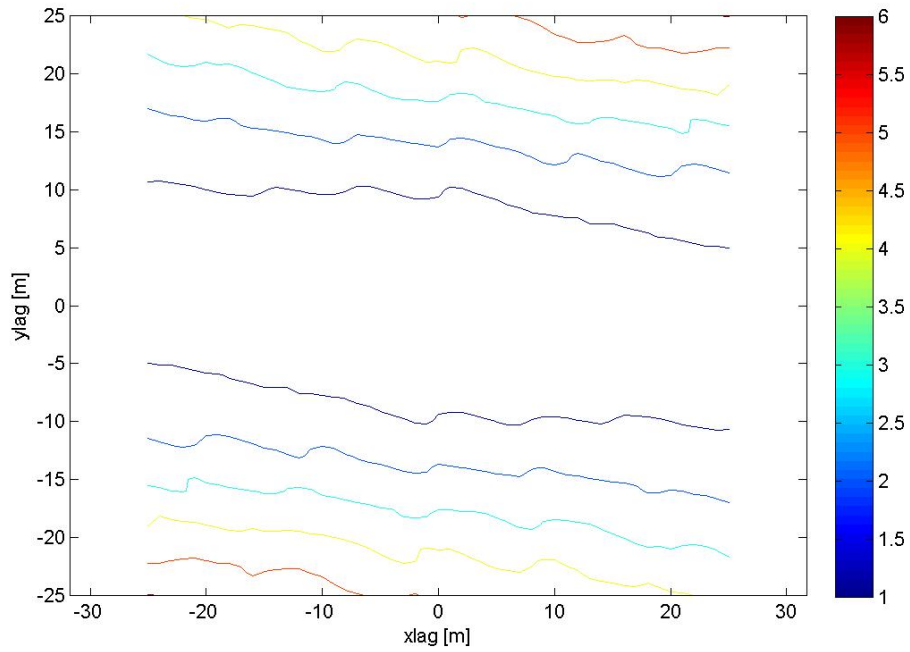
Bias for NN is 0.0005214 m, and for TIN -0.000865 m. RMSE for NN is 0.028123 m, and for TIN 0.022446 m.



OKSx_2013

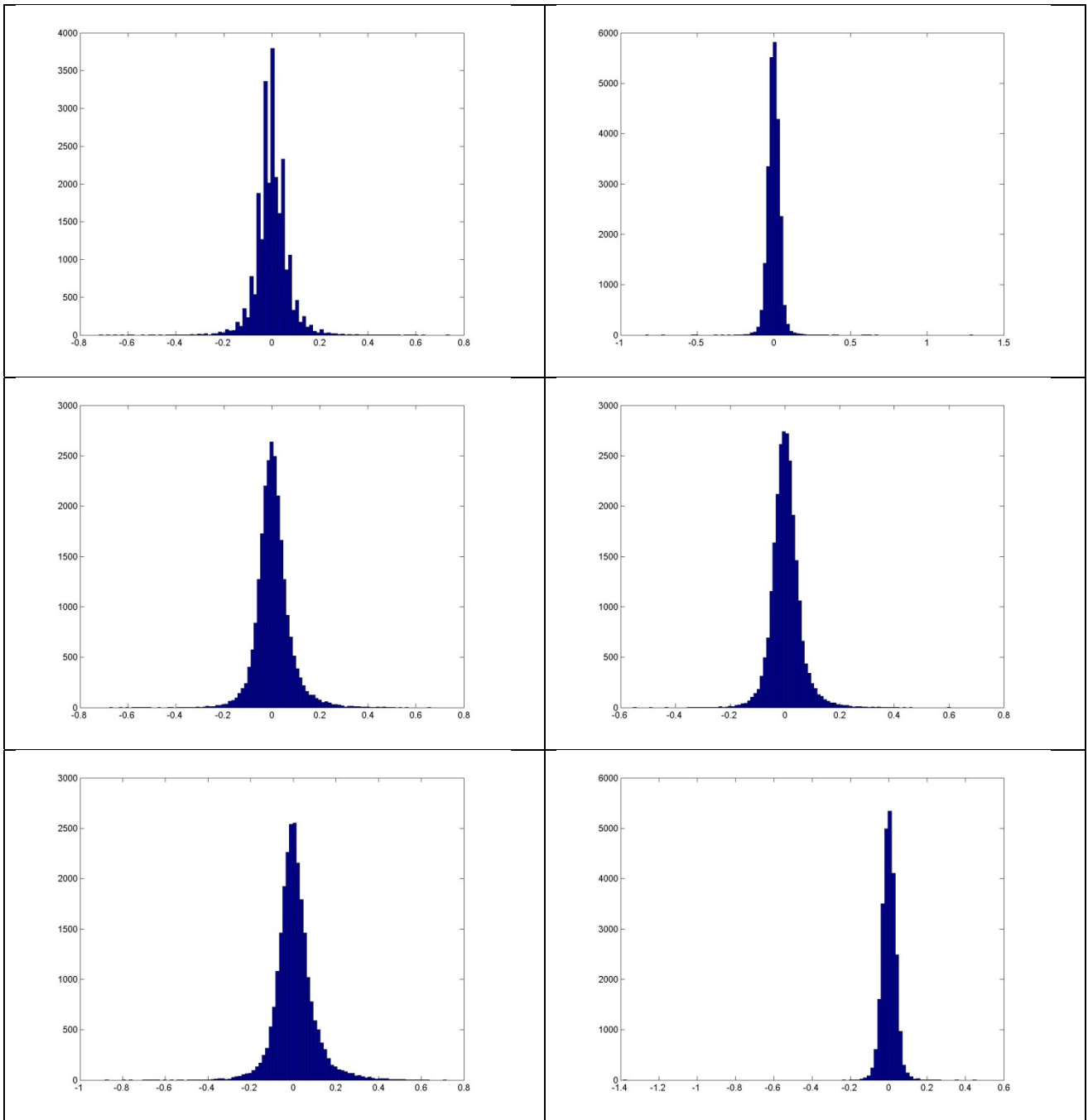


OKSx_2013

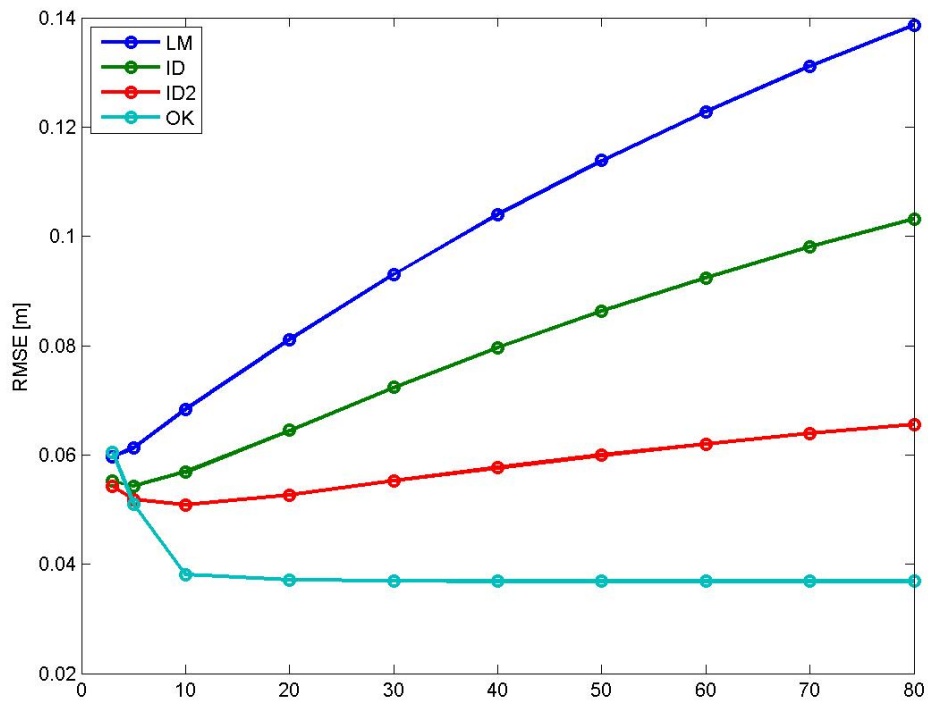
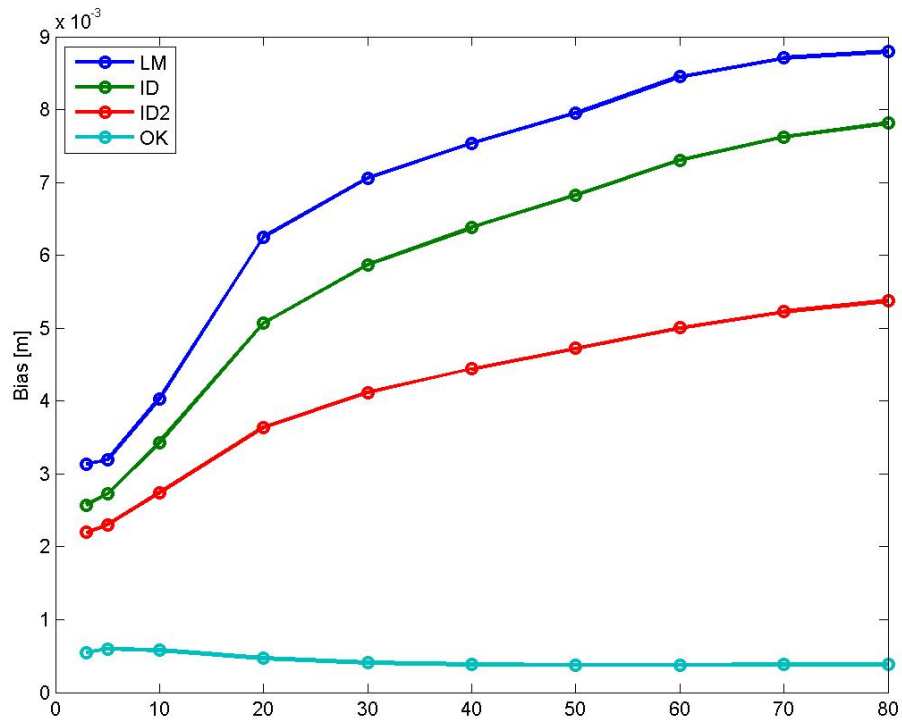


A modified Gaussian semivariogram model is used (power is 1.75 and not 2), no nugget effect.

OKSx_2013

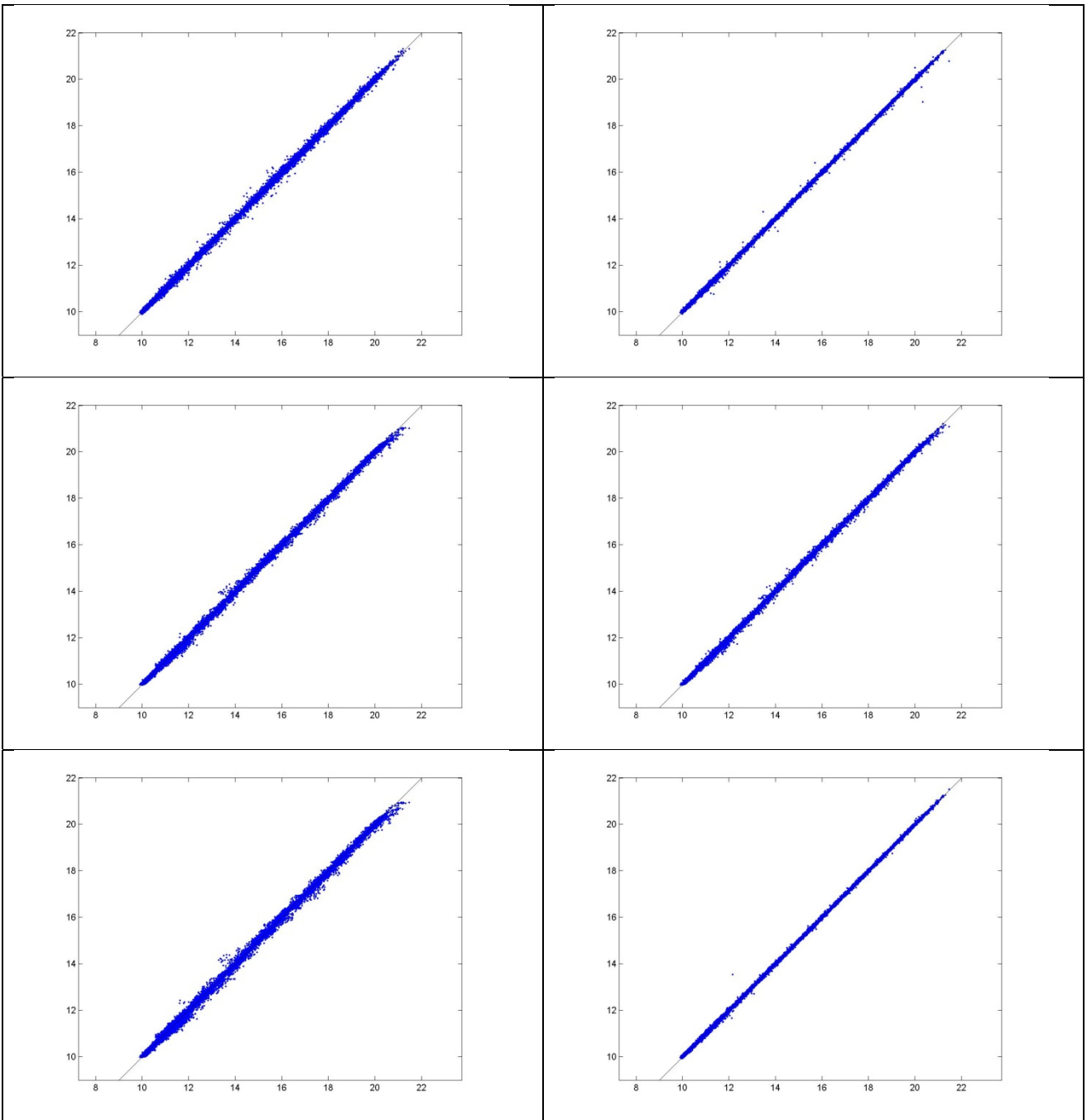


OKSx_2013

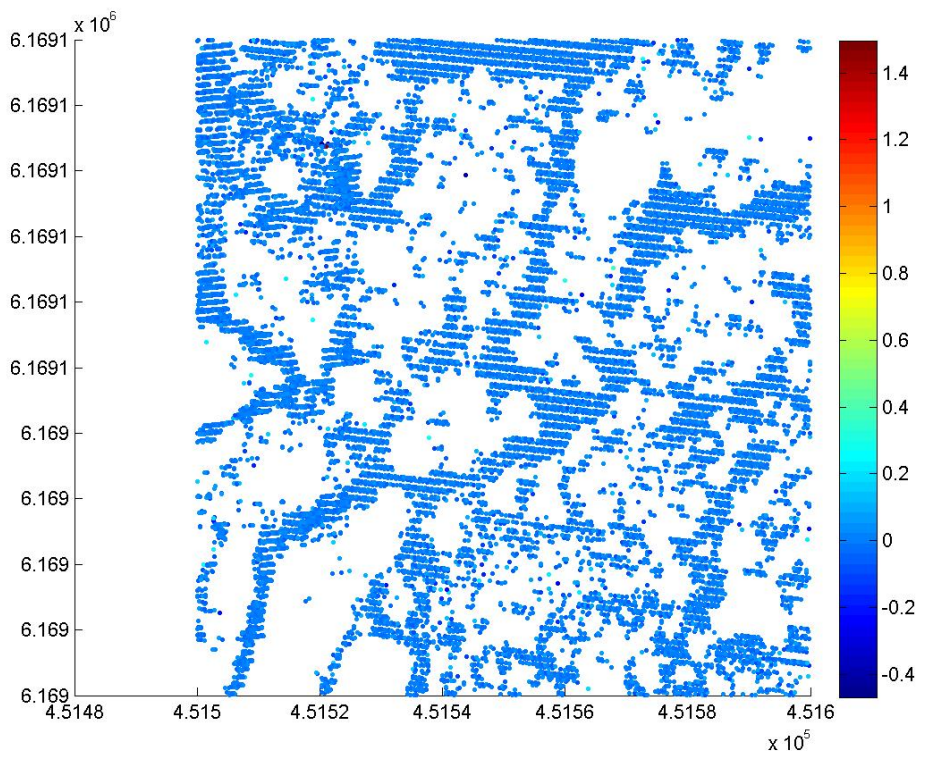
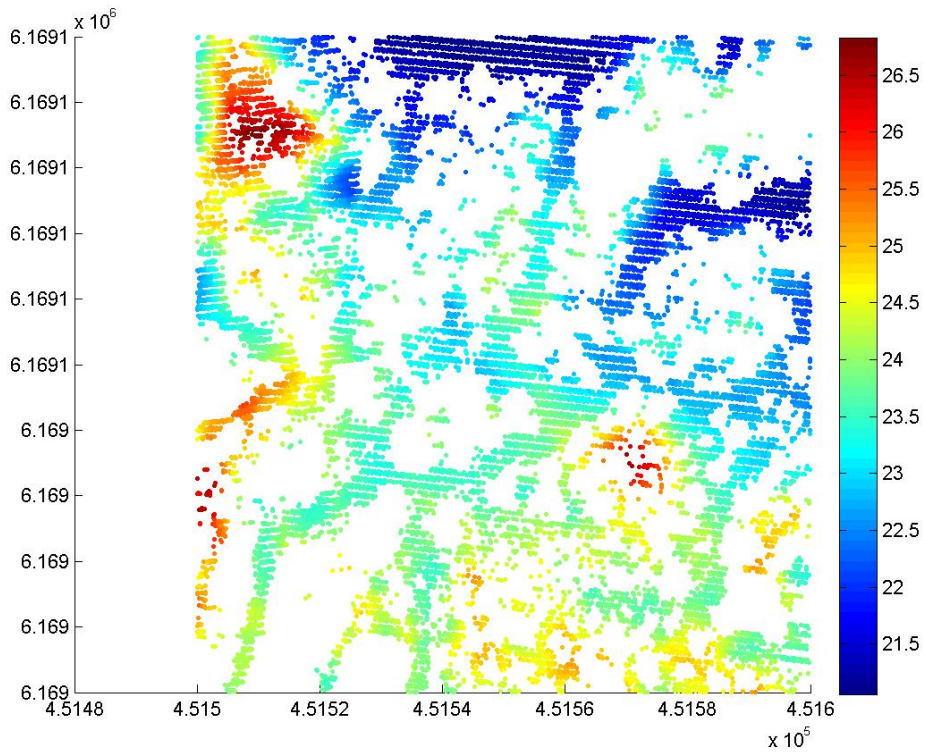


Bias for NN is 0.001561 m, and for TIN -0.0007537 m. RMSE for NN is 0.067633 m, and for TIN 0.041353 m.

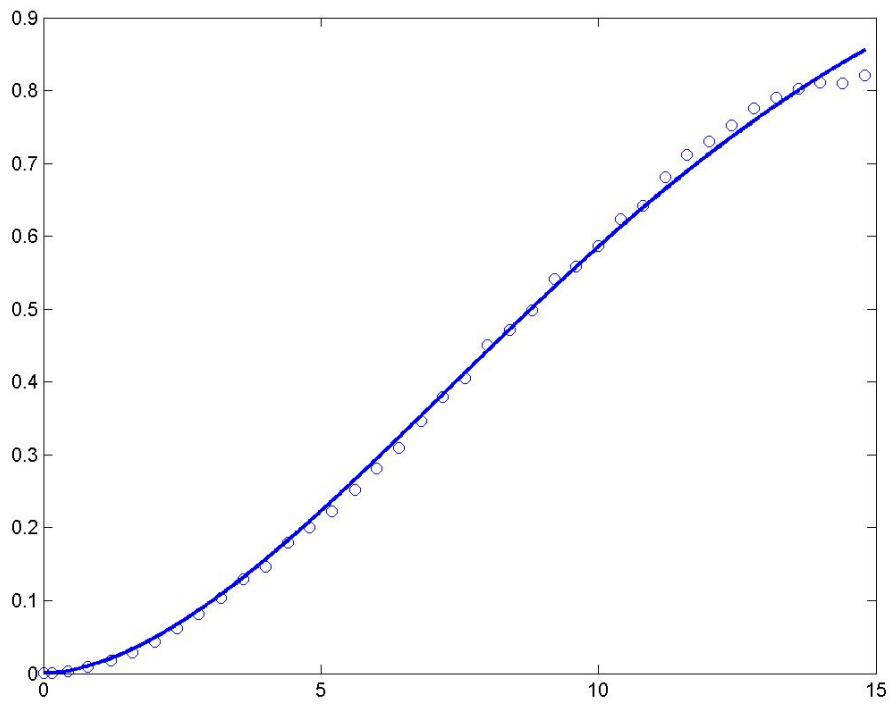
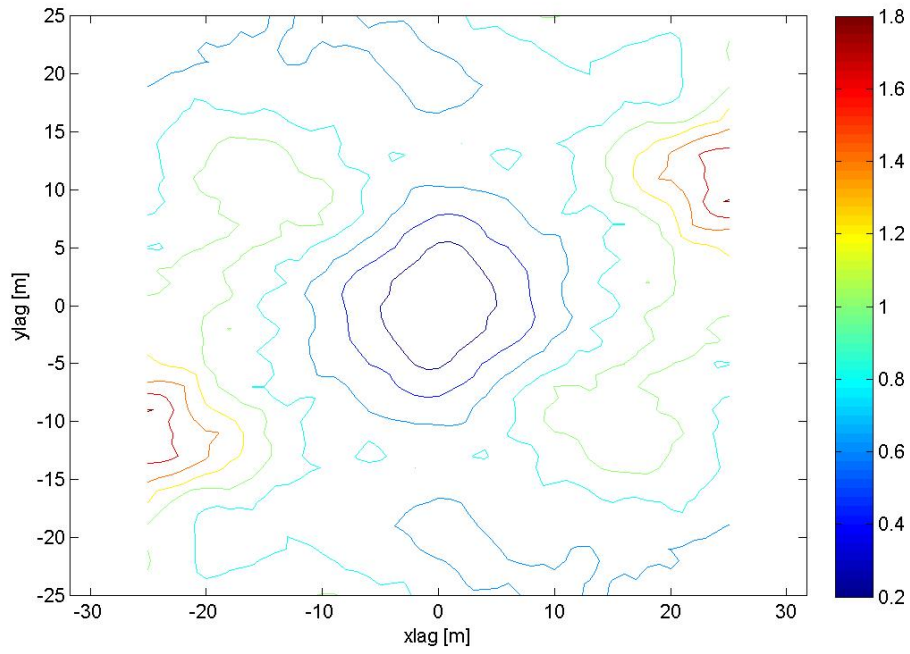
OKSx_2013



OKSy_2013

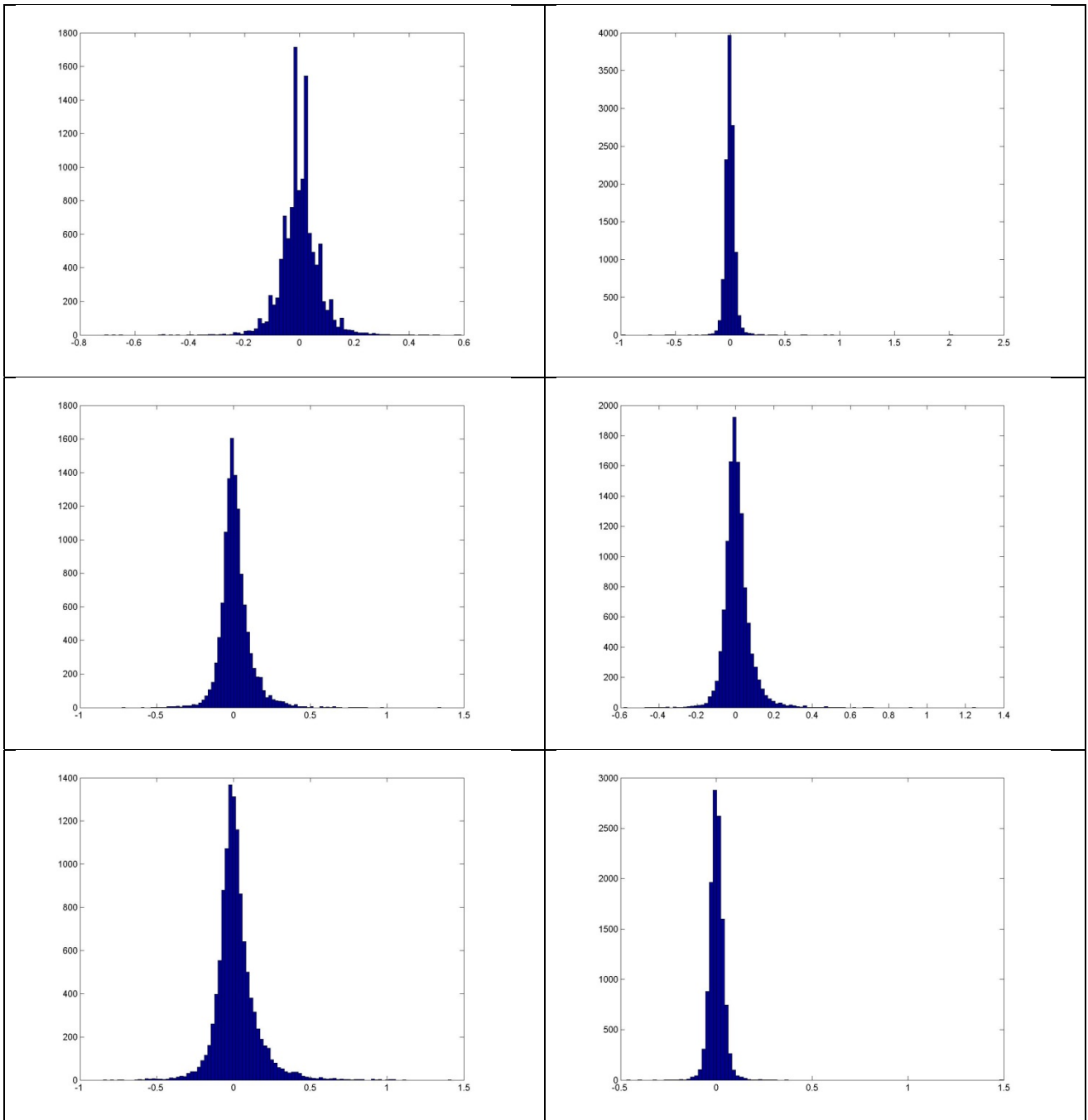


OKSy_2013

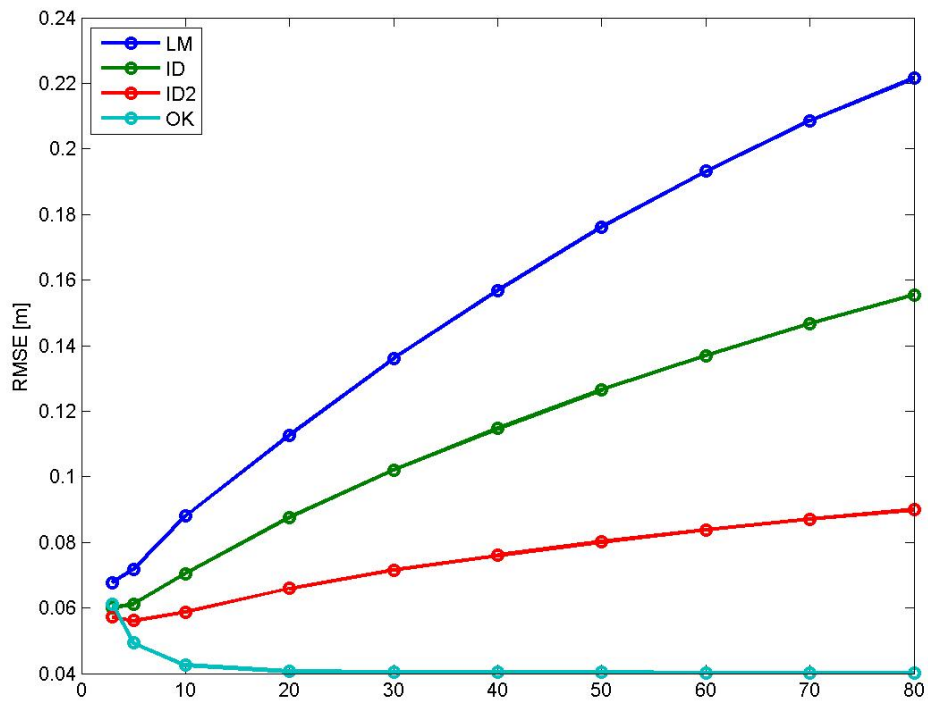
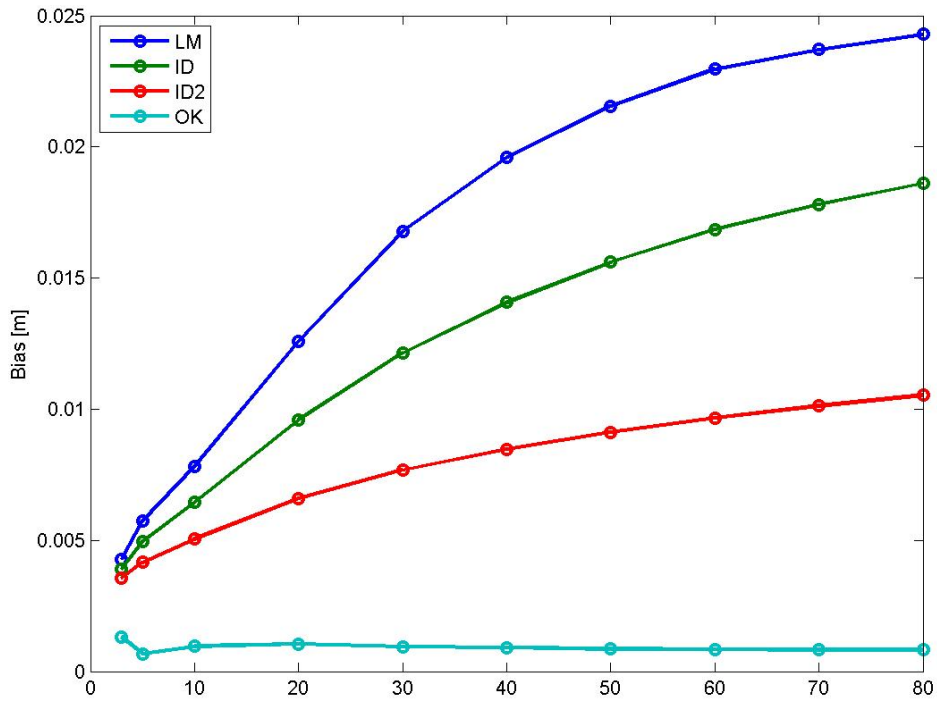


A modified Gaussian semivariogram model is used (power is 1.75 and not 2), no nugget effect.

OKSy_2013



OKSy_2013



Bias for NN is 0.002464 m, and for TIN -0.0006516 m. RMSE for NN is 0.070317 m, and for TIN 0.052671 m.

OKSy_2013

