



UW Biostatistics Working Paper Series

6-27-2008

Influence of prediction approaches for spatially-dependent air pollution exposure on health effect estimation

Sun-Young Kim

University of Washington, puha0@u.washington.edu

Lianne Sheppard

University of Washington, sheppard@u.washington.edu

Ho Kim

Seoul National University, hokim@snu.ac.kr

Suggested Citation

Kim, Sun-Young; Sheppard, Lianne; and Kim, Ho, "Influence of prediction approaches for spatially-dependent air pollution exposure on health effect estimation" (June 2008). *UW Biostatistics Working Paper Series*. Working Paper 331. <http://biostats.bepress.com/uwbiostat/paper331>

This working paper is hosted by The Berkeley Electronic Press (bepress) and may not be commercially reproduced without the permission of the copyright holder.

Copyright © 2011 by the authors

Influence of prediction approaches for spatially-dependent air pollution exposure on health effect estimation

Sun-Young Kim (1), Lianne Sheppard (2)(1), Ho Kim (3)

(1) Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, USA

(2) Department of Biostatistics, University of Washington, Seattle, USA

(3) Department of Biostatistics and Epidemiology, Seoul National University, Seoul, Korea



ABSTRACT

Background: Air pollution studies increasingly estimate individual-level exposures from area-based measurements by using exposure prediction methods such as nearest monitor and kriging predictions. However, little is known about the properties of these methods for health effects estimation. This simulation study explores how two common prediction approaches for fine particulate matter (PM_{2.5}) affect relative risk estimates for cardiovascular events in a single geographic area.

Methods: We estimated two sets of parameters to define correlation structures from 2002 PM_{2.5} data in the Los Angeles (LA) area and selected additional parameters to evaluate different correlation features. For each structure, annual average PM_{2.5} was generated at 22 existing monitoring sites and 2,000 pre-selected individual locations in LA.

Associated survival time until cardiovascular event was simulated for 10,000 hypothetical subjects. Using PM_{2.5} generated at monitoring sites, we predicted PM_{2.5} at subject locations by nearest monitor and kriging interpolation. Finally, relative risks (RRs) of the effect of PM_{2.5} on time to cardiovascular event were estimated.

Results: Health effect estimates for cardiovascular events had higher or similar coverage probability for kriging compared to nearest monitor exposures. The lower mean square error of nearest monitor prediction resulted from more precise but biased health effect estimates. The difference between these approaches dramatically moderated when spatial correlation increased and geographical characteristics were included in the mean model.

Conclusions: When the underlying exposure distribution has a large amount of spatial dependence, both kriging and nearest monitor predictions gave good health effect estimates. For exposure with little spatial dependence, kriging exposure was preferable

but gave very uncertain estimates.



1. INTRODUCTION

Studies of health effects of long-term air pollution have traditionally assigned exposures to individuals using measurements made at a few fixed site monitoring stations.^{1,2} This approach was based on the assumption that the spatially-averaged ambient air pollution concentrations measured at monitoring sites in an administrative area, such as a community or a metropolitan area, is representative of the exposure of individuals residing at different locations in the area. However, such exposure data do not capture all the individual spatial heterogeneity in exposure and may result in less accurate or reliable health effect estimates.

In order to improve exposure assessment, some more recent studies used predicted individual air pollution exposure instead of an average area-wide monitored concentration. One prediction approach assigns exposure based on nearest monitor to the participant's residential location.³⁻⁶ Another common approach is to interpolate predictions by applying a geo-statistical method such as kriging.⁷⁻¹⁰ Kriging predicts individual concentrations corresponding to residential locations after estimating the parameters of a model of the spatial surface of air pollution concentration. Although both methods have been used in practice, little is known about how these prediction methods affect health effect estimation. Using assumptions derived from an analysis of annual average concentration of particulate matter less than or equal to 2.5 microns in aerodynamic diameter (PM_{2.5}) and a previous analysis of the effect of particulate matter on cardiovascular events,⁴ we conducted a simulation study to investigate the impact on the health effect estimate of using predicted exposure from these two exposure prediction approaches.

2. METHODS

2.1. Data and assumptions

2.1.1. Population

We selected 2,000 residential locations from the Los Angeles-Long Beach-Santa-Ana urbanized area of California (LA urbanized area) as defined by the US census classification (figure 1). Random selection of locations was made using a uniform distribution stratified by census tract to obtain up to three residential locations per tract with the number of individual sites proportional to the population size. To obtain 10,000 subjects, five subjects were assumed to live around each residential site, and have identical PM_{2.5} exposure.

2.1.2. PM_{2.5} data and analysis

The twenty two PM_{2.5} monitoring stations used in the simulation were located in five counties (Los Angeles, Orange, Riverside, San Bernardino, and Ventura counties) surrounding the LA urbanized area. These stations were nearly identical to those used in Jerrett et al, 2005.⁸ Using annual average concentrations from 2002, we estimated parameters for a spatial model of PM_{2.5} by fitting mean and covariance models. We fit a covariance model to the empirical variogram, the squared differences of PM_{2.5} against distance. The covariance model has a specific functional form with three parameters: the range, partial sill, and nugget. The range is the distance where the variogram curve reaches the maximum variability. The total variance, named sill, is partitioned into the partial sill and nugget. The partial sill is the component of variance due to spatial variability, while the nugget is considered measurement error. The mean model is a regression model with covariates assumed to affect PM_{2.5}. In this analysis, we only

considered longitude (X) and latitude (Y) as possible covariates. After exploration of several covariance and mean models, the best fitting model, chosen by cross-validation, was a spherical covariance model with a second-order polynomial mean model using the XY coordinates as predictors. This model is referred to as TEM 1 in the tables and figures. It is similar to the one used in previous studies in the LA area.^{8,9} For sensitivity analysis, we also estimated spatial parameters from a constant mean model (TEM 4).

2.1.3. Disease

We used an exponential relative risk (RR) model for time to first cardiovascular event and used the RR, baseline rate, and drop-out patterns consistent with a previous study of postmenopausal women in the Women's Health Initiative (WHI) cohort.⁴ Miller et al reported a cardiovascular and cerebrovascular disease incidence proportion of 0.032 for a median follow-up of 6 years among 65,893 women and estimated overall hazard ratio of 1.24 for a $10 \mu\text{g}/\text{m}^3$ increase in long-term average PM2.5 concentration. In the simulation, we assumed a baseline incidence rate of 0.032 per year and a beta coefficient of 0.0215 ($=\log(1.24)/10$) for the RR parameter.

2.2. Simulation

Given the residence locations and model parameters, we generated true PM2.5 exposures for 2,000 residence and 22 monitor locations, and 10,000 outcomes (times to cardiovascular disease event) in 1,000 simulations. We made the simplifying assumption that PM2.5 concentration equals PM2.5 exposure. Using only monitoring data, we then predicted PM2.5 exposure by two methods: nearest monitor and kriging. The effect of PM2.5 on cardiovascular disease incidence was estimated conditional on predicted or true exposure in a Cox proportional hazards regression analysis. Across all simulations,

locations were fixed for the 22 monitors (real locations) and 2,000 residences; the true exposures and health outcomes were generated in every simulation.

2.2.1. True models

1) Underlying exposure model

We assumed exposure to PM_{2.5} followed a multivariate normal distribution with variability consisting of spatial and non-spatial components parameterized by the partial sill (σ^2), range (ϕ), and nugget (τ^2) in the spherical covariance model. True exposure (W) at location s_i was generated as

$$W(s_i) = \mu(s_i) + v(s_i) + \varepsilon(s_i) \quad (1)$$

The true exposure was composed of the mean ($\mu(s_i)$) and two error terms, the spatial ($v(s_i)$) and non-spatial errors ($\varepsilon(s_i)$). Variances of these errors are the partial sill (σ^2) and nugget (τ^2), respectively. Correlation of spatial errors is a function of the range (ϕ).¹¹ To represent different spatial correlation scenarios, we focused primarily on six true exposure models with different mean and covariance parameters (table 1).

Parameters for the two models were estimated from analyses of PM_{2.5} data in LA (section 2.1.2). In the first true exposure model (TEM 1), most of the variability of PM_{2.5} was dominated by the geographical covariates. The mean was determined by a second-order polynomial function of XY coordinates,

$$(\mu(s_i) = (-46771) + (6.69 * X) + (24.07 * Y) + (-0.09 * X^2) + (-0.31 * Y^2) + (-0.0016 * X * Y)),$$

while the partial sill, range, and nugget were relatively small

($\sigma^2 = 5.81$ $\phi = 44$ km, $\tau^2 = 0.43$). The other model (TEM 4) had a constant mean

($\mu(s_i) = 18.35 \mu\text{g}/\text{m}^3$) so all the spatial dependence was parameterized in the covariance

model. Thus, the partial sill and range were larger ($\sigma^2 = 46.9$, $\phi = 120$ km) and there was no nugget ($\tau^2 = 0$). To explore sensitivity to the exposure model structure, arbitrary parameter values were chosen for four other models (TEM 2, TEM 3, TEM 5, and TEM 6). These models had the same constant mean and partial sill as the TEM 4 but different range parameters to represent a range of spatial correlations from none to high. Figure 2 shows examples of each exposure surface; the spatial surface is smoother in the first, fourth, fifth, and sixth models with large spatial dependence from the variogram or the second-order mean models. In order to better understand how modeled exposure affected health effect estimates, we conducted a sensitivity analysis by simulating an additional set of forty true exposure models. We assigned a constant mean to half and a polynomial mean to the other half (mean identical to the first true exposure model). Each of the twenty pairs had range parameters (ϕ) between 10 and 500 km, partial sill parameters (σ^2) between 0.1 and 90, and no nugget ($\tau^2 = 0$).

2) Underlying disease model

Survival time to cardiovascular event occurrence (T) followed an exponential distribution with mean, $\gamma = \frac{1}{\lambda_0 \exp(\beta(W - \bar{W}))}$, i.e.

$$T \sim \text{Exp}(\gamma) \quad (2)$$

Censoring was based on two mechanisms: cardiovascular disease-free survival for ten years (study duration) and loss to follow-up (drop-out). Ten percent of all subjects dropped out in the first ten years with time determined by the uniform distribution so the probability of drop-out was constant over all study years.

2.2.2. Fitted models

1) Predicted exposure

For each of the six exposure models, two versions of predicted PM2.5 were estimated from the simulated (“measured”) PM2.5 concentration at the twenty-two monitor locations. Nearest monitor exposures (W^*) assigned the measured PM2.5 at the closest monitor to each individual residential location. Kriged exposures (W^{**}) interpolated measured PM2.5 by universal kriging methods. Kriging prediction consists of two parts: estimating the spatial structure and interpolating the spatial surface. We estimated parameters from the monitor measurements under a variogram with a spherical covariance and a second-order polynomial mean model. Given the estimated parameters, we predicted PM2.5 at individual residence locations.

2) Fitted disease model

The effect of PM2.5 on time to cardiovascular event was estimated by the Cox proportional hazards model, conditional on true ($W' = W$) or predicted exposure ($W' = W^*$ or W^{**}),

$$\lambda(t) = \lambda_0 \exp(\hat{\beta}(W' - \bar{W}')). \quad (3)$$

3) Assessment of validity and reliability

To evaluate validity and reliability of health effect estimates, bias, variance, mean square error, and 95 percent coverage probability were estimated. Coverage probability is the proportion of 95 percent confidence intervals of health effect estimates that include the true β .

In order to gain insight into properties of health effect estimates based on properties of exposures and their prediction, we also investigated validity and reliability of predicted exposure by summarizing average prediction error (difference of predicted

PM2.5 minus true PM2.5), variance of true and predicted exposures at subject locations, mean square prediction error, and 1 minus mean square prediction error relative to true exposure variance (constrained to always be greater than or equal to 0), which is an estimate of R-square.

3. RESULTS

Summary statistics of true and predicted PM2.5 in the first simulation are shown in the online appendix. Across the exposure models, the true PM2.5 exposure means for the 2,000 residential sites were between 17.10 and 26.08 $\mu\text{g}/\text{m}^3$. Predicted PM2.5 had smaller standard deviations and ranges than true PM2.5, particularly for kriging prediction. Relative to the variation in the true exposure data, kriged PM2.5 varied much less in the second and third true exposure models (TEM 2 and TEM 3) which assumed little spatial correlation. A more detailed picture emerges from the true versus predicted exposure scatter plots displayed in figure 3. Predicted PM2.5 was more correlated with true PM2.5 when the true exposure models had more spatial structure. Although models 1 and 4 (TEM 1 and TEM 4) were based on different fits to the same data, model 1 predictions were more correlated with true exposures.

Table 2 gives summary statistics for the quality of predictions over all simulations. Across the models, kriged PM2.5 had smaller average prediction error and variance than nearest monitor PM2.5. Mean square prediction error was smaller for kriging as well, except for model 4. Likewise, estimated R-square was generally higher for kriged predictions.

Table 2 also summarizes qualities of health effect estimates. Kriged PM2.5 health effect estimates were less biased and more variable than health effects from nearest

monitor PM_{2.5}. Estimates with the highest bias were from exposures with the least amount of spatial structure. However, because variance dominated bias, nearest monitor predictions had smaller mean square error than kriging predictions. This trend was consistent across all six exposure models. Moreover, the same pattern was found for higher (RR=1.34) and lower (RR=1.14) true effects (results not shown). Coverage for kriged exposure was generally better than for nearest monitor, but was mostly less than the target 95 percent. With respect to the relationship between predicted exposure and health effect estimates, spatial exposure structures with higher estimated R-square produced higher coverage of health effect estimates.

Figure 4 shows scatter plots of the relationship between health effect estimates. Predictions from the first exposure model (TEM 1) gave the best health effect estimates from kriged exposure. Similarly, the fourth, fifth, and sixth exposure models with larger spatial correlation (TEM 4, TEM 5, and TEM 6) produced a stronger association between health effect estimates for true and predicted exposures than the exposure models with no or little spatial dependence (TEM 2 and TEM 3). Nearest monitor prediction underestimated true effects on average, while kriging only underestimated on average when there was little spatial correlation in the exposure. Note that since there are estimates in the second and fourth quadrants of the nearest vs. kriged exposure scatter plots, there are multiple realizations for each exposure model where the kriging and nearest monitor prediction approaches would lead to health effect estimates with opposite signs.

Figure 5 shows a sensitivity analysis of the relationship between spatial correlation (represented by the range) and health effect estimate properties (mean square

error and coverage) across exposure models with different covariance parameters and mean structure. The models that produced the best health effect estimates (as defined by good coverage and small mean square error) had large range parameters relative to the partial sill. As the range increased, coverage increased towards 95 percent and mean square error decreased. Moreover, the difference between prediction approaches for both coverage and mean square error became smaller. Addition of mean structure had a dramatic effect on mean square error but relatively small impact on coverage. In particular, with mean structure in the model there was only a small difference in mean square error between kriging and nearest monitor exposures for all range parameters.

4. DISCUSSION

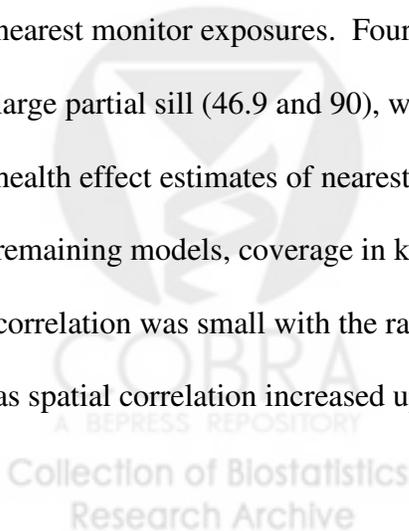
We explored two exposure prediction approaches in a simulation study to estimate the effect of PM_{2.5} on cardiovascular events. Kriging prediction estimated more accurate exposure because it had smaller average mean square prediction error. However, kriged exposure predictions were also less variable which has implications for health effect estimates. Overall, health effect estimates were better with kriged exposure when comparing results based on coverage probability. In contrast, for mean square error kriging generally performed worse than nearest monitor prediction due to greater variance in the health effect estimates resulting from smaller exposure variance. Differences between the two approaches diminished in more spatially-dependent exposure structures where the dependence is represented by a longer range and/or the polynomial mean model.

Health effect estimates for both exposure prediction approaches had better properties as spatial correlation in the covariance model increased and a spatially varying

mean structure was included. Coverage converged to 95 percent and mean square error decreased as the range became larger. This improvement was consistent across different partial sill parameters (figure 5). In addition, when the spatially varying mean model was included, coverage improved slightly while mean square error reduced more dramatically. In particular, mean square error of kriging prediction dramatically declined and even became smaller than that of nearest monitor prediction.

The mean square error and coverage criteria led to generally different conclusions about the preference of nearest monitor versus kriging predictions. Coverage is a very important property of health effect estimates since interval estimates are the basis for inference. Coverage of 95 percent means that the 95 percent confidence interval estimated in an observational study has the correct interpretation, i.e. that the interval has a 95 percent probability of covering the true effect. While nearest monitor prediction had generally smaller estimated mean square error, this was due to less variable but mostly biased effect estimates that produced generally smaller coverage below 95 percent.

When we examined the expanded set of exposure models with different range, partial sill, and mean parameters, all models except five showed better coverage for kriged versus nearest monitor exposures. Four models among them had medium range (120 km) and large partial sill (46.9 and 90), with and without mean structure. The coverages of the health effect estimates of nearest monitor were slightly better (by 1-4 percent). In the remaining models, coverage in kriging was much better (by over 50 percent) when spatial correlation was small with the range of 10 km and became slightly better (by 1-3 percent) as spatial correlation increased up to the range of 500 km (figure 5).



Health effect estimates from the two prediction methods showed features typically attributed to classical and Berkson measurement errors depending upon the spatial structure. Nearest monitor health effect estimates were always attenuated as seen from the lower fitted regression line for nearest monitor compared to true exposure (figure 4). Attenuation was also present for kriged exposure with no spatial correlation or little range (TEM 2 and TEM 3). These underestimated effects behave like attenuation due to classical measurement error. In contrast, health estimates from kriged predictions of exposure models with strong spatial structure showed little or no attenuation of the regression slope but the estimates were more variable than for true exposures. This is behavior typically attributed to a Berkson measurement error model, although in this application the error structure is not independent and identically distributed so the Berkson model does not apply exactly. The standard errors of the health effect parameter estimates in such Berkson-like measurement error models can be corrected with a sandwich variance estimator¹². Since we also showed kriged exposures in models with poor spatial structure gave health effect estimates that were both attenuated and more variable, our results suggest this correction will not suffice. Szpiro et al identify two additional sources of error that result from uncertainty in the estimation of the parameters in the prediction model¹³. They propose methods to correct the measurement error biases from using kriging predictions in a health effects analysis. In future work, we plan to evaluate the performance of their methods for our examples, and we expect the corrected coverage probabilities will be closer to the target 95 percent.

One of the features of our approach is that we considered a monitoring network and exposure models based on existing data in order to ensure the relevance of our results.

Health effect estimates from exposure models derived from analysis of Los Angeles data (TEM 1 and TEM 4) showed good properties. Although the Los Angeles area has a great deal of spatial variability of PM_{2.5} and a relatively large number of monitors compared to other areas, the monitoring network is still relatively sparse from a spatial statistics perspective and thus it may be challenging in practice to model the spatial structure. More work is needed to determine the best approaches to spatial prediction for air pollution exposure data. As a sensitivity analysis for one aspect of this question, we looked at the impact of monitor density on our simulation results. First we increased the number of monitors by five up to forty-two in the first true exposure model (TEM 1). Locations of additional monitors were uniformly distributed within the area covered by the existing twenty-two monitors. The denser network did produce marginally better exposure predictions and health effect estimates. With twenty additional monitors, mean square error for health effect estimates decreased by 16 and 15 percent in nearest monitor and kriging predictions, respectively. However, the coverage probability was almost identical to that of the original number of monitors. We also explored the influence of a sparser monitoring network, a more frequent occurrence in air pollution epidemiology, by reducing the number of monitors by five down to twelve in total. With only twelve monitors (ten less than original monitoring network), our health effect estimates yielded similar coverage probability estimates and a small increase of mean square error estimates for kriged exposures. In contrast, nearest monitor prediction performed much worse. Mean square error increased by over 50 percent. Coverage probability dropped by around 20 percent. This sensitivity analysis suggests that a large number of monitors may not be necessary for kriging prediction in a health effect analysis, at least when there

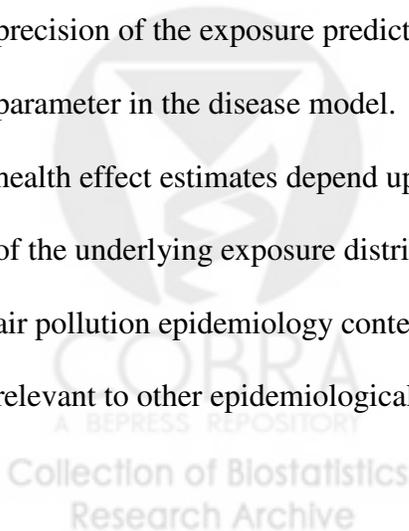
is adequate spatial variability and the form of the spatial model is not in question. Health effect estimates were more sensitive to monitor reduction when individual exposure was predicted using nearest monitor.

We observed that for the kriged exposure, a small percentage of health effect estimates resulted in large outliers (omitted from figure 4). In further evaluation of those simulations, we found the exposure model covariance parameter estimates were unreasonably large. Under the spatial structure with very small range, poorly estimated covariance parameters induced the most extreme outlier occurrences. As part of our analysis, we examined both universal and ordinary kriging fitting procedures for all exposure models. We found that estimation using ordinary kriging (i.e. with a constant mean) produced consistently poorer covariance parameter estimates across all exposure models. The universal kriging approach gave more stable parameter estimates even when there was no underlying mean structure in the true exposure model. While it is not practical to investigate a large number of mean and covariance models in a simulation study, further research should be done to determine whether other spatial statistical models or fitting algorithms would produce more reasonable parameter estimates in such cases.

In order to focus this research on features of spatially dependent exposure and prediction that affect health effect parameter estimates, we made some strong assumptions. Most important, we assumed ambient air pollution concentration was identical to ambient source exposure. This assumption does not reflect the difference of exposure from concentration due to time activity and local sources or the attenuation of ambient exposure by infiltration. Future research will need to incorporate this exposure

feature. Another topic that needs additional research is investigation of the importance of the multivariate normal distribution for the underlying true exposure. It will be valuable to determine whether our conclusion that kriged exposure is preferable to nearest monitor prediction still holds when this distributional assumption is relaxed, since kriging assumes a multivariate normal spatial surface. Finally, we assumed there was no additional spatial structure in the disease model. There may be spatial dependence in the health outcome in addition to the spatial structure induced from the exposure distribution. Work is needed to incorporate spatial dependence in the disease model and determine how separable the two sources of spatial structure may be in the health analysis.

While the focus of this paper is specific to the field of air pollution epidemiology, this research is broadly relevant to inference in epidemiological studies in the presence of limited exposure data. We present a problem where there are no exposure measurements for study subjects but where it is possible to use an external dataset to develop an exposure prediction model to plug into the health effect analysis. We evaluate the merits of two exposure prediction approaches directly in terms of the parameter of interest in the epidemiological study. We learned that it is not sufficient to assess the accuracy and precision of the exposure predictions when the target of inference is the health effect parameter in the disease model. We showed that the bias and precision of the resulting health effect estimates depend upon the approach to exposure prediction and the features of the underlying exposure distribution. While the details of our results are specific to the air pollution epidemiology context, our general approach and overarching conclusions are relevant to other epidemiological study settings.



In summary, based on the generally better coverage and lower bias properties of the health effect parameter estimate of interest, we conclude that kriging exposure prediction is preferable to nearest monitor prediction for estimation of air pollution health effects, particularly when the underlying true exposure is less spatially correlated. Both prediction approaches performed well when there is good spatial structure in the underlying exposure distribution.



REFERENCES

1. Dockery DW, Pope CA 3rd, Xu X, et al. An association between air pollution and mortality in six U.S. cities. *N Engl J Med.* 1993;329:1753-1759.
2. Pope CA 3rd, Burnett RT, Thun MJ, et al. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA.* 2002;287:1132-1141.
3. Basu R, Woodruff TJ, Parker JD, Saulnier L, Schoendorf KC. Comparing exposure metrics in the relationship between PM_{2.5} and birth weight in California. *J Expo Anal Environ Epidemiol.* 2004;14:391-396.
4. Miller KA, Siscovick DS, Sheppard L, et al. Long-term exposure to air pollution and incidence of cardiovascular events in women. *N Engl J Med.* 2007;356:447-458.
5. Ritz B, Wilhelm M, Zhao Y. Air pollution and infant death in southern California, 1989-2000. *Pediatrics.* 2006;118:493-502.
6. Wilhelm M, Ritz B. Local variations in CO and particulate air pollution and adverse birth outcomes in Los Angeles County, California, USA. *Environ Health Perspect.* 2005;113:1212-1221.
7. Finkelstein MM, Jerrett M, Sears MR. Environmental inequality and circulatory disease mortality gradients. *J Epidemiol Community Health.* 2005;59:481-487.
8. Jerrett M, Burnett RT, Ma R, et al. Spatial analysis of air pollution and mortality in Los Angeles. *Epidemiology.* 2005;16:727-736.
9. Kunzli N, Jerrett M, Mack WJ, et al. Ambient air pollution and atherosclerosis in Los Angeles. *Environ Health Perspect.* 2005;113:201-206.
10. Leem JH, Kaplan BM, Shim YK, et al. Exposures to air pollutants during pregnancy and preterm delivery. *Environ Health Perspect.* 2006;114:905-910.

11. Banerjee S, Carlin BP, Gelfand AE. *Hierarchical modeling and analysis for spatial data*. Chapman & Hall/CRC; 2004:21-68.
12. Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM. *Measurement error in nonlinear models*. 2nd ed. Chapman & Hall/CRC; 2006.
13. Szpiro AA, Sheppard L, Lumley T. Accounting for errors from predicting exposures in environmental epidemiology and environmental statistics. *UW Biostatistics Working Paper Series*. Working Paper 330. (June 19, 2008).
<http://www.bepress.com/uwbiostat/paper330>



Figure legends

Figure 1. Locations of 2,000 hypothetical study subjects, and 22 PM2.5 monitors in the Los Angeles area.

Figure 2. Spatial surface of the six true exposure models. Darker shading indicates higher concentration. Black dots on the spatial surface for the TEM 2 represent the locations of the 22 PM2.5 monitoring sites in LA.

Figure 3. Relationship between true and predicted exposures for the six true exposure models in which sources of spatial variability come from geographic characteristics (TEM 1), little spatial correlation (TEM 2 and 3), medium spatial correlation (TEM 4), and high spatial correlation (TEM 5 and 6). Green line is the X-Y identity.

Figure 4. Relationship between estimated health effects from true and predicted exposures for the six true exposure models in which sources of spatial variability come from geographic characteristics (TEM 1), little spatial correlation (TEM 2 and 3), medium spatial correlation (TEM 4), and high spatial correlation (TEM 5 and 6) after excluding outlying estimates. Green and red lines display the X-Y identity line and the best-fit line between health effect estimates of true and predicted exposures, respectively.

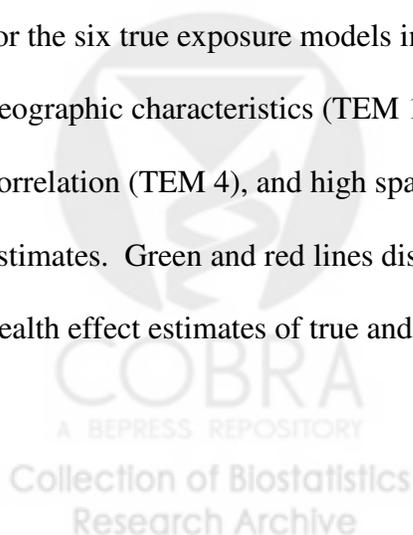


Figure 5. Relationship of range versus mean square error (MSE) and coverage probability of health effect estimates for kriging exposure and differences of kriging from nearest monitor exposures for all forty true exposure models. Solid and dotted lines represent constant and polynomial mean models, respectively. Horizontal line in red for the difference figures is at zero to show no difference. Several extreme values for MSE and the MSE difference occurred when partial sill is equal to 0.1 and are not shown. Differences are for estimates (of MSE or coverage) for kriging exposure minus nearest monitor exposure



Table 1. Source of spatial variability and parameter value in the six true exposure models

True exposure model		Parameter value			
		Nugget (τ^2)	Partial sill (σ^2)	Range (km) (ϕ)	Mean ($\mu\text{g}/\text{m}^3$) ($\mu(s_i)$)
					$-46771+6.69X+24.07Y$
TEM 1	Geographical characteristics	0.43	5.81	44	$-0.09X^2-0.31Y^2-0.0016XY$ ¹⁾
TEM 2	No spatial correlation	46.9	0	0	18.35
TEM 3	Low spatial correlation	0	46.9	10	18.35
TEM 4	Medium spatial correlation	0	46.9	120	18.35
TEM 5	High spatial correlation	0	46.9	277	18.35
TEM 6	Highest spatial correlation	0	46.9	500	18.35

1) 'X' and 'Y' mean X and Y coordinates. Original latitude and longitude coordinates were converted into kilometers in the UTM coordinate system

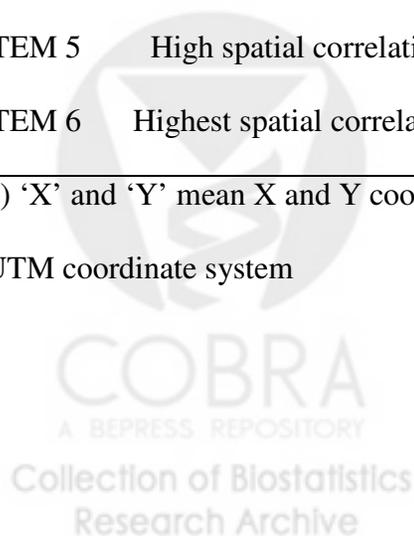


Table 2. Summary estimates of validity and reliability of predicted exposure and health effect estimate by the six true exposure models across 1,000 simulations

Exposure model		Exposure ¹⁾				Health effect ²⁾			
True	Fitted	MAPE ³⁾	RMV ⁴⁾	RMMSPE ⁵⁾	Estimated R ² ⁶⁾	Bias ²	Variance	MSE	CP ⁷⁾
TEM 1	True		2.93			0	47	48	0.97
	Nearest	0.43	2.59	2.44	0.29	36	93	129	0.83
	Kriging	0.00	2.29	2.07	0.47	0	136	136	0.93
TEM 2	True		6.84			0	8	8	0.95
	Nearest	0.06	6.31	9.67	0	463	13	476	0.00
	Kriging	0.04	2.07	7.45	0	472	291	763	0.51
TEM 3	True		6.79			0	8	8	0.96
	Nearest	0.06	6.44	8.92	0	326	21	347	0.02
	Kriging	0.03	2.01	7.26	0	339	686	1025	0.54
TEM 4	True		4.25			0	27	27	0.95
	Nearest	-0.01	4.03	3.22	0.35	34	54	88	0.75

	Kriging	-0.02	2.42	3.91	0.20	0	727	728	0.71
TEM 5	True		2.88			0	59	59	0.96
	Nearest	0.01	2.74	2.13	0.36	30	108	139	0.85
	Kriging	0.01	2.09	2.15	0.40	1	400	401	0.87
TEM 6	True		2.16			0	108	108	0.94
	Nearest	-0.01	2.06	1.58	0.37	32	143	175	0.92
	Kriging	-0.04	1.71	1.44	0.47	0	332	332	0.95

1) Summary of 1,000 realizations (average prediction error, variance, mean square prediction error, and estimated R-square)

for 2,000 exposures in each simulation

2) Bias², variance, and mean square error (MSE) for health effect estimates in 1000 simulations were multiplied by 10⁶

3) Mean of average prediction error for the difference of predicted PM2.5 minus true PM2.5

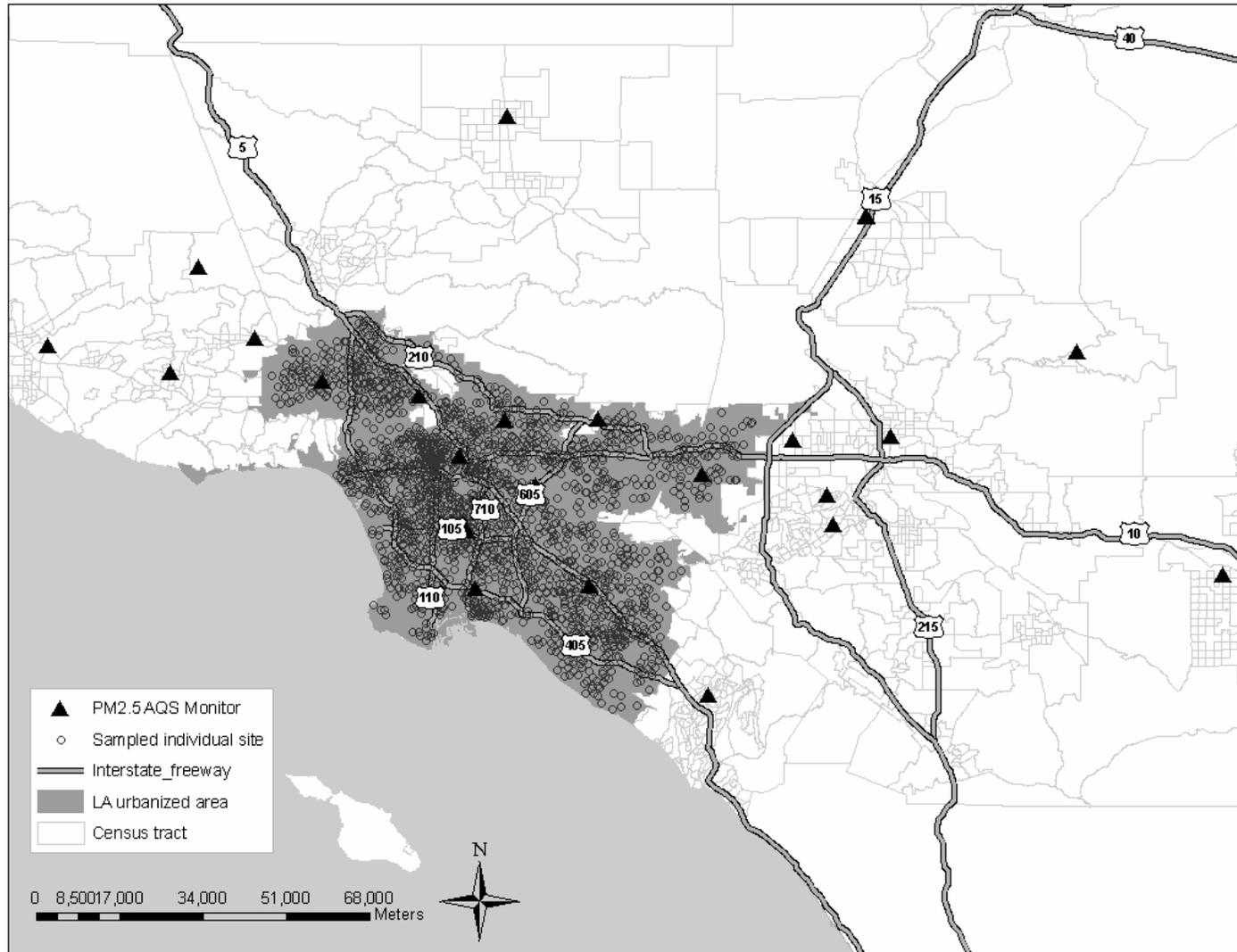
4) Root mean of variance

5) Root mean of mean square prediction error

6) Mean of 1 minus mean square prediction error divided by variance of true exposure. Estimates were constrained to have a minimum of zero.

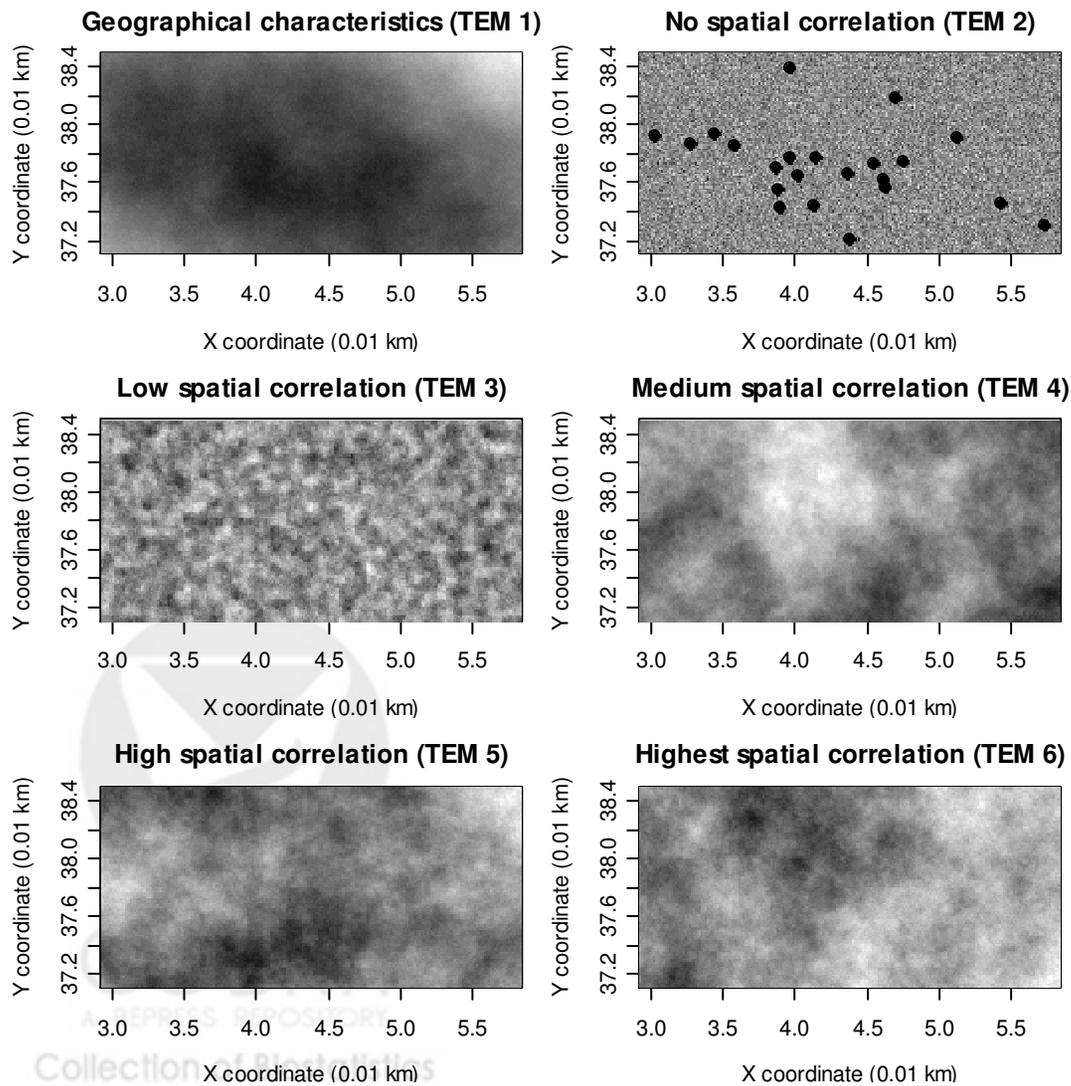
7) Coverage probability of health effect estimate

Figure 1



Research Archive

Figure 2



UC Berkeley
Collection of Biological
Research Archive

Figure 3

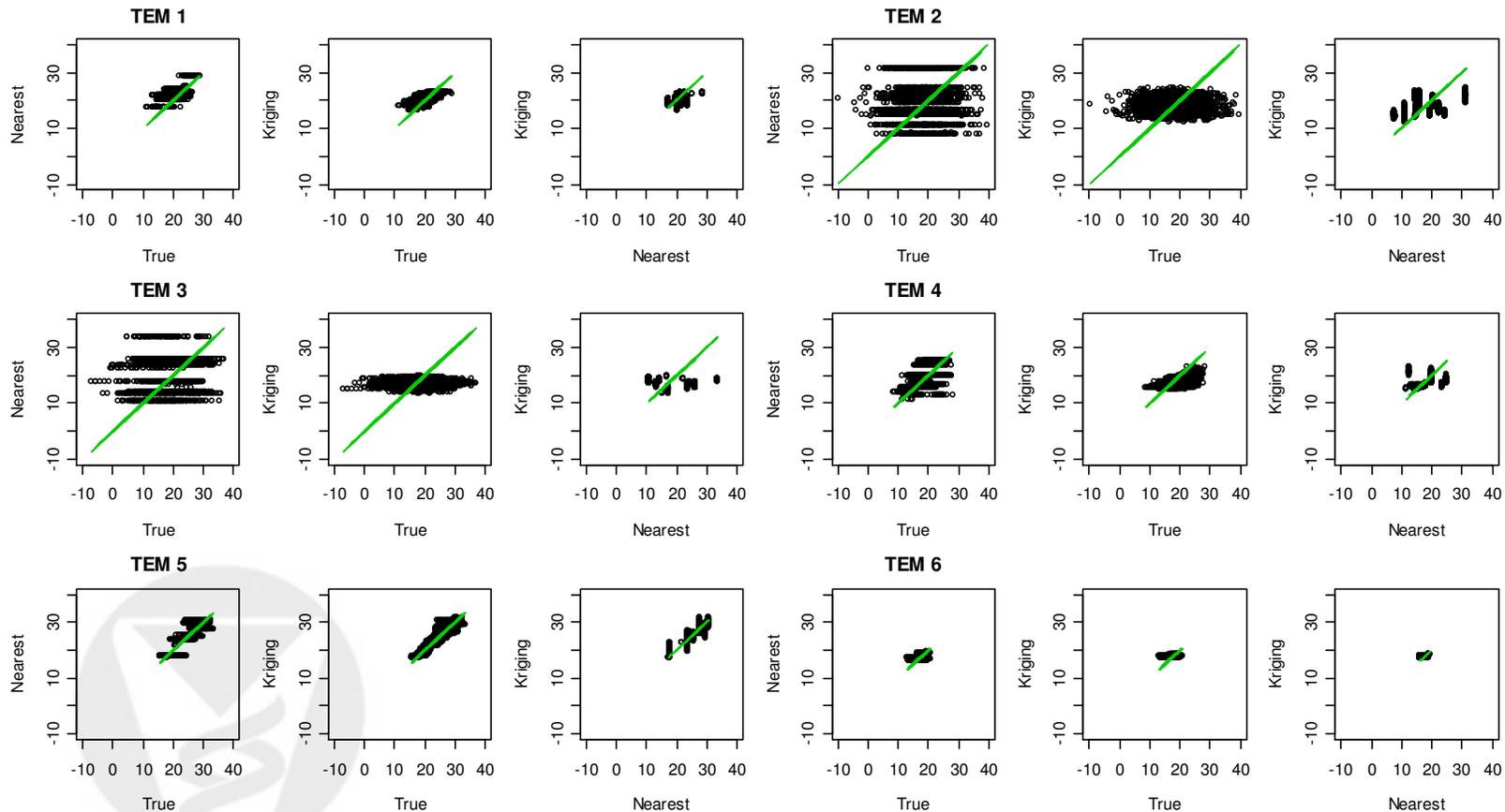


Figure 4

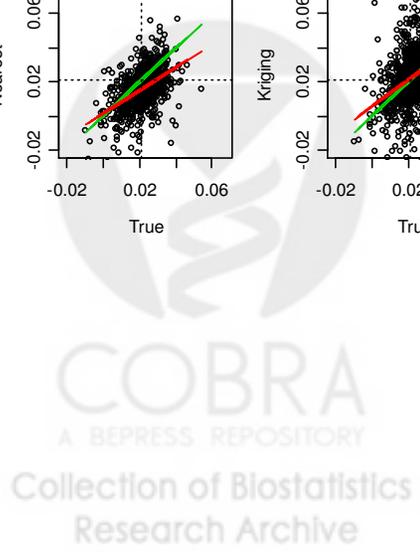
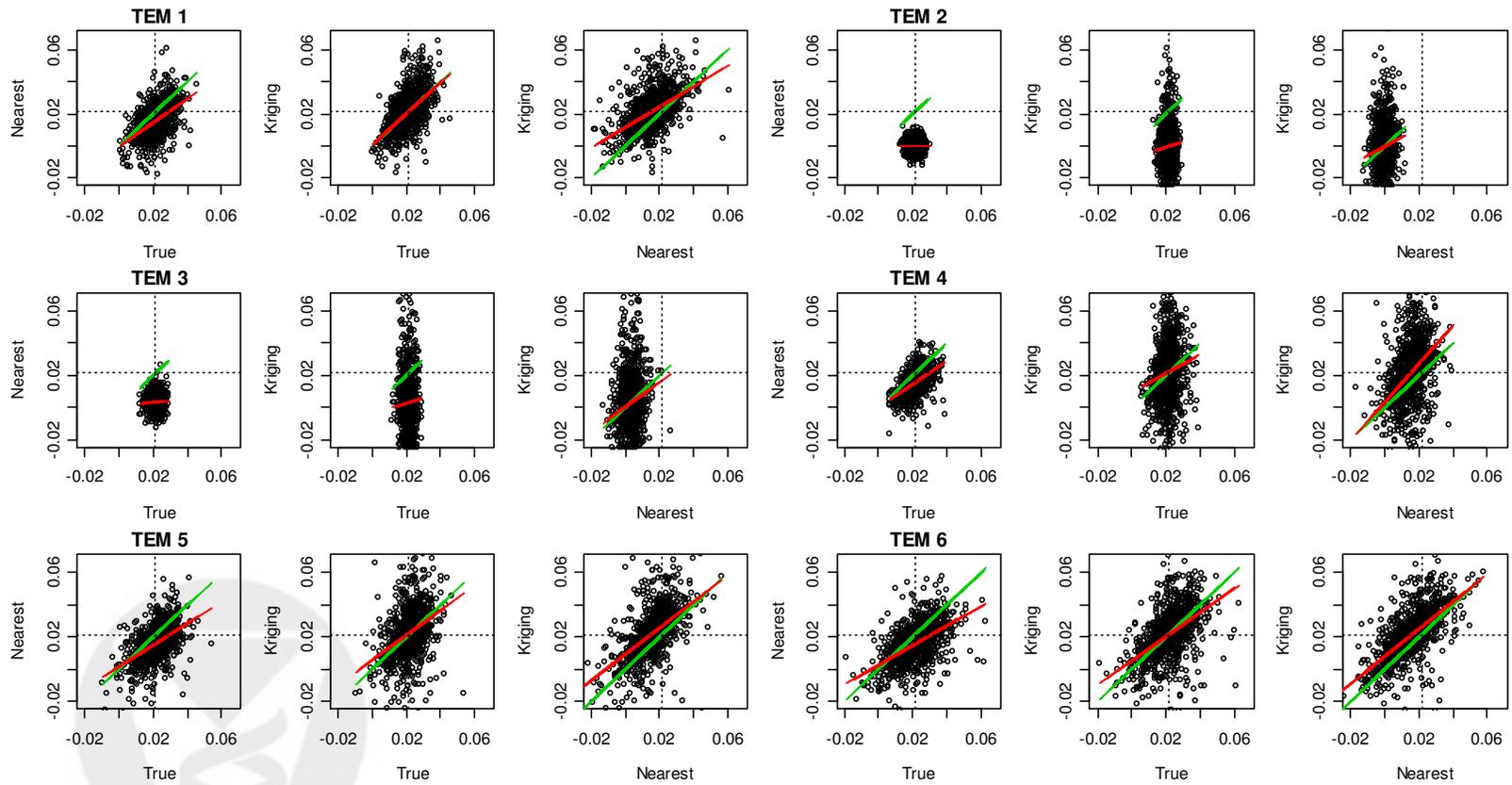
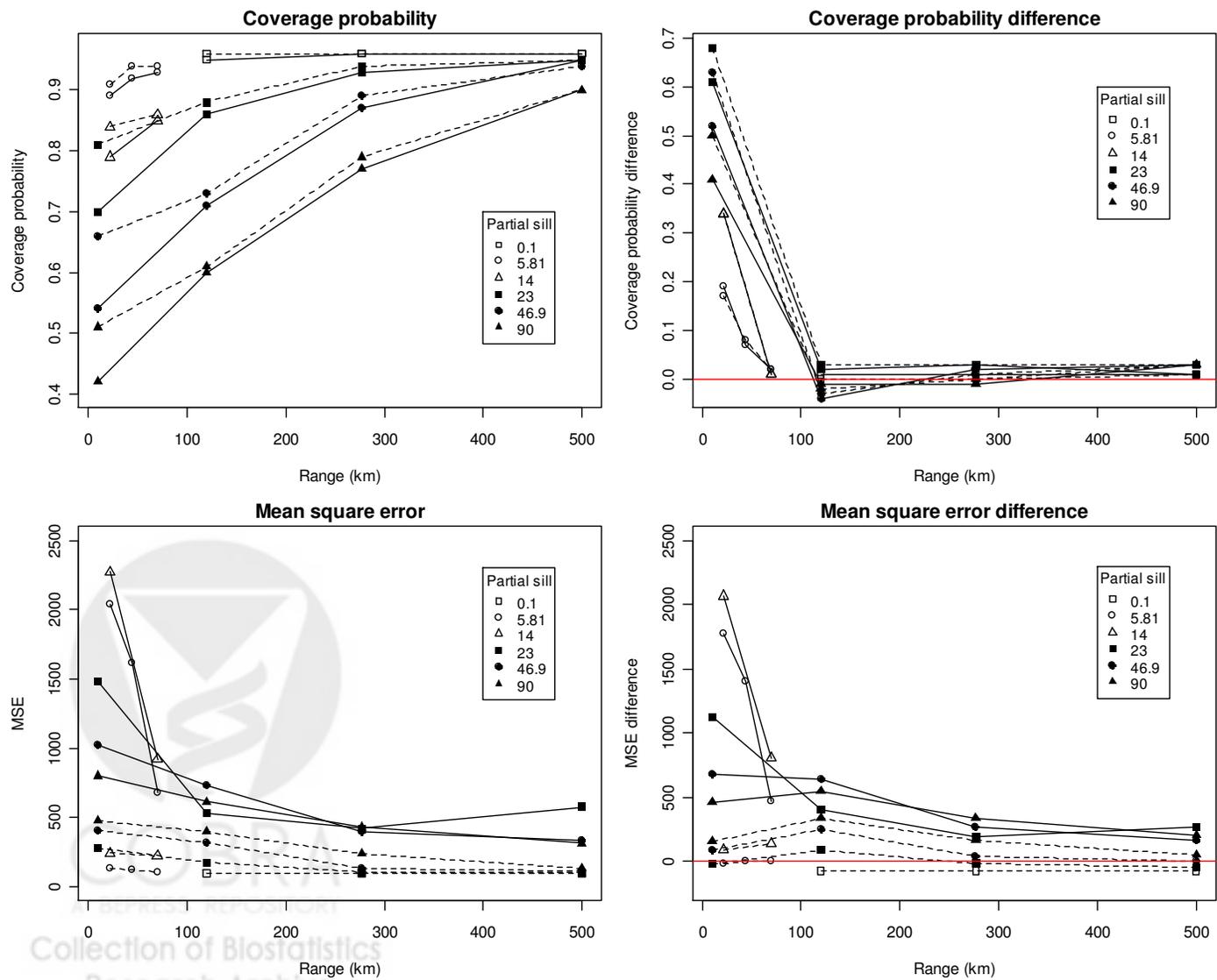


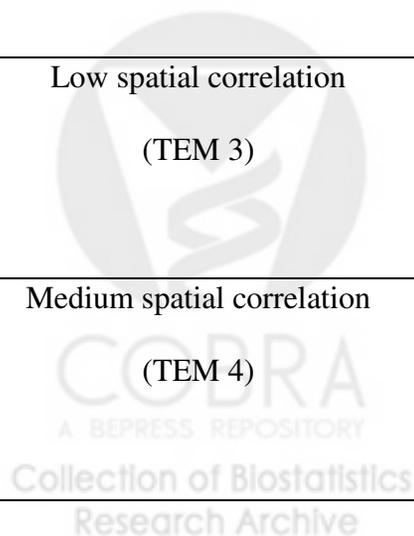
Figure 5



Online Appendix

Table. Summary statistics of PM2.5 concentration by the six true exposure models in the first simulation

True exposure model	Fitted exposure model	Min	Q1	Median	Q3	Max	Mean	SD
Geographical characteristics (TEM 1)	True	11.34	19.64	20.92	22.38	28.98	20.94	2.36
	Nearest	17.40	20.45	20.69	21.68	28.58	21.64	2.02
	Kriging	16.38	20.02	20.85	21.75	22.94	20.79	1.26
No spatial correlation (TEM 2)	True	-9.71	13.41	18.44	22.84	39.77	18.26	6.82
	Nearest	7.76	15.04	16.24	20.47	31.49	17.89	6.14
	Kriging	12.41	15.70	17.84	19.87	24.50	17.91	2.59
Low spatial correlation (TEM 3)	True	-7.21	14.08	19.02	23.58	36.70	18.73	6.77
	Nearest	10.63	14.12	17.49	23.62	33.75	19.20	6.14
	Kriging	13.54	16.15	17.01	18.11	19.62	17.03	1.25
Medium spatial correlation (TEM 4)	True	8.62	16.84	19.48	22.39	28.28	19.46	3.61
	Nearest	11.47	16.85	17.94	23.58	25.16	18.82	3.66
	Kriging	14.77	15.94	16.28	17.67	23.26	16.95	1.52



High spatial correlation (TEM 5)	True	15.47	23.84	26.57	29.15	33.28	26.08	3.79
	Nearest	17.34	23.83	27.47	29.43	30.52	25.69	3.94
	Kriging	17.17	23.45	26.44	28.61	31.98	25.70	3.64
Highest spatial correlation (TEM 6)	True	13.00	16.11	17.07	18.13	20.73	17.10	1.47
	Nearest	16.08	16.46	17.07	17.40	19.42	17.21	1.01
	Kriging	17.36	17.44	17.54	17.77	18.53	17.64	0.25

