# Prediction of fine particulate matter chemical components with the spatio-temporal model for the Multi-Ethnic Study of Atherosclerosis cohort

Sun-Young Kim, Lianne Sheppard, Silas Bergen, Adam A. Szpiro, Paul D. Sampson, Joel D. Kaufman, Sverre Vedal

Text summary: This supplementary information includes 5 tables and 10 figures to provide additional information for data summary, intermediate results of modeling procedures, and modeling results.

Table S1. List of geographic variables

| Category | Measure | Variable description |
| --- | --- | --- |
| Traffic | Distance to the nearest road | Any road, A1, intersection |
| | Sum of road lengths within buffers of 0.05-5 km [1] | A1, A2+A3 [3], truck route, intersections |
| Population | Sum within buffers of 0.5-15 km [1] | Population in block groups |
| Land use (Urban) | Percent within buffers of 0.3-15 km [1] | Urban or built-up land |
| | | (residential, commercial, industrial, transportation, urban) |
| | | Developed low, medium, and high density |
| | | Developed open space |
| Land use (Rural) | Percent within buffers of 0.3-15 km [1] | Agricultural land (cropland, groves, feeding) |
| | | Rangeland (herbaceous, shrub) |
| | | Forest land (deciduous, evergreen, mixed) |
| | | Water (streams, lakes, reservoirs, bays) |
| | | Wetland |
| | | Barren land (beaches, dry salt flats, sand, mines, rock) |
| | | Tundra |
| | | Perennial snow or ice |
| Position | Coordinates | Longitude, latitude |
| Source | Distance to the nearest source | Coastline |
| | | Commercial area |
| | | Railroad |
| | | Railyard |
| | | Airport |
| | | Major airport |
| | | Large port |
| | | City hall [4] |
| Emissions | Sum within buffers of 3-30 km [1] | $PM_{2.5}$ |
| | | $PM_{10}$ |
| | | CO |

| | | |
|---|---|---|
| | | $SO_2$<br>$NO_X$ |
| Vegetation | Quantiles within buffers of 0.25-10 km [1] | Normalized Difference Vegetation Index (NDVI) |
| Imperviousness | Percent within buffers of 0.05-5 km [1] | Impervious surface |
| Elevation | Elevation above sea levels | Absolute elevation |
| | Counts of points above or below the location on buffers of 1 and 5 km [1] | Relative elevation |
| Residual oil [2] | Distance to the nearest boiler | Residual oil grade 4 or 6 boiler |
| | Sum within buffers of 0.1-3 km [1] | Residual oil grade 4 or 6 boiler |

1. 0.05-5 km buffers for traffic and impervious surface: 50, 100, 300, 400, 500, 750, 1000, 1500, 3000, 5000 m; 0.5-15 km buffers for population: 500, 1000, 1500, 2000, 2500, 3000, 5000, 10000, 15000 m; 0.3-15 km buffers for land use: 300, 400, 500, 750, 1000, 1500, 3000, 5000, 10000, 15000 m; 3-30 km buffers for emission: 3, 15, and 30 km; 0.25-10 km for vegetation: 250, 500, 1000, 2500, 5000, 7500, 10000 m; 1-5 km for residual oil: 100, 150, 300, 500, 750, 1500, 3000 m
2. New York only
3. Road classification based on CFCC (Census Feature Class Code): primary highways with limited access are coded as A1, primary roads without limited access as A2, and secondary connecting roads as A3
4. Considered as a surrogate for possible pollution sources in central parts of metropolitan areas

Table S2. List of selected geographical variables for four $PM_{2.5}$ components

| Category | Variable name [1] | Variable description |
|---|---|---|
| Traffic | ll.a1.s<radius> [2] | Sum of CFCC (Census Feature Class Code) A1 roads within a <radius> meter buffer in meters |
| | ll.a23.s<radius> | Sum of CFCC A2 and A3 roads within a <radius> meter buffer in meters |
| | tl.s<radius> | Sum of truck rout lengths within a <radius> meter buffer in meters |
| | interchange12.s<radius> | Sum of intersections of A1 and A2 roads within a <radius> meter buffer |
| | interchange3.s<radius> | Sum of intersections of A3 and any roads within a <radius> meter buffer |
| | intersect.s<radius> | Sum of intersections of any roads within a <radius> meter buffer |
| | log10.m.to.a1 | Log-transformed (base 10) meters to nearest A1 road |
| | log10.m.to.interchange3 | Log-transformed (base 10) meters to nearest intersection of A3 roads |
| | log10.m.to.truck | Log-transformed (base 10) meters to nearest truck route |
| | log2.interchange12.s<radius> | Log-transformed (base 2) sum of intersections of A1 and A2 roads within a <radius> meter buffer |
| | log2.interchange3.s<radius> | Log-transformed (base 2) sum of intersections of A3 and any roads within a <radius> meter buffer |
| | log2.intersect.s<radius> | Log-transformed (base 2) sum of intersections of any roads within a <radius> meter buffer |
| Land use | lu.bays.p<radius> | Percentage of 1980 land use type of bays and estuaries within a <radius> meter buffer |
| | lu.comm.p<radius> | Percentage of 1980 land use type of commercial and service areas within a <radius> meter buffer |
| | lu.crop.p<radius> | Percentage of 1980 land use type of cropland and pasture within a <radius> meter buffer |
| | lu.forest.p<radius> | Percentage of 1980 land use type of deciduous forest land within a <radius> meter buffer |
| | lu.green.p<radius> | Percentage of 1980 land use type of evergreen forest land within a <radius> meter buffer |
| | lu.grove.p<radius> | Percentage of 1980 land use type of orchards, groves, vineyards, nurseries within a <radius> meter buffer |
| | lu.industrial.p<radius> | Percentage of 1980 land use type of industrial areas within a <radius> meter buffer |
| | lu.oth.urban.p<radius> | Percentage of 1980 land use type of other urban or built-up land within a <radius> meter buffer |
| | lu.transition.p<radius> | Percentage of 1980 land use type of transitional areas within a <radius> meter buffer |
| | lu.transport.p<radius> | Percentage of 1980 land use type of transportation, communications and utilities within a <radius> meter buffer |
| | lc.anyforest.p<radius> | Percentage of 1980 land use type of woody wetland or forest area within a <radius> meter buffer |
| | lc.openbasic.p<radius> | Percentage of 1980 land use type of any forest, water, crop, shrub, pasture, herb, grass, and barren areas within a <radius> meter buffer |
| | lc.openplus.p<radius> | Percentage of 1980 land use type of any forest, water, crop, shrub, pasture, herb, grass, barren, and developed open space within a <radius> meter buffer |
| | lc.water.p<radius> | Percentage of 1980 land use type of open water within a <radius> meter buffer |

| | | |
|---|---|---|
| | rlu.grass.p<radius> | Percentage of 2006 land use type of grasslands and herbaceous vegetation within a <radius> meter buffer |
| | rlu.water.p<radius> | Percentage of 2006 land use type of open water within a <radius> meter buffer |
| | rlu.dev.hi.p<radius> | Percentage of 2006 land use type of developed high intensity within a <radius> meter buffer |
| | rlu.dev.med.p<radius> | Percentage of 2006 land use type of developed medium intensity within a <radius> meter buffer |
| | rlu.dev.open.p<radius> | Percentage of 2006 land use type of developed open space within a <radius> meter buffer |
| | rlc.dev.medhi.p<radius> | Percentage of 2006 land use type of developed high and medium intensity within a <radius> meter buffer |
| | rlc.openbasic.p<radius> | Percentage of 2006 land use type of any forest, water, crop, shrub, pasture, herb, grass, and barren areas within a <radius> meter buffer |
| | rlc.anyforest.p<radius> | Percentage of 2006 land use type of woody wetland or forest area within a <radius> meter buffer |
| Position | long | GPS longitude coordinate in decimal degrees |
| Source | m.to.rr | Meters to the nearest railroad |
| | log10.m.to.comm | Log-transformed (base 10) meters to nearest commercial zone |
| | log10.m.to.l.airp | Log-transformed (base 10) meters to nearest large airport |
| | log10.m.to.rr | Log-transformed (base 10) meters to nearest railroad |
| Emissions | em.CO.s30000 | Sum of CO emissions in tons per year from tall stacks within 30 km, minus the emissions from tall stacks within 3 km |
| Vegetation | ndvi.q25.a<radius> | Average 2006 NDVI value at the first quantile within a <radius> meter buffer |
| | ndvi.q75.a<radius> | Average 2006 NDVI value at the third quantile within a <radius> meter buffer |
| Imperviousness | imp.a<radius> | Average impervious surface value (percent imperviousness) within a <radius> meter buffer |
| Elevation | elev.1k.above | Count of points (out of 24 concentric points) more than 20 m uphill of the location on 1 km buffer |
| | elev.1k.below | Count of points more than 20 m downhill of the location on 1 km buffer |
| | elev.1k.rabove | Squre root-transformed count of points more than 20 m uphill of the location on 1 km buffer |
| | elev.1k.rbelow | Squre root-transformed count of points more than 20 m downhill of the location on 1 km buffer |
| | elev.5k.below | Count of points more than 50 m downhill of the location on 5 km buffer |
| Residual oil | oil.edf4.s<radius> | Sum of potential mega BTU (British Thermal Unit) from oil number 4 per hour within a <radius> meter buffer |
| | log10.m.to.6oil | Log-transformed (base 10) meters to nearest Number 6 oil boiler |

1. The list of variables include variables seen in Figure S4
2. Radius of buffers were presented in Table S2

Table S3. Provisional cross-validation statistics and selected variables from trend-adjusted "long-term average" concentrations at NPACT/MESA Air home-outdoor sites

| | | Cross-validation statistics [1] | | Geographic variables [2] | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| City | Pollutant | RMSE (μg/m$^3$) | $R^2$ | Traffic | Land use (urban) | Land use (rural) | Position | Source | Emission | Vegetation | Imperviousness | Elevation | Residual oil [3] |
| LA | Sulfur | 0.17 | 0.21 | ■ | ■ | | | ■ | | | | ■ | |
| | Silicon | 0.03 | 0.47 | ■ | | ■ | | | | | | | |
| | EC | 0.24 | 0.80 | ■ | ■ | ■ | | | | | | ■ | |
| | OC | 0.68 | 0.18 | | ■ | ■ | | | | | | ■ | |
| Chicago | Sulfur | 0.16 | 0.35 | ■ | | ■ | | ■ | | | | | |
| | Silicon | 0.03 | 0.19 | ■ | ■ | ■ | | | | | | | |
| | EC | 0.18 | 0.48 | ■ | | ■ | | ■ | | | | | |
| | OC | 0.35 | 0.47 | ■ | ■ | | | ■ | | | | | |
| St. Paul | Sulfur | 0.08 | 0.17 | ■ | | | | ■ | | ■ | ■ | ■ | |
| | Silicon | 0.03 | 0.10 | | ■ | ■ | | | | ■ | | | |
| | EC | 0.12 | 0.39 | ■ | ■ | | | | | ■ | | | |
| | OC | 0.13 | 0.60 | ■ | ■ | ■ | | | | | | | |
| Baltimore | Sulfur | 0.15 | 0.12 | | ■ | ■ | | | | | | | |
| | Silicon | 0.00 | 0.59 | | | | ■ | ■ | | ■ | | ■ | |
| | EC | 0.17 | 0.58 | ■ | | ■ | | | | | | ■ | |
| | OC | 0.20 | 0.70 | ■ | | ■ | | | ■ | ■ | | | |
| NY | Sulfur | 0.41 | 0.13 | | ■ | ■ | | | | | | ■ | |
| | Silicon | 0.04 | 0.04 | | | | | | | | | ■ | ■ |
| | EC | 0.60 | 0.55 | | ■ | ■ | | | | | | | |
| | OC | 0.40 | 0.44 | ■ | ■ | ■ | | | | | | | |
| Winston -Salem | Sulfur | 0.18 | 0.25 | | ■ | ■ | | | | | | | |
| | Silicon | 0.00 | 0.21 | ■ | | | ■ | | | | | ■ | |
| | EC | 0.13 | 0.43 | | ■ | ■ | | | | | | ■ | |
| | OC | 0.36 | 0.19 | | ■ | ■ | | ■ | | | | | |

1. Provisional cross-validation approach based on lasso variable selection followed by all subset universal kriging
2. List of geographic variables for each category is shown in Table S1
3. Considered only for New York

Table S4. Area-specific summary statistics for the spatio-temporal and national spatial models of predicted long-term average concentrations of four $PM_{2.5}$ components for 5,493 MESA participants residing within 10 kilometers of any NPACT/MESA Air monitors based on addresses for 2000-2002

| | | | Spatio-temporal model | | National spatial model | |
|---|---|---|---|---|---|---|
| Pollutant | City | N | Mean (μg/m$^3$) | SD (μg/m$^3$) | Mean (μg/m$^3$) | SD (μg/m$^3$) |
| Sulfur | LA | 1,073 | 1.13 | 0.04 | 0.53 | 0.02 |
| | Chicago | 999 | 1.24 | 0.06 | 0.75 | 0.03 |
| | Minneapolis-St. Paul | 898 | 0.81 | 0.01 | 0.53 | 0.01 |
| | Baltimore | 775 | 1.67 | 0.06 | 1.00 | 0.03 |
| | NY | 856 | 1.46 | 0.13 | 0.78 | 0.02 |
| | Winston-Salem | 892 | 1.69 | 0.08 | 0.94 | 0.02 |
| Silicon | LA | 1,073 | 0.14 | 0.02 | 0.14 | 0.00 |
| | Chicago | 999 | 0.11 | 0.01 | 0.06 | 0.00 |
| | Minneapolis-St. Paul | 898 | 0.10 | 0.01 | 0.08 | 0.00 |
| | Baltimore | 775 | 0.09 | 0.01 | 0.08 | 0.00 |
| | NY | 856 | 0.12 | 0.01 | 0.07 | 0.00 |
| | Winston-Salem | 892 | 0.10 | 0.01 | 0.11 | 0.00 |
| EC | LA | 1,073 | 1.98 | 0.34 | 0.83 | 0.14 |
| | Chicago | 999 | 1.40 | 0.20 | 0.74 | 0.13 |
| | Minneapolis-St. Paul | 898 | 0.85 | 0.08 | 0.58 | 0.07 |
| | Baltimore | 775 | 1.35 | 0.21 | 0.68 | 0.15 |
| | NY | 856 | 2.38 | 0.44 | 1.13 | 0.13 |
| | Winston-Salem | 892 | 1.10 | 0.08 | 0.57 | 0.07 |
| OC | LA | 1,073 | 2.33 | 0.29 | 2.43 | 0.20 |
| | Chicago | 999 | 1.92 | 0.28 | 1.71 | 0.23 |
| | Minneapolis-St. Paul | 898 | 1.73 | 0.13 | 2.09 | 0.19 |
| | Baltimore | 775 | 2.19 | 0.38 | 2.12 | 0.24 |
| | NY | 856 | 2.18 | 0.39 | 1.64 | 0.20 |
| | Winston-Salem | 892 | 2.63 | 0.16 | 1.95 | 0.29 |

Table S5. Proportion of total variance of the predictions captured by the long-term mean, temporal trend, and spatio-temporal residuals across MESA Air monitoring home-outdoor sites

| City | Pollutant | Long-term mean [1] | | Temporal trend [1] | Spatio-temporal residual [1] |
|---|---|---|---|---|---|
| | | Regression [2] | Regression+kriging | | |
| LA | Sulfur | 0.00 | 0.01 | 0.82 | 0.17 |
| | Silicon | 0.10 | 0.20 | 0.32 | 0.48 |
| | EC | 0.04 | 0.35 | 0.50 | 0.15 |
| | OC | 0.04 | 0.28 | 0.36 | 0.36 |
| Chicago | Sulfur | 0.02 | 0.04 | 0.52 | 0.43 |
| | Silicon | 0.07 | 0.08 | 0.76 | 0.15 |
| | EC | 0.26 | 0.31 | 0.54 | 0.16 |
| | OC | 0.19 | 0.19 | 0.60 | 0.22 |
| St. Paul | Sulfur | 0.00 | 0.00 | 0.56 | 0.44 |
| | Silicon | 0.02 | 0.02 | 0.60 | 0.38 |
| | EC | 0.39 | 0.41 | 0.41 | 0.18 |
| | OC | 0.12 | 0.13 | 0.71 | 0.15 |
| Baltimore | Sulfur | 0.01 | 0.01 | 0.79 | 0.20 |
| | Silicon | 0.11 | 0.11 | 0.69 | 0.19 |
| | EC | 0.48 | 0.48 | 0.35 | 0.17 |
| | OC | 0.10 | 0.10 | 0.72 | 0.18 |
| NY | Sulfur | 0.09 | 0.09 | 0.54 | 0.37 |
| | Silicon | 0.11 | 0.11 | 0.39 | 0.51 |
| | EC | 0.66 | 0.64 | 0.36 | 0.00 |
| | OC | 0.46 | 0.46 | 0.43 | 0.11 |
| Winston-Salem | Sulfur | 0.01 | 0.01 | 0.84 | 0.15 |
| | Silicon | 0.02 | 0.02 | 0.74 | 0.24 |
| | EC | 0.17 | 0.21 | 0.57 | 0.22 |
| | OC | 0.06 | 0.06 | 0.43 | 0.51 |

1. Sum of ratios of long-term mean including regression and kriging, temporal trend, and spatio-temporal residual is equal to 1; Total variance used as denominator for calculating ratios was sum of variances of long-term mean, temporal trend, and spatio-temporal residual instead of variance of predictions given correlation structure between three parts

2. Ratio of regression part for long-term mean is separately presented to show its contribution to total variability

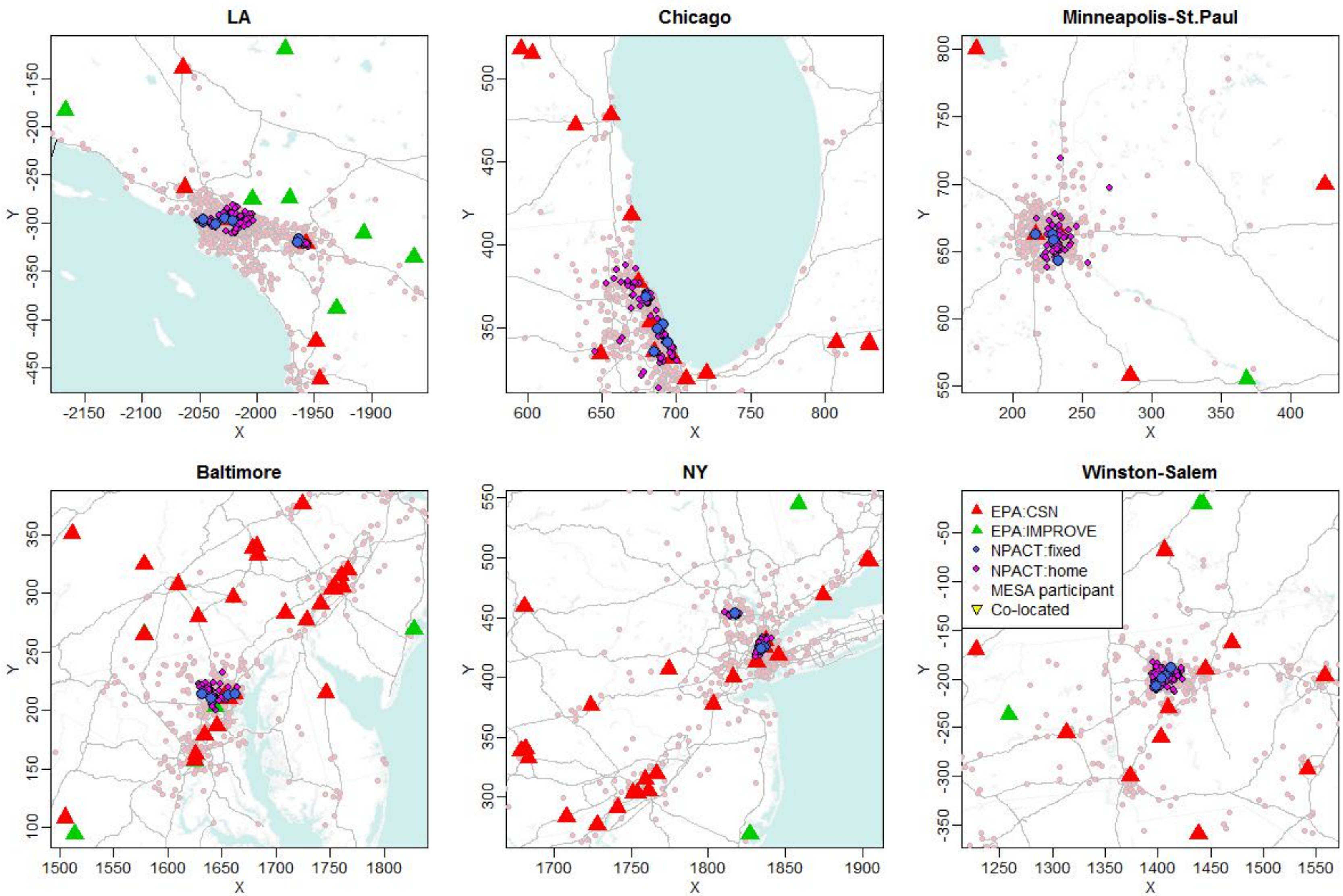



Figure S1. Regulatory and NPACT/MESA Air monitoring sites for $PM_{2.5}$ components located within 200 km from city centers in the six MESA city areas

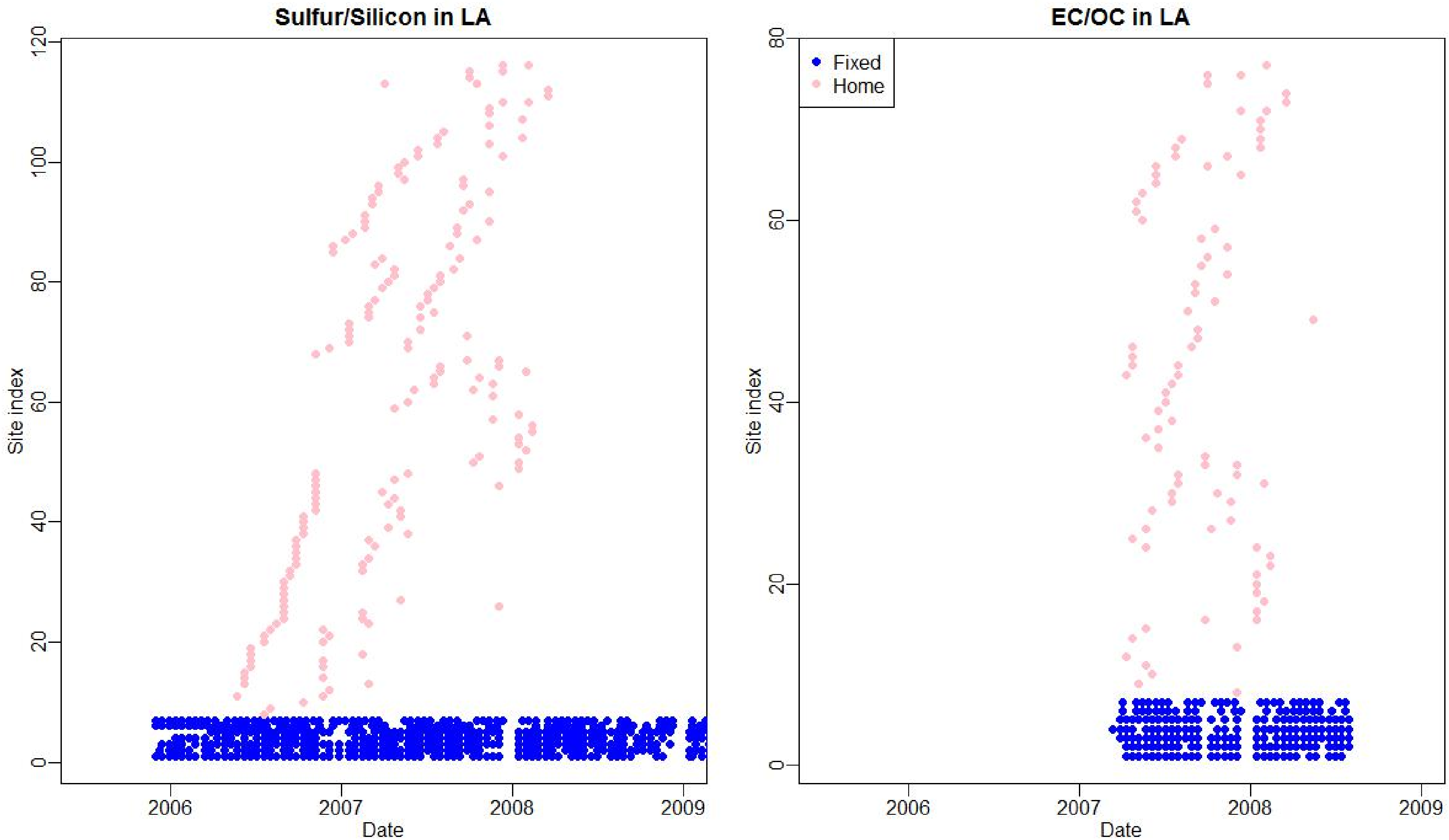
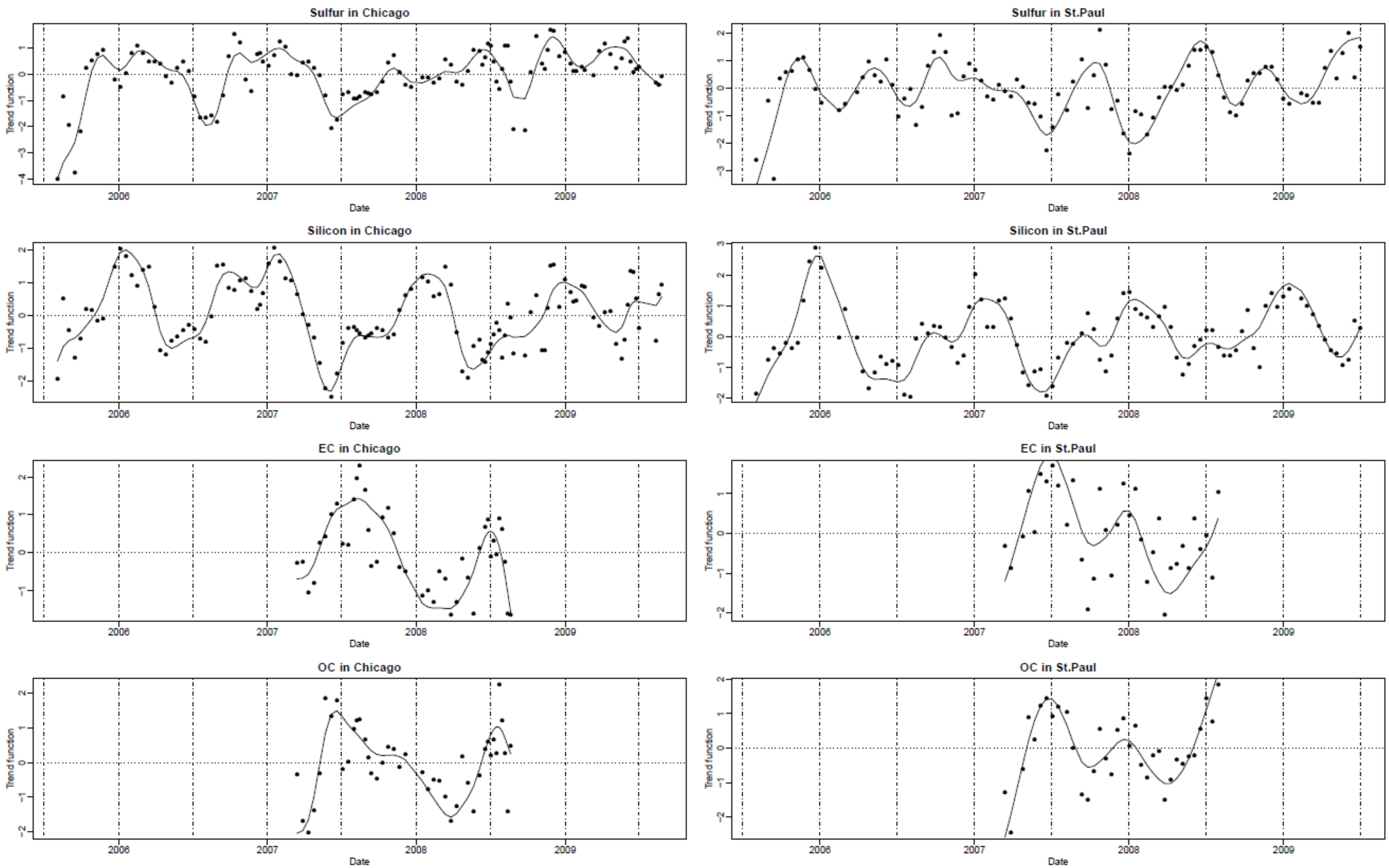
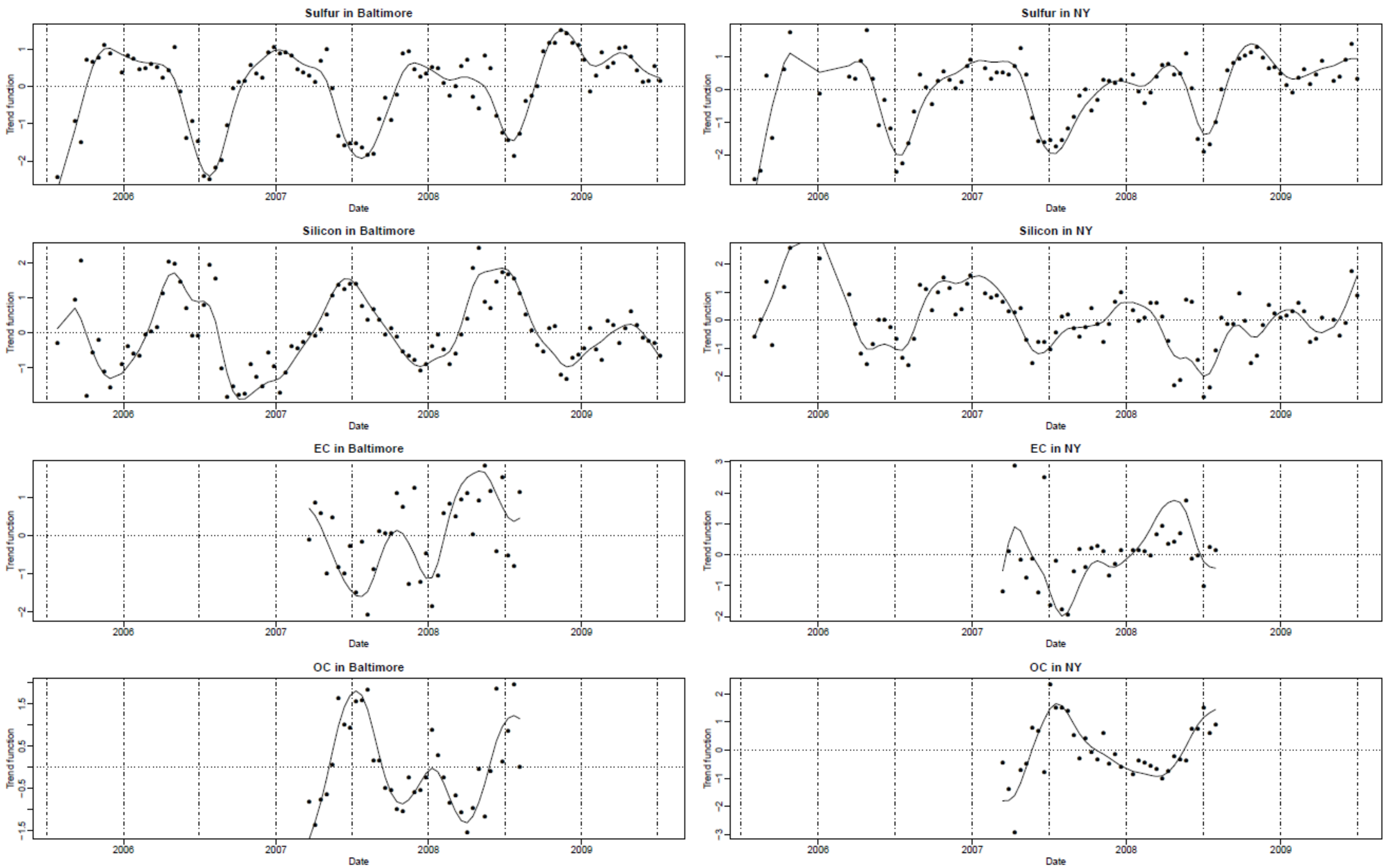



Figure S2. Temporal and spatial sampling design for silicon and EC for regulatory and NPACT/MESA Air monitors in Los Angeles


Sulfur in Chicago
Sulfur in St.Paul
Silicon in Chicago
Silicon in St.Paul
EC in Chicago
EC in St.Paul
OC in Chicago
OC in St.Paul
Trend function
Date
2006
2007
2008
2009


Sulfur in Baltimore
Sulfur in NY
Silicon in Baltimore
Silicon in NY
EC in Baltimore
EC in NY
OC in Baltimore
OC in NY
Trend function
Date
2006
2007
2008
2009

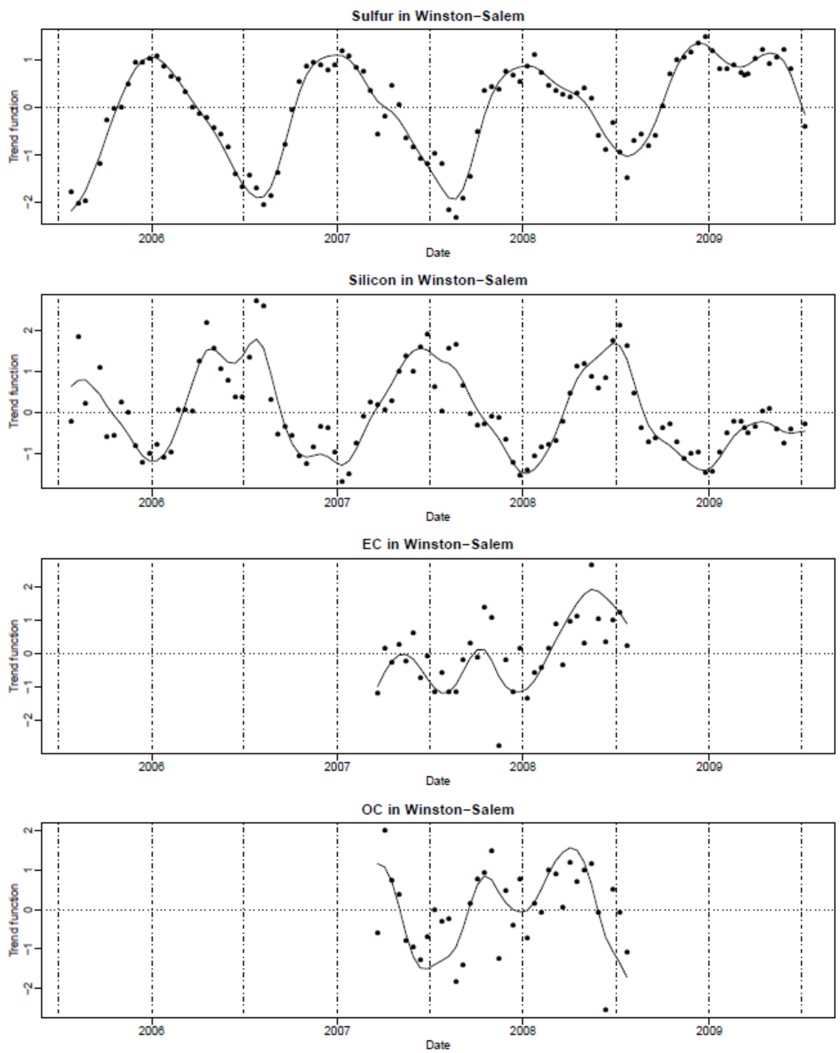
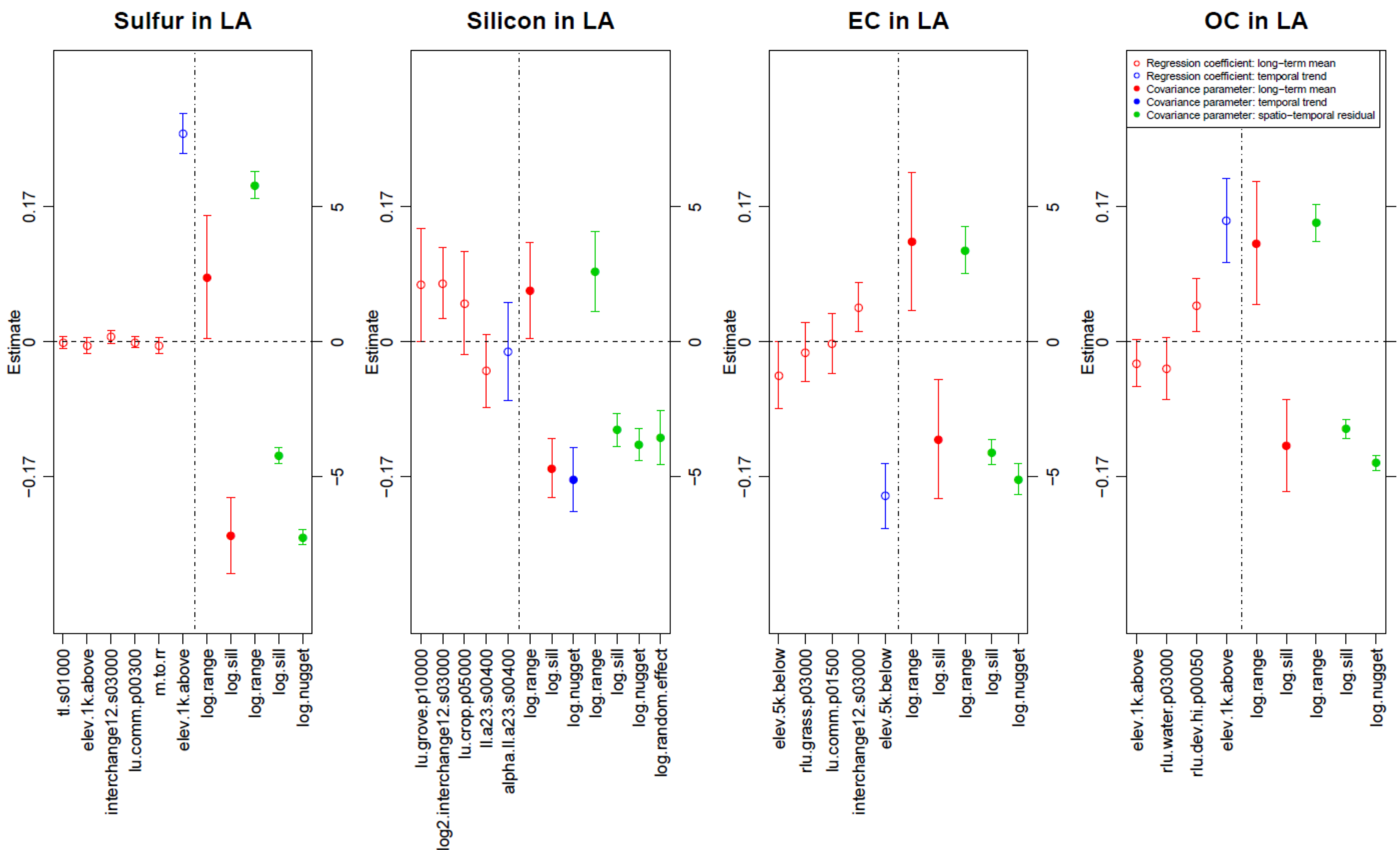
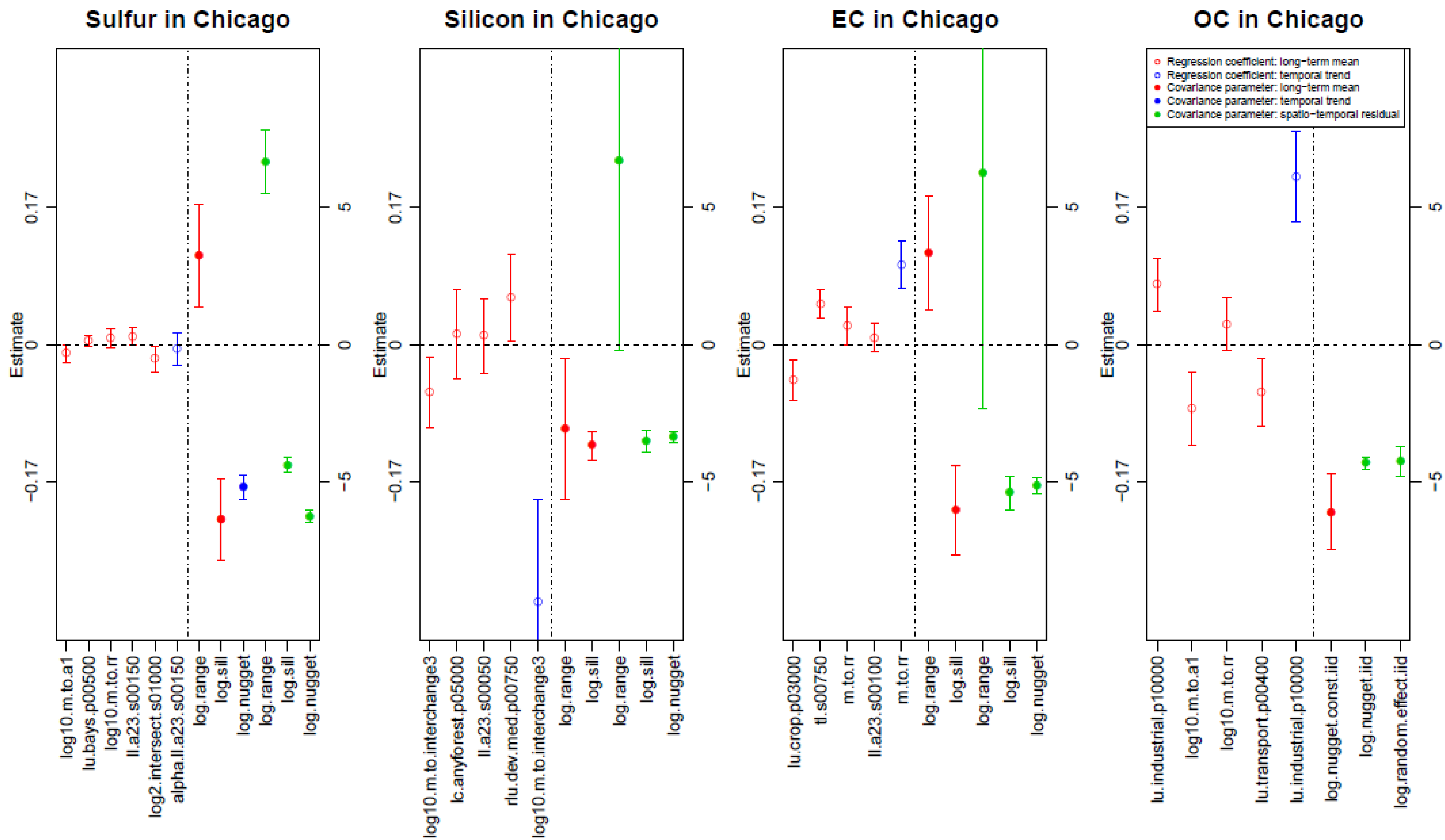
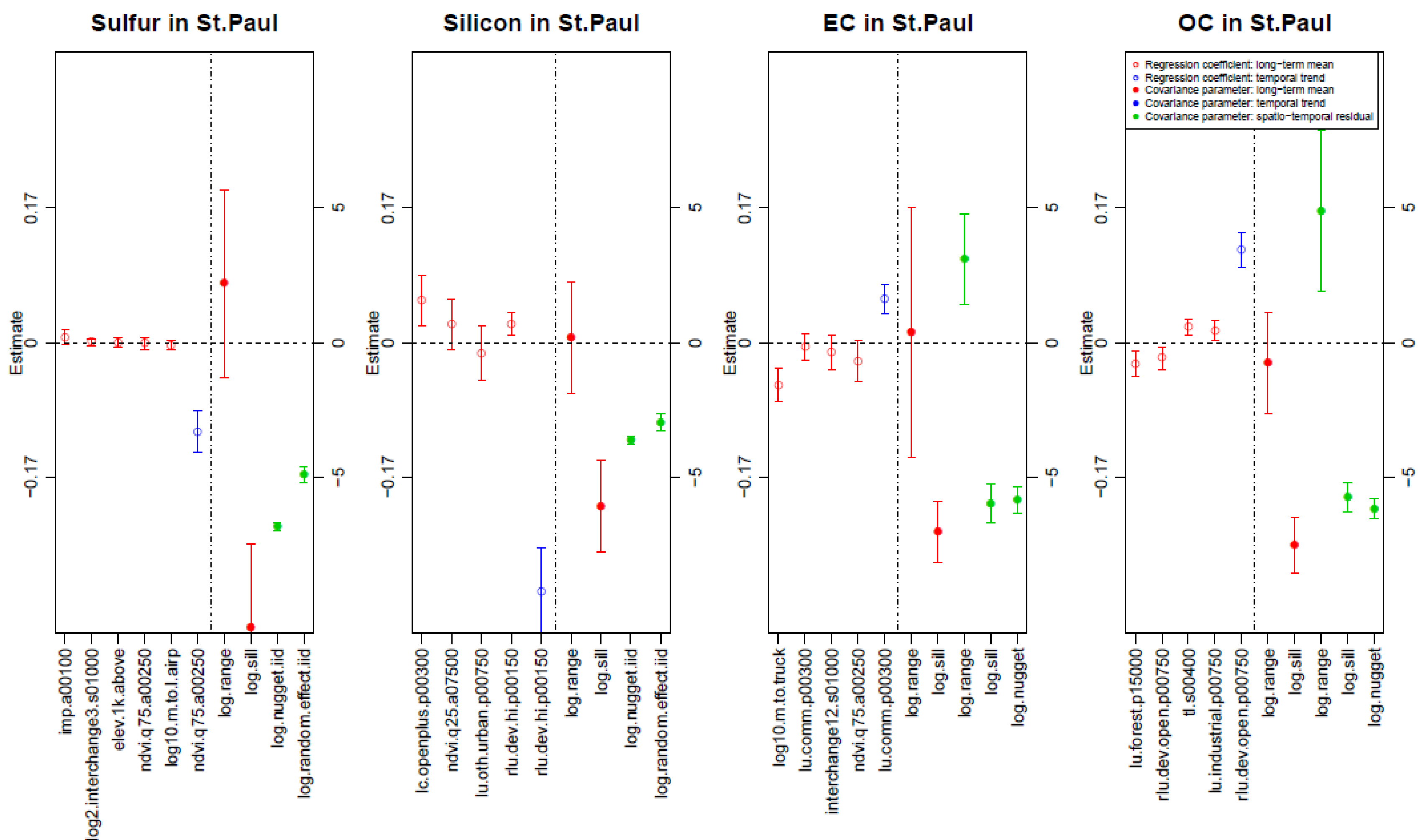
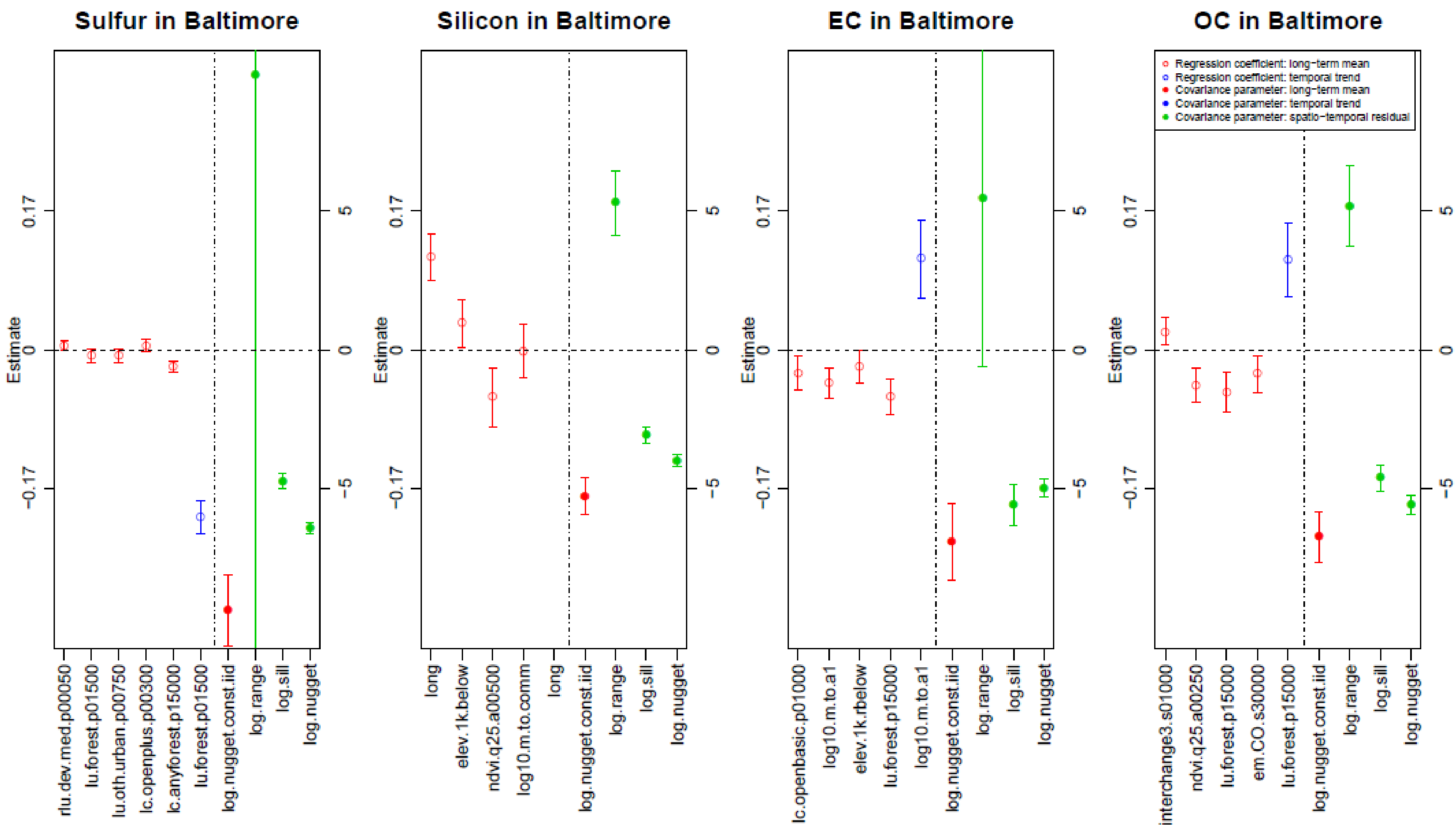


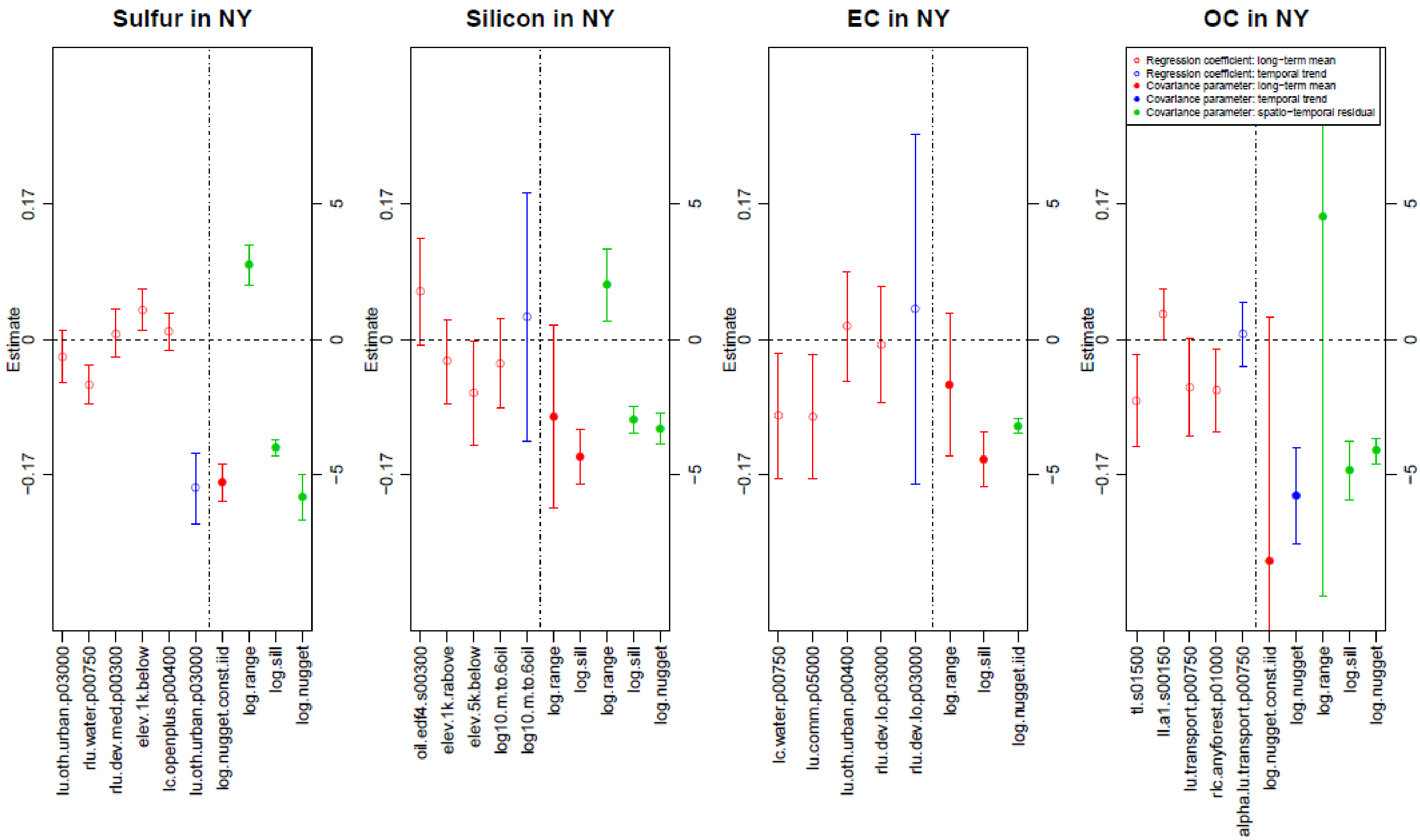

Figure S3. Estimated smooth temporal trends for four log-transformed $PM_{2.5}$ components in the MESA five city areas (except Los Angeles)


Sulfur in LA
Silicon in LA
EC in LA
OC in LA
Estimate
0.17
0
−0.17
5
0
−5
Regression coefficient: long−term mean
Regression coefficient: temporal trend
Covariance parameter: long−term mean
Covariance parameter: temporal trend
Covariance parameter: spatio−temporal residual
tl.s01000
elev.1k.above
interchange12.s03000
lu.comm.p00300
m.to.rr
elev.1k.above
log.range
log.sill
log.range
log.sill
log.nugget
lu.grove.p10000
log2.interchange12.s03000
lu.crop.p05000
ll.a23.s00400
alpha.ll.a23.s00400
log.range
log.sill
log.nugget
log.range
log.sill
log.nugget
log.random.effect
elev.5k.below
rlu.grass.p03000
lu.comm.p01500
interchange12.s03000
elev.5k.below
log.range
log.sill
log.range
log.sill
log.nugget
elev.1k.above
rlu.water.p03000
rlu.dev.hi.p00050
elev.1k.above
log.range
log.sill
log.range
log.sill
log.nugget


Sulfur in Chicago
Silicon in Chicago
EC in Chicago
OC in Chicago
Estimate
0.17
0
-0.17
5
0
-5
log10.m.to.a1
lu.bays.p00500
log10.m.to.rr
ll.a23.s00150
log2.intersect.s01000
alpha.ll.a23.s00150
log.range
log.sill
log.nugget
log.range
log.sill
log.nugget
log10.m.to.interchange3
lc.anyforest.p05000
ll.a23.s00050
rlu.dev.med.p00750
log10.m.to.interchange3
log.range
log.sill
log.range
log.sill
log.nugget
lu.crop.p03000
tl.s00750
m.to.rr
ll.a23.s00100
m.to.rr
log.range
log.sill
log.range
log.sill
log.nugget
lu.industrial.p10000
log10.m.to.a1
log10.m.to.rr
lu.transport.p00400
lu.industrial.p10000
log.nugget.const.iid
log.nugget.iid
log.random.effect.iid
Regression coefficient: long−term mean
Regression coefficient: temporal trend
Covariance parameter: long−term mean
Covariance parameter: temporal trend
Covariance parameter: spatio−temporal residual


Sulfur in St.Paul
Silicon in St.Paul
EC in St.Paul
OC in St.Paul
Estimate
0.17
0
-0.17
5
0
-5
imp.a00100
log2.interchange3.s01000
elev.1k.above
ndvi.q75.a00250
log10.m.to.l.airp
ndvi.q75.a00250
log.range
log.sill
log.nugget.iid
log.random.effect.iid
lc.openplus.p00300
ndvi.q25.a07500
lu.oth.urban.p00750
rlu.dev.hi.p00150
rlu.dev.hi.p00150
log10.m.to.truck
lu.comm.p00300
interchange12.s01000
ndvi.q75.a00250
lu.comm.p00300
log.nugget
lu.forest.p15000
rlu.dev.open.p00750
tl.s00400
lu.industrial.p00750
rlu.dev.open.p00750
Regression coefficient: long−term mean
Regression coefficient: temporal trend
Covariance parameter: long−term mean
Covariance parameter: temporal trend
Covariance parameter: spatio−temporal residual


Sulfur in Baltimore
Silicon in Baltimore
EC in Baltimore
OC in Baltimore
Estimate
0.17
0
−0.17
5
0
−5
Regression coefficient: long−term mean
Regression coefficient: temporal trend
Covariance parameter: long−term mean
Covariance parameter: temporal trend
Covariance parameter: spatio−temporal residual
rlu.dev.med.p00050
lu.forest.p01500
lu.oth.urban.p00750
lc.openplus.p00300
lc.anyforest.p15000
lu.forest.p01500
log.nugget.const.iid
log.range
log.sill
log.nugget
long
elev.1k.below
ndvi.q25.a00500
log10.m.to.comm
long
log.nugget.const.iid
log.range
log.sill
log.nugget
lc.openbasic.p01000
log10.m.to.a1
elev.1k.rbelow
lu.forest.p15000
log10.m.to.a1
log.nugget.const.iid
log.range
log.sill
log.nugget
interchange3.s01000
ndvi.q25.a00250
lu.forest.p15000
em.CO.s30000
lu.forest.p15000
log.nugget.const.iid
log.range
log.sill
log.nugget



Sulfur in NY
Silicon in NY
EC in NY
OC in NY
Estimate
0.17
0
−0.17
5
0
−5
Regression coefficient: long−term mean
Regression coefficient: temporal trend
Covariance parameter: long−term mean
Covariance parameter: temporal trend
Covariance parameter: spatio−temporal residual
lu.oth.urban.p03000
rlu.water.p00750
rlu.dev.med.p00300
elev.1k.below
lc.openplus.p00400
lu.oth.urban.p03000
log.nugget.const.iid
log.range
log.sill
log.nugget
oil.edf4.s00300
elev.1k.rabove
elev.5k.below
log10.m.to.6oil
log10.m.to.6oil
log.range
log.sill
log.range
log.sill
log.nugget
lc.water.p00750
lu.comm.p05000
lu.oth.urban.p00400
rlu.dev.lo.p03000
rlu.dev.lo.p03000
log.range
log.sill
log.nugget.iid
tl.s01500
ll.a1.s00150
lu.transport.p00750
rlc.anyforest.p01000
alpha.lu.transport.p00750
log.nugget.const.iid
log.nugget
log.range
log.sill
log.nugget

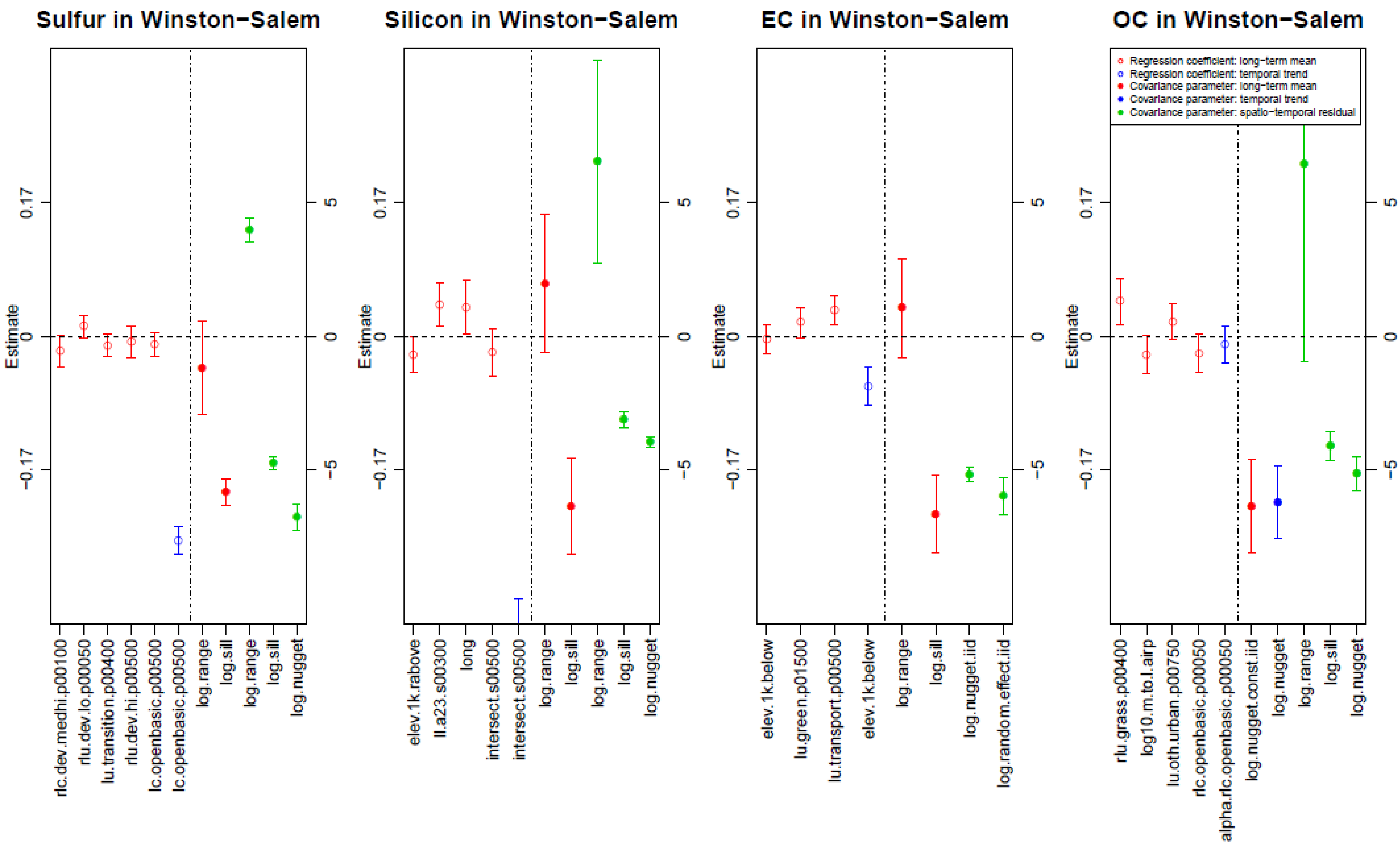
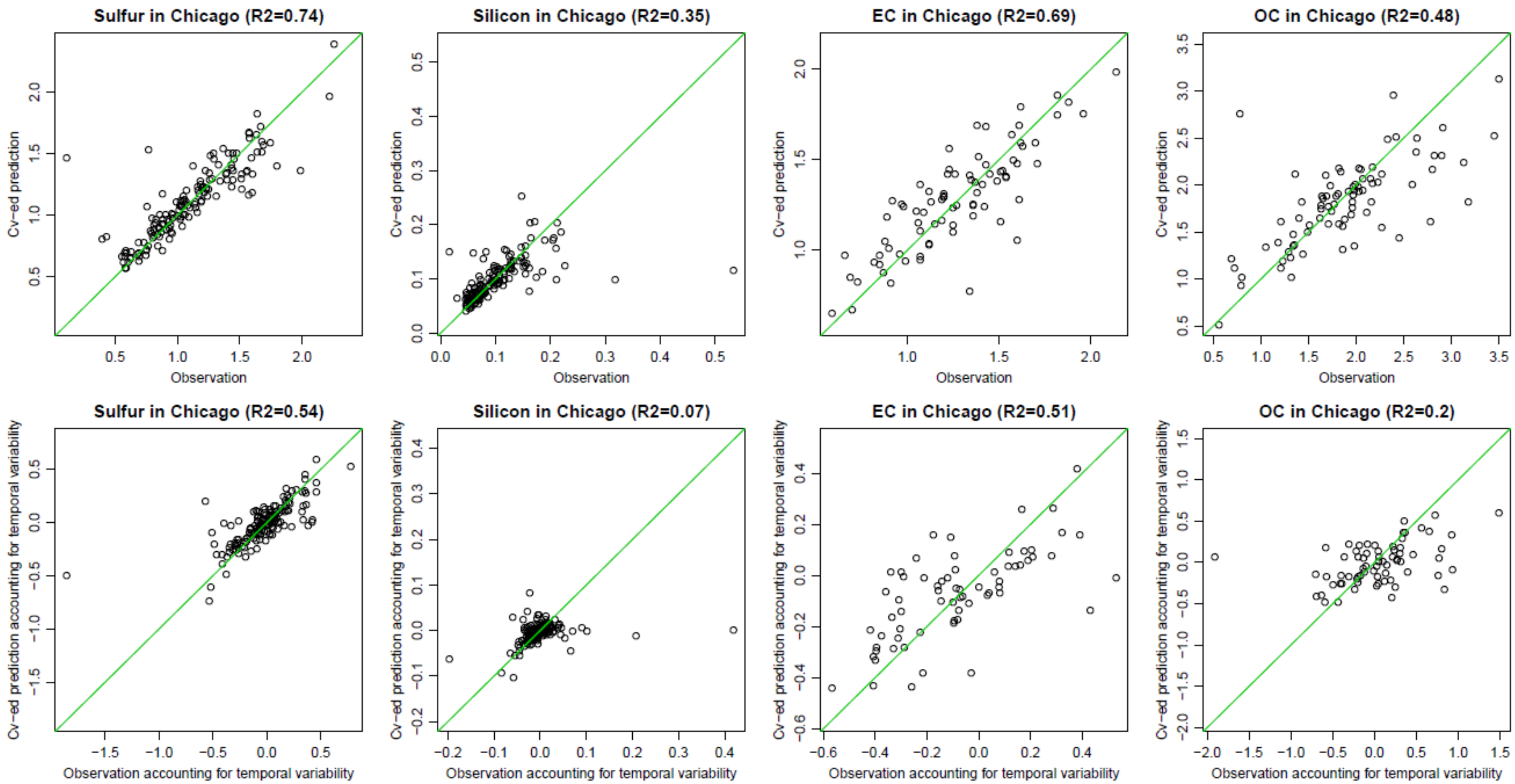
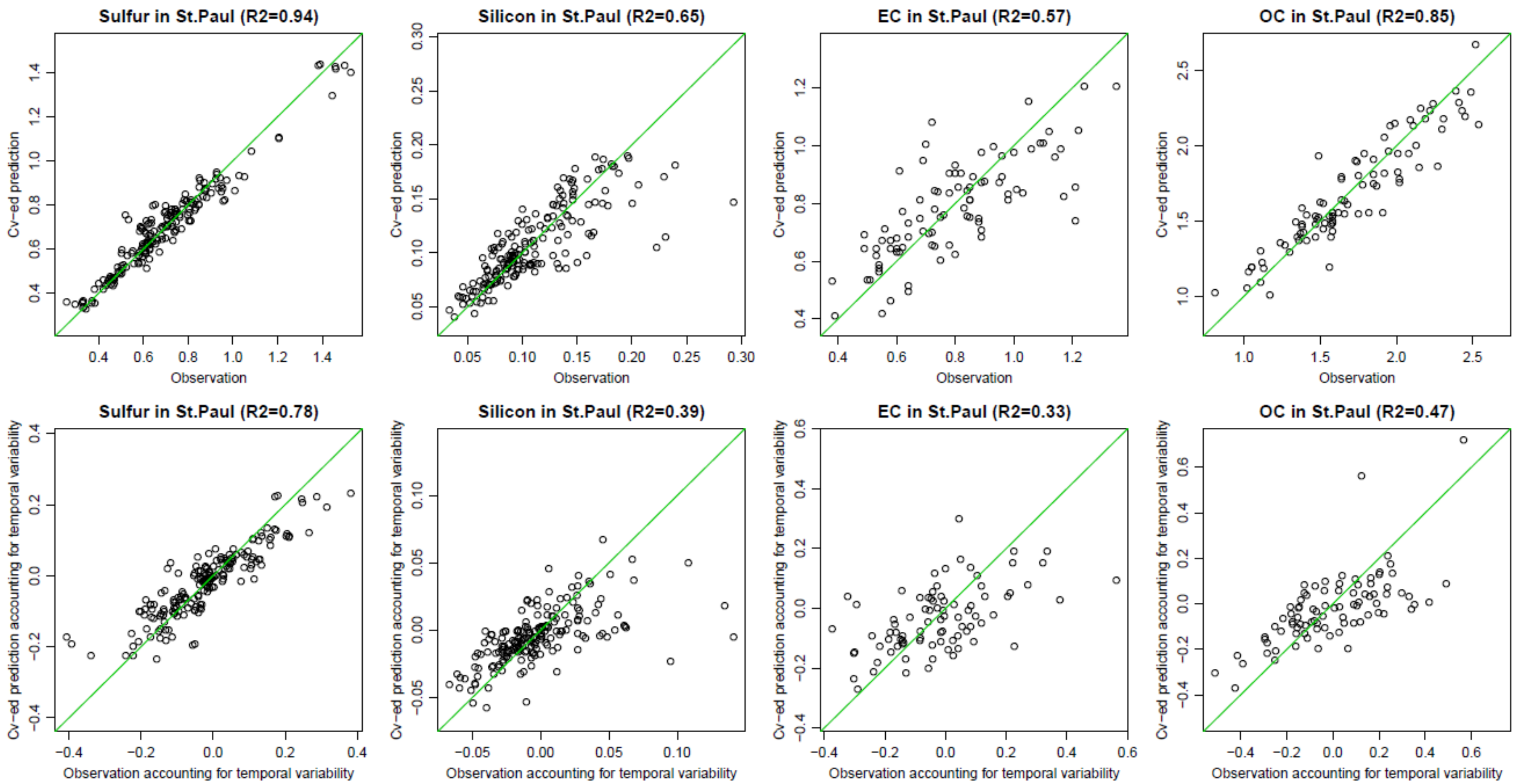
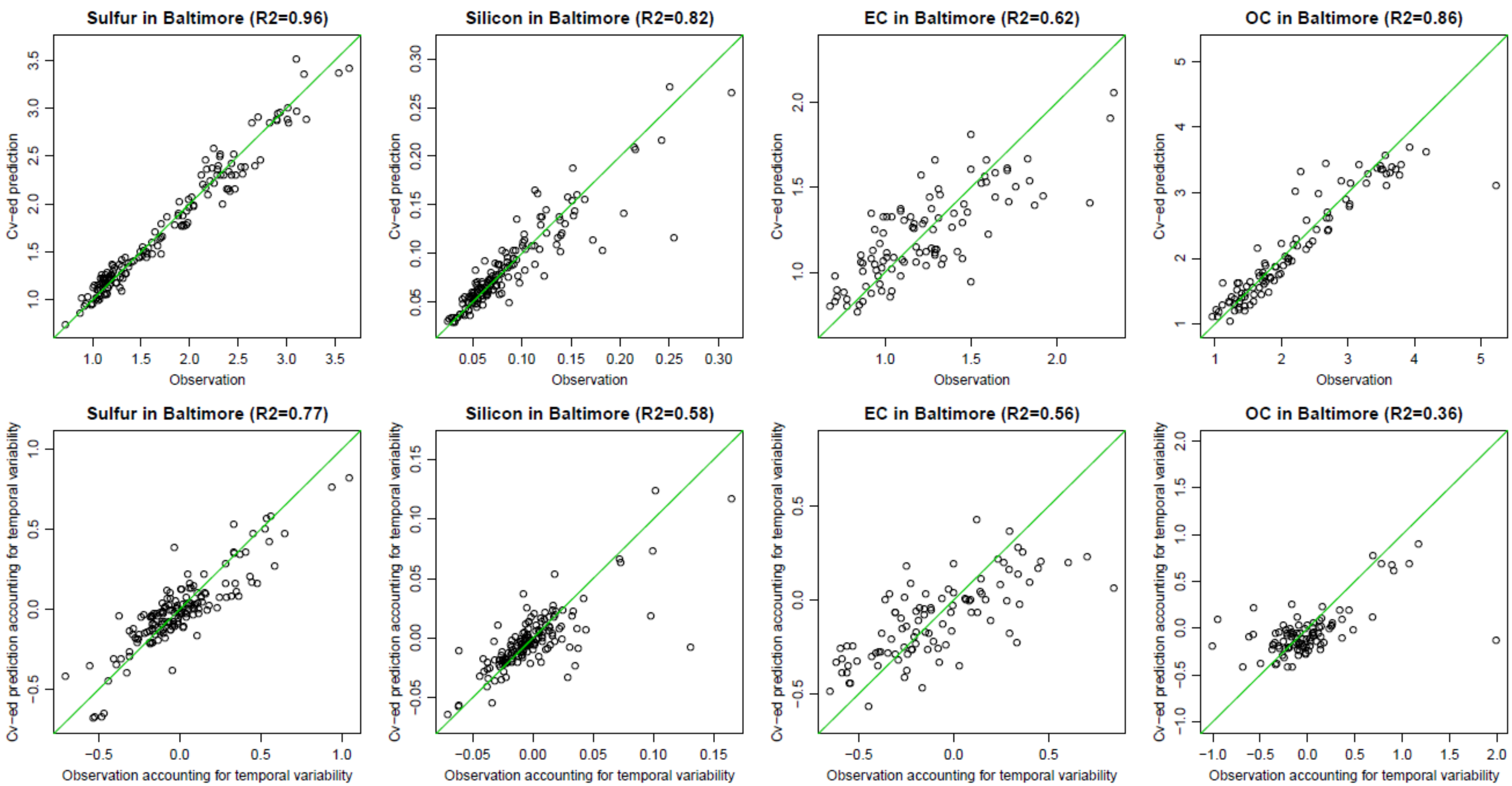
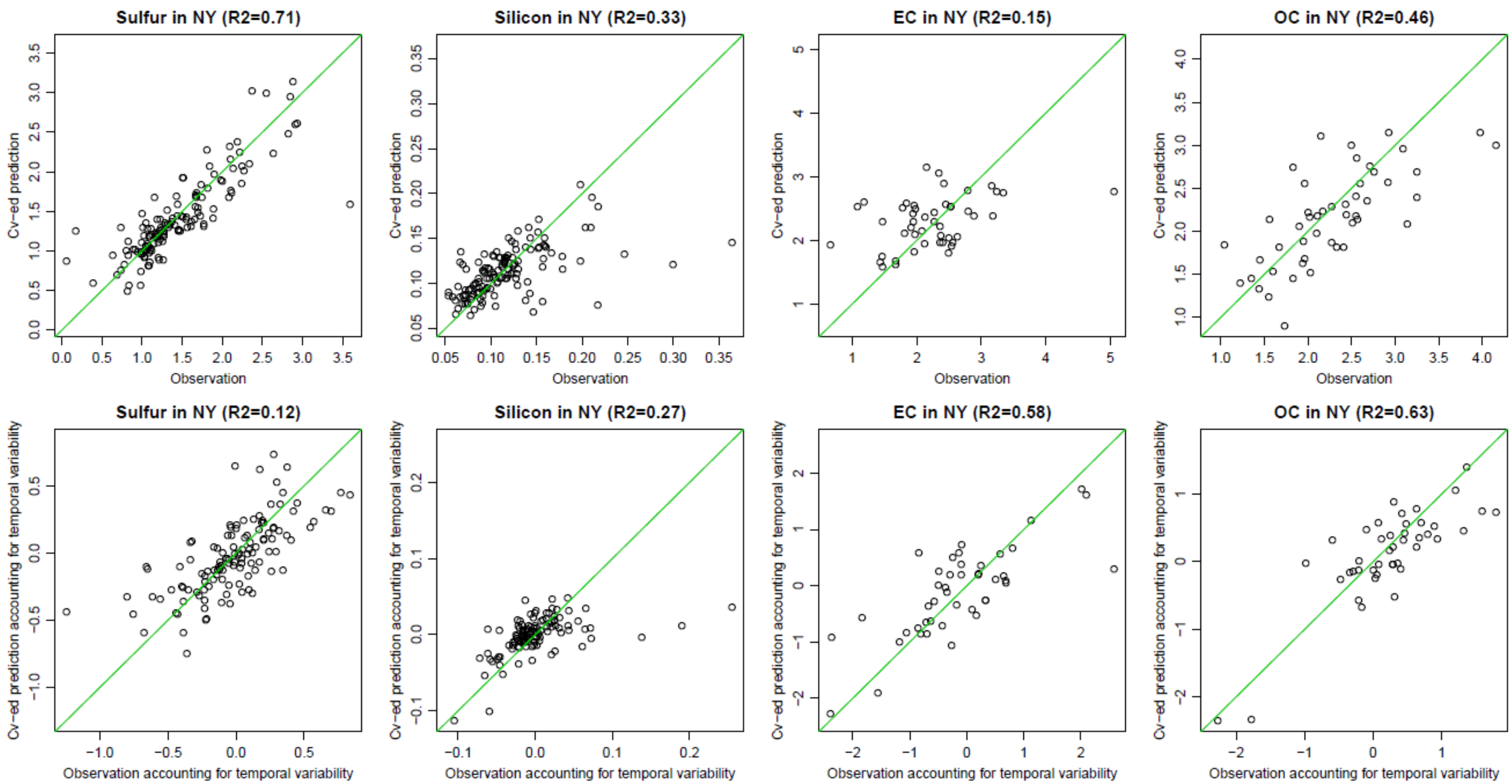



Figure S4. Estimated parameters for the selected geographical variables (scaled) and covariance structure in the spatio-temporal model for the four log-transformed $PM_{2.5}$ components in the six MESA cities


Sulfur in Chicago (R2=0.74)
Silicon in Chicago (R2=0.35)
EC in Chicago (R2=0.69)
OC in Chicago (R2=0.48)
Cv−ed prediction
Observation
Sulfur in Chicago (R2=0.54)
Silicon in Chicago (R2=0.07)
EC in Chicago (R2=0.51)
OC in Chicago (R2=0.2)
Cv−ed prediction accounting for temporal variability
Observation accounting for temporal variability


Sulfur in St.Paul (R2=0.94)
Silicon in St.Paul (R2=0.65)
EC in St.Paul (R2=0.57)
OC in St.Paul (R2=0.85)
Cv−ed prediction
Observation
Sulfur in St.Paul (R2=0.78)
Silicon in St.Paul (R2=0.39)
EC in St.Paul (R2=0.33)
OC in St.Paul (R2=0.47)
Cv−ed prediction accounting for temporal variability
Observation accounting for temporal variability


Sulfur in Baltimore (R2=0.96)
Silicon in Baltimore (R2=0.82)
EC in Baltimore (R2=0.62)
OC in Baltimore (R2=0.86)
Cv−ed prediction
Observation
Sulfur in Baltimore (R2=0.77)
Silicon in Baltimore (R2=0.58)
EC in Baltimore (R2=0.56)
OC in Baltimore (R2=0.36)
Cv−ed prediction accounting for temporal variability
Observation accounting for temporal variability


Sulfur in NY (R2=0.71)
Silicon in NY (R2=0.33)
EC in NY (R2=0.15)
OC in NY (R2=0.46)
Cv−ed prediction
Observation
Sulfur in NY (R2=0.12)
Silicon in NY (R2=0.27)
EC in NY (R2=0.58)
OC in NY (R2=0.63)
Cv−ed prediction accounting for temporal variability
Observation accounting for temporal variability

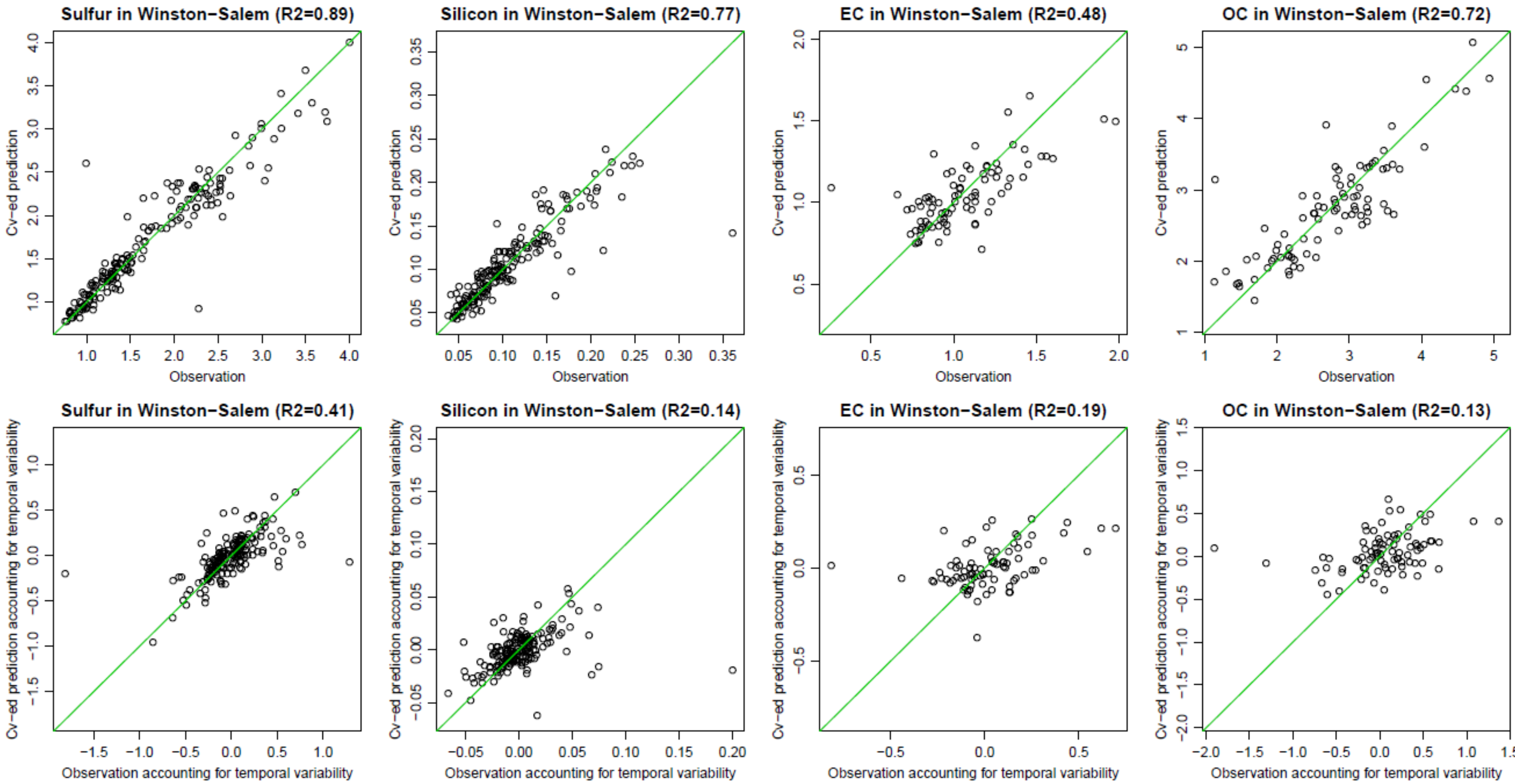



Figure S5. City-specific scatter plots of observations and cross-validated predictions from the spatio-temporal model for 2-week average concentrations (top) and for 2-week average concentrations accounting for temporal variability (bottom) across home-outdoor sites in five MESA city areas (except Los Angeles)

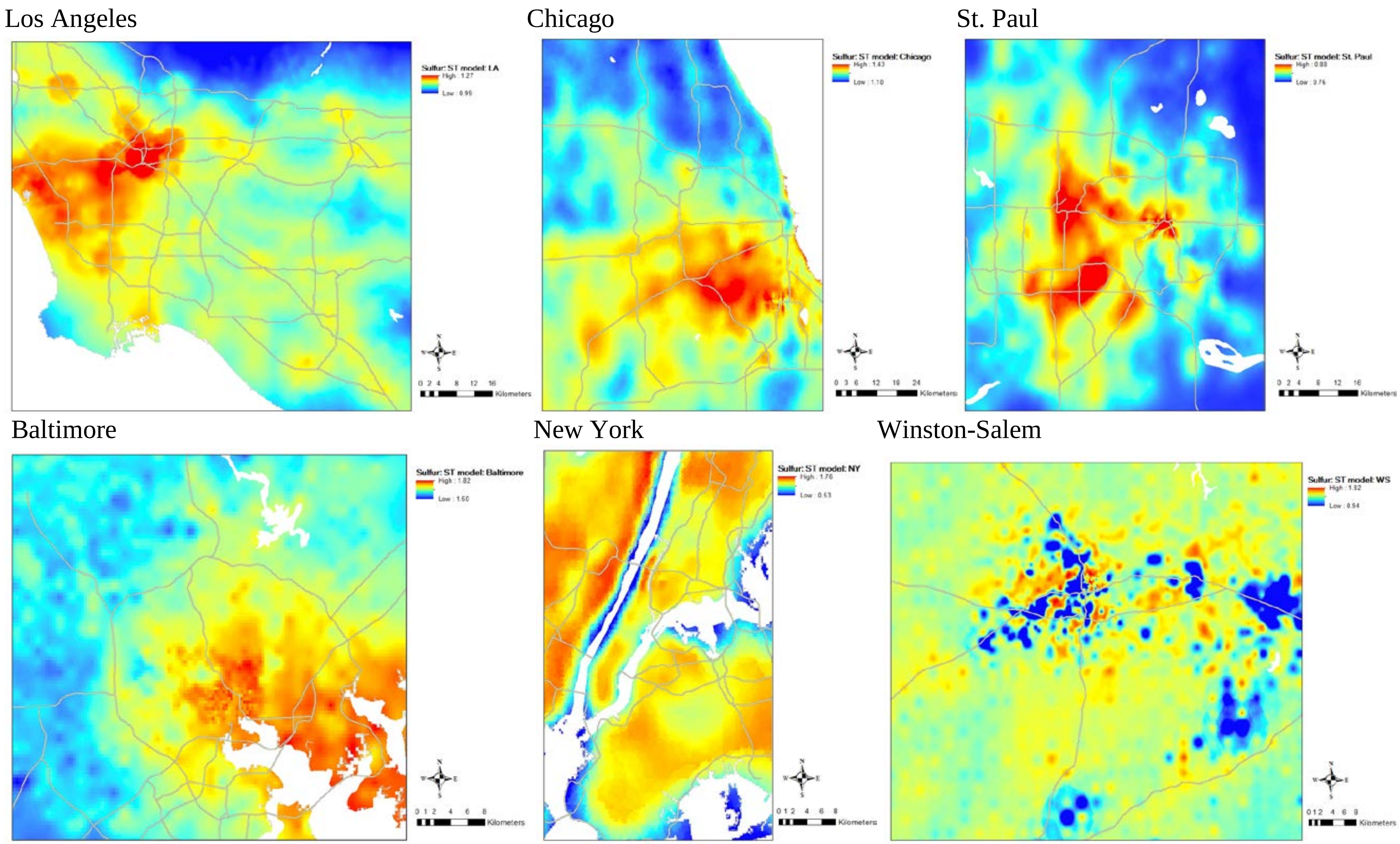



Figure S6. Maps of predicted long-term concentrations of sulfur (μg/m$^3$) from the spatio-temporal model in the six MESA city areas

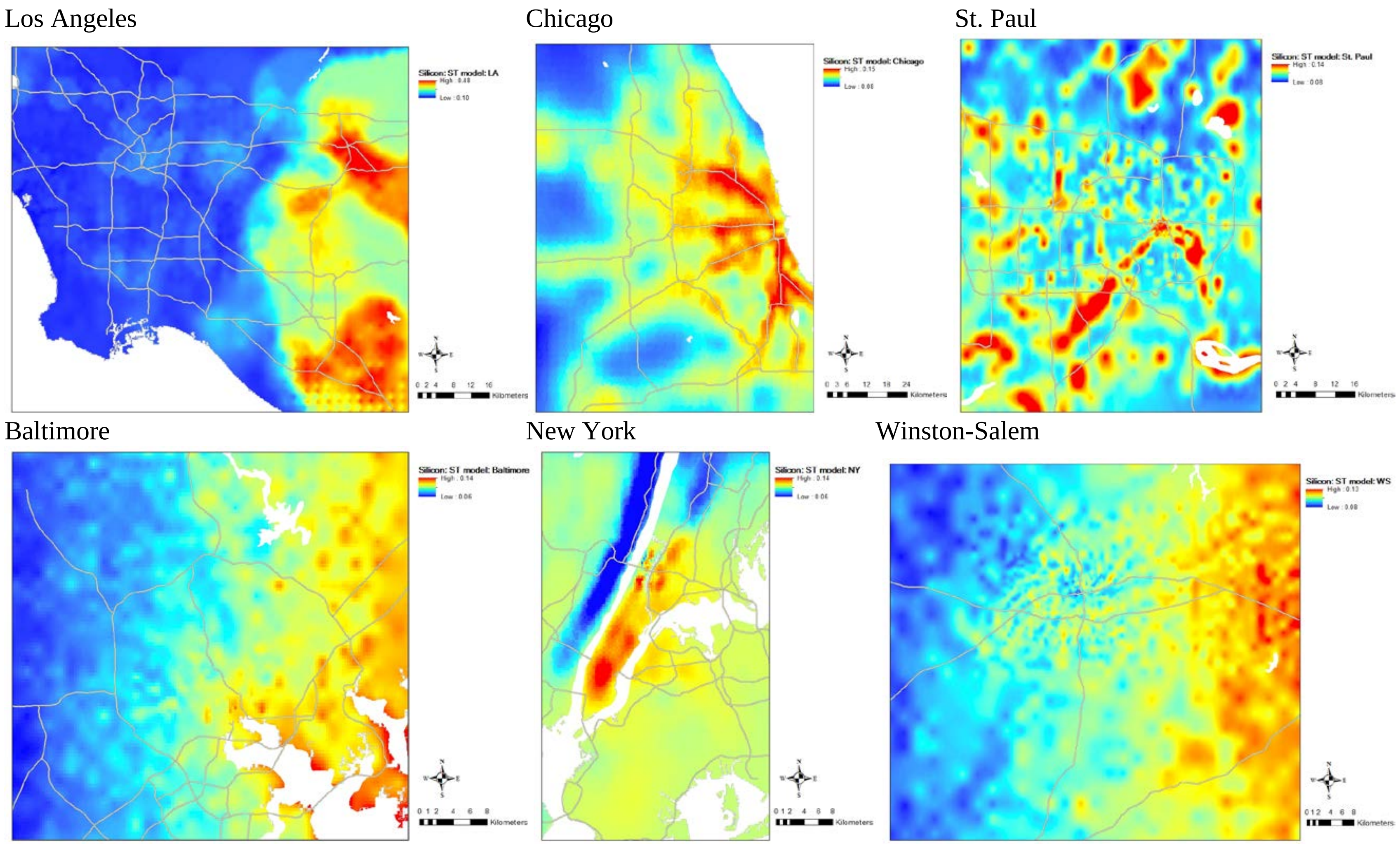



Figure S7. Maps of predicted long-term concentrations of silicon ($\mu g/m^3$) from the spatio-temporal model in the six MESA city areas

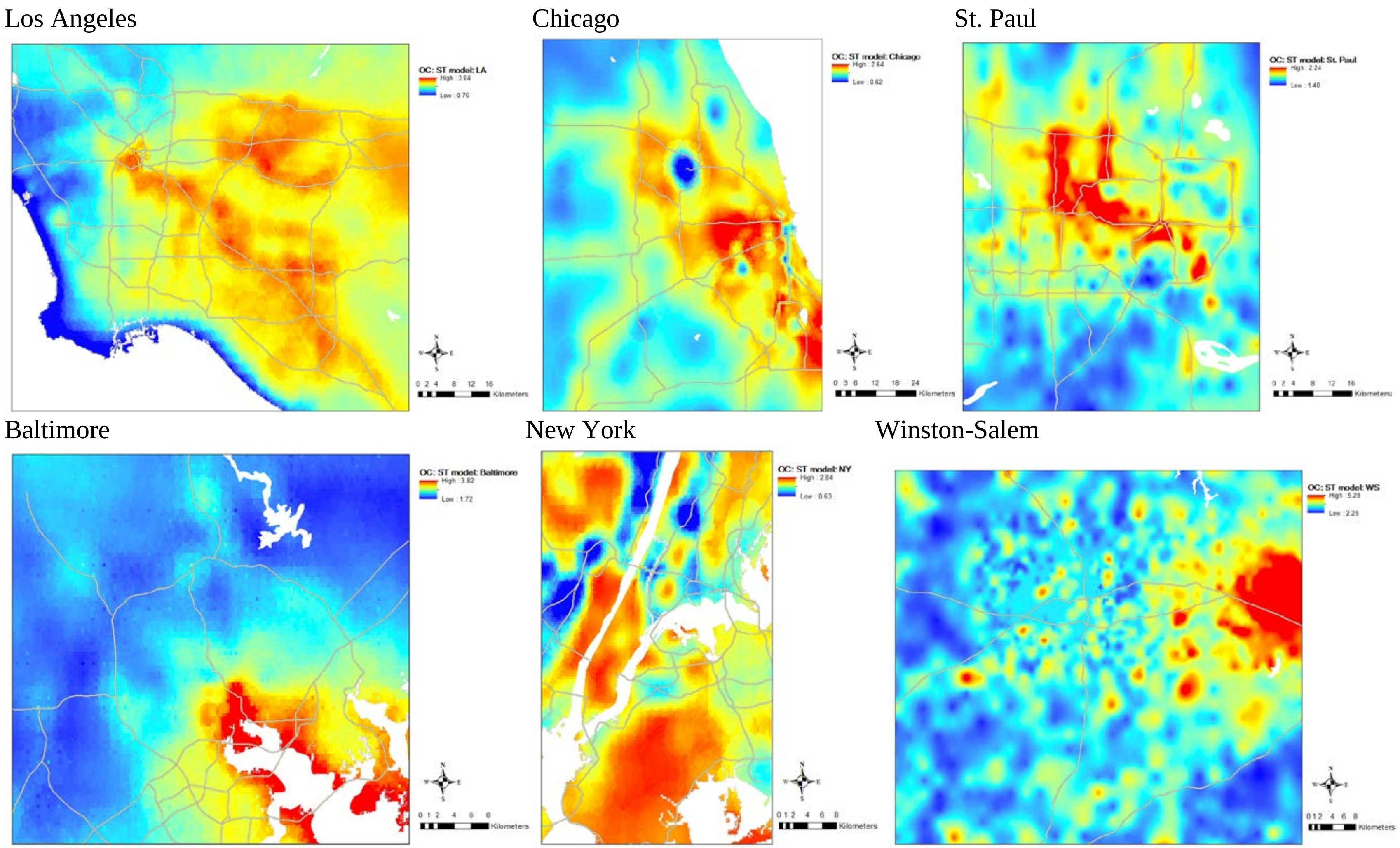
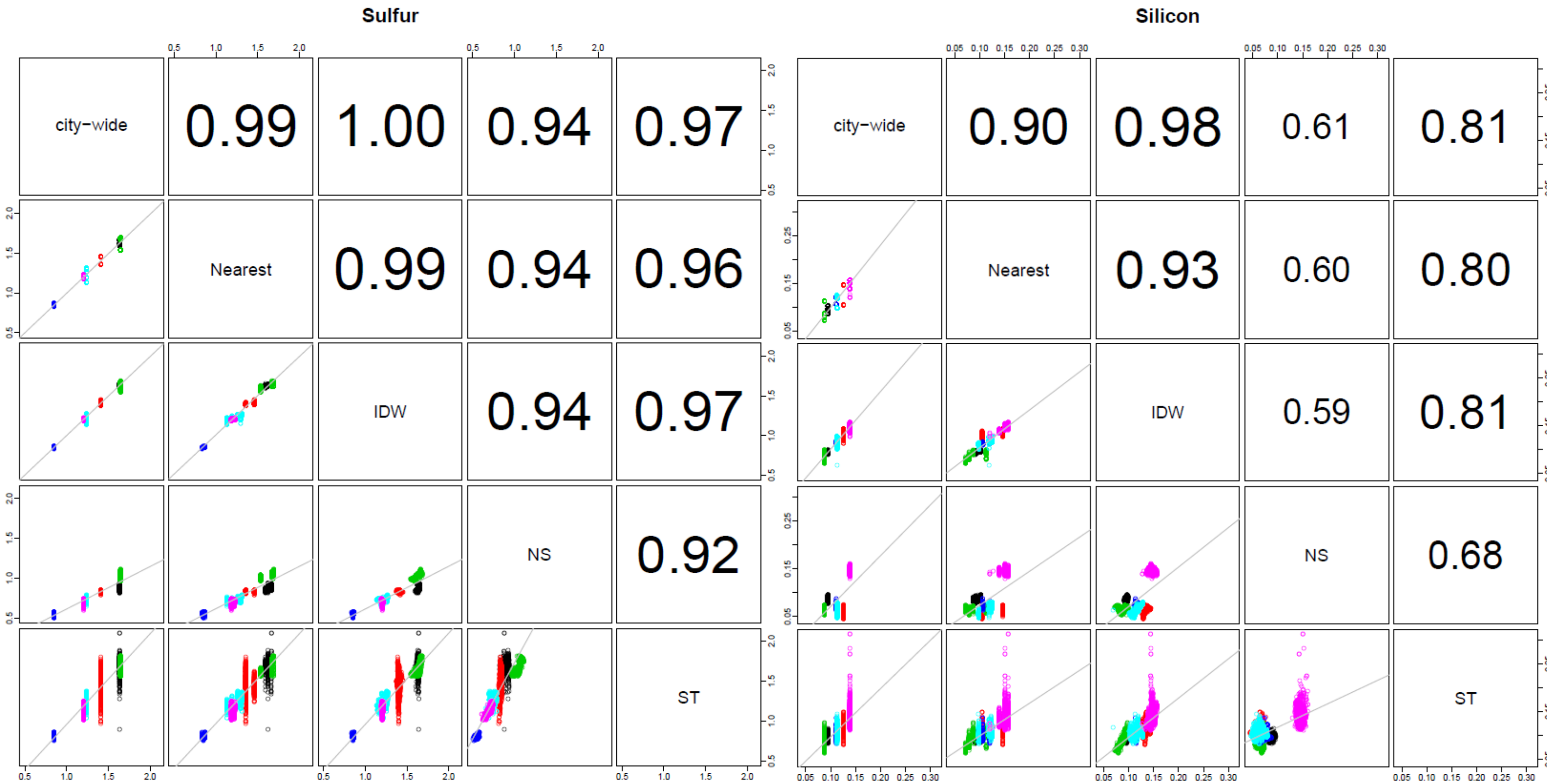



Figure S8. Maps of predicted long-term concentrations of OC (μg/m$^3$) from the spatio-temporal model in the six MESA city areas


Sulfur
city−wide
0.99
1.00
0.94
0.97
Nearest
0.99
0.94
0.96
IDW
0.94
0.97
NS
0.92
ST
Silicon
city−wide
0.90
0.98
0.61
0.81
Nearest
0.93
0.60
0.80
IDW
0.59
0.81
NS
0.68
ST

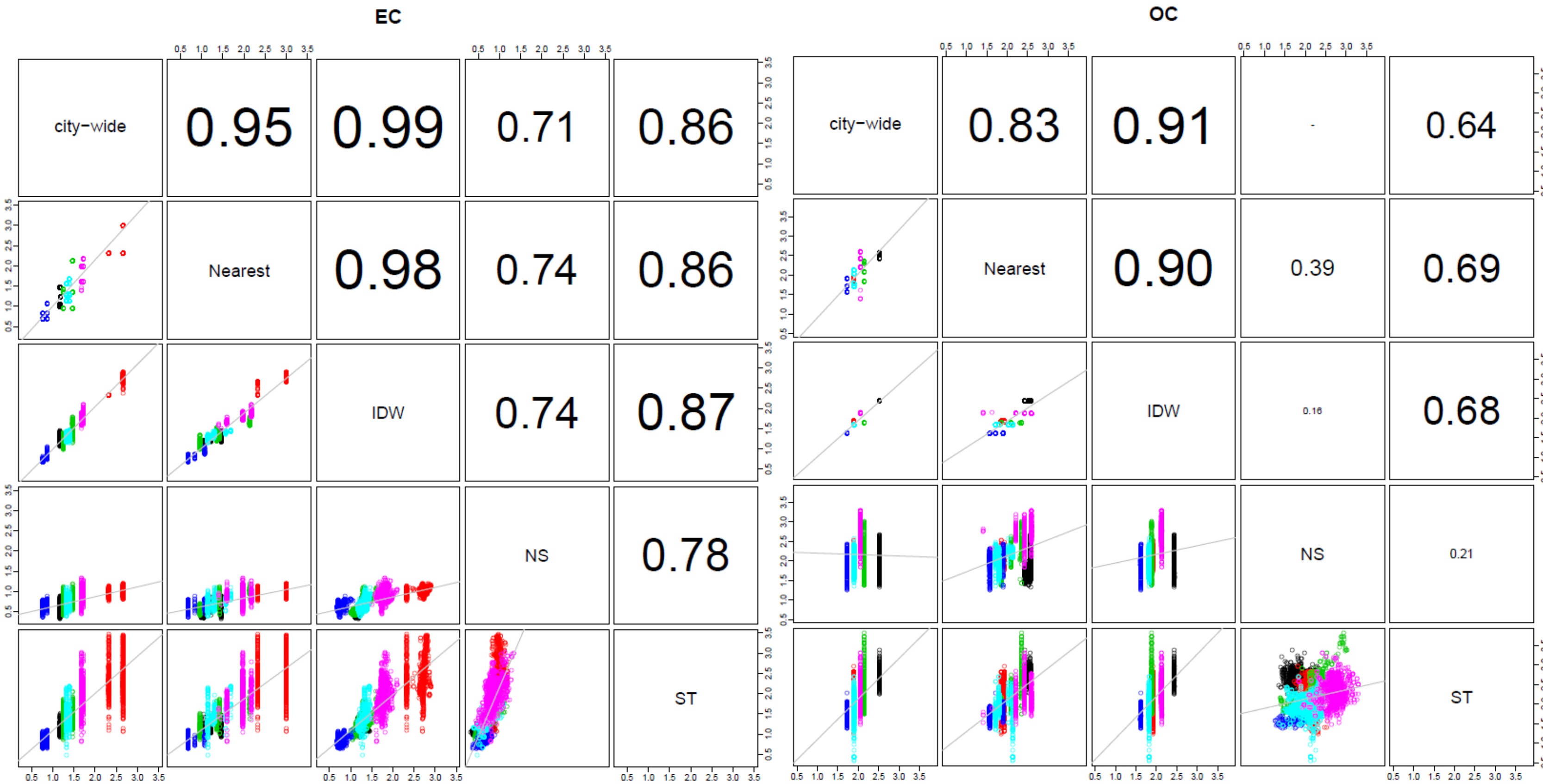



Figure S9. Scatter plots and correlation coefficients of predicted long-term concentrations of four $PM_{2.5}$ components between the five different prediction models (city-wide, nearest monitor, IDW (inverse distance weighting), NS (national spatial model), and ST (spatio-temporal model)) for MESA participant addresses within 10 kilometers of any NPACT/MESA Air monitors in six MESA city areas (color code: black = Winston-Salem, red = NY, green = Baltimore, blue = St. Paul, light blue = Chicago, and pink = LA).

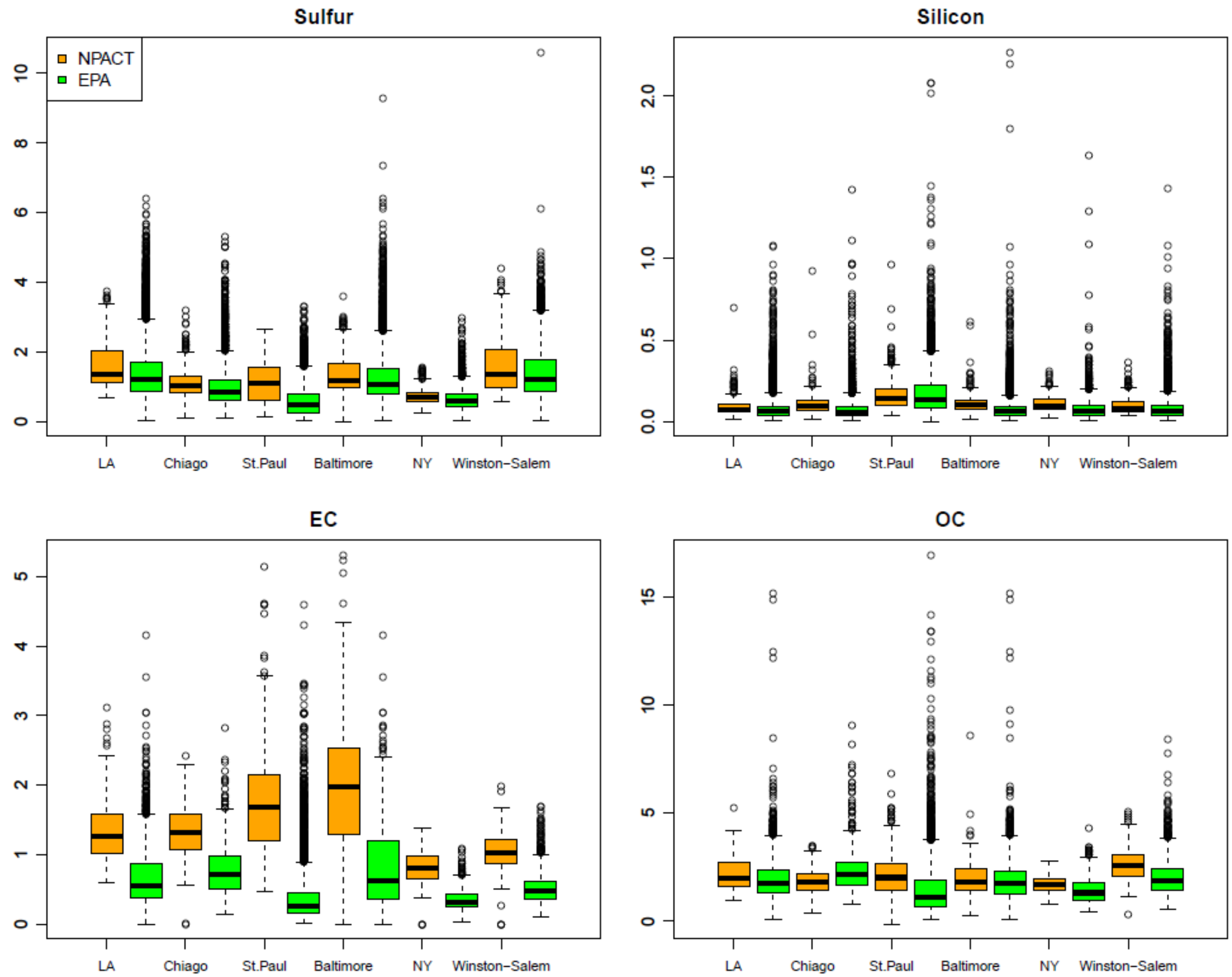



Figure S10. Box plots of measurements for four $PM_{2.5}$ components by NPACT/MESA Air and CSN/IMPROVE monitoring campaign by the six MESA city areas defined by 200 kilometers within the centers of six MESA cities (2 week samples for NPACT/MESA Air and daily samples on every 3rd and 6th day schedule for CSN/IMPROVE)