GIS Approaches for the Estimation of Residential-Level Ambient PM Concentrations

Duanping Liao,1 Donna J. Peuquet,2 Yinkang Duan,1 Eric A. Whitsel,3,4 Jianwei Dou,2 Richard L. Smith,5 Hung-Mo Lin,1 Jiu-Chiuan Chen,3 and Gerardo Heiss3

1Department of Health Evaluation Sciences, Pennsylvania State University College of Medicine, Hershey, Pennsylvania; USA; 2Department of Geography, Pennsylvania State University, College Park, Pennsylvania, USA; 3Department of Epidemiology, Department of Medicine, and 5Department of Statistics, University of North Carolina–Chapel Hill, Chapel Hill, North Carolina, USA

Large-scale, population-based epidemiologic investigations of the health effects of ambient air pollution often rely on measurements from a network of air quality monitors maintained by the U.S. Environmental Protection Agency (U.S. EPA 1995a, 1995b, 2005). The Air Quality System (AQS) is the only national ambient air pollution database currently available for public use in the United States. The availability of individual-level health outcome and covariable data from national-scale studies that often characterize participants over the course of several years enables researchers to study the acute effects of ambient air pollution using individual-level data (Liao et al. 2004, 2005a; Sullivan et al. 2005; Wellenius et al. 2005; Whitsel et al. 2004). This approach requires measures of daily particulate matter (PM) exposures, ideally assessed as close to the individual level as possible, such as at participant residences or in immediate proximity to participants themselves. Because daily measures of ambient PM concentrations from the AQS are unavailable in the large majority of locations, spatial estimation methods using geographic information systems (GIS) are increasingly being considered to estimate geocoded location-specific ambient PM concentrations, such as kriging methods. Important methodologic and practical issues still need to be resolved, however. This study was designed to (a) assess the feasibility of large-scale kriging estimation of daily residential-level ambient PM concentrations, (b) perform and compare cross-validations of different kriging models, (c) contrast three popular kriging approaches, and (d) determine and contrast the most appropriate kriging approaches, and (e) compare the SEs of the kriging estimations.

Researchers have developed kriging methods and geostatistical models to estimate concentrations of ambient particulate matter (PM) at locations not monitored by the AQS. Spatial kriging has been used to estimate PM concentrations from the AQS database. These data served as pollutant- and site-specific daily source data for our study (Liao et al. 2005b). We geocoded 94,135 addresses of Women’s Health Initiative (WHI) Clinical Trial (CT) participant residences and examination sites in the contiguous 48 United States and District of Columbia, after assessing geocoding vendor error (Whitsel et al. 2004, 2005). Daily PM10 and PM2.5 concentrations and the associated estimation errors (SEs) are estimated at these geographic locations by the Environmental Epidemiology of Arrhythmogenesis in WHI study (Whitsel 2006).

We used ArcView GIS (version 8.3) and its Geostatistical Analyst Extension (ESRI Inc., Redlands, CA) for semivariogram determination and cross-validation and for subsequent spatial estimation of daily location-specific PM concentrations. Three frequently referenced spatial models (spherical, exponential, Gaussian) (Cressie 1993a; Davis 2002) were considered using the weighted least-squares method (Gribov et al. 2004; Jian et al. 1996) to obtain the “optimal” daily semivariogram parameters (range, partial sill, and nugget). Based on the daily semivariograms, we performed ordinary kriging to estimate the daily mean PM concentration and its SE at each of the 94,135 geocoded addresses. Next, we performed the standard cross-validation—an iterative procedure that omits site-specific PM data points one at a time and refits the model using the remaining data to estimate the PM concentration at the site of the omitted observation. We assessed the validity (also termed “goodness of fit”) of each semivariogram using three cross-validation parameters readily available from the

© Environmental Health Perspectives is published by the National Institute of Environmental Health Sciences, National Institutes of Health, U.S. Department of Health and Human Services, P.O. Box 12233, Research Triangle Park, NC 27709. Copyright © 2006. All rights reserved. This is a U.S. Government work and is in the public domain in the USA. EHP articles may be reproduced or distributed in any format for noncommercial use without written permission from the author(s), as long as the work is properly cited. Request permission for commercial use from Permissions@lww.com. DOI: 10.1289/ehp.9169

Address correspondence to D. Liao, Department of Health Evaluation Sciences, Pennsylvania State University College of Medicine, 600 Centerview Dr., A210, Hershey, PA 17033 USA. Telephone (717) 531-4149; Fax: (717) 531-5779. E-mail: dliao@psu.edu

We acknowledge the contributions of WHI investigators and institutions (Appendix).

The National Institute of Environmental Health Sciences funded this ancillary study (5-R01-ES012258). The National Heart, Lung, and Blood Institute, U.S. Department of Health and Human Services, funded the Women’s Health Initiative (WHI) program. The authors declare they have no competing financial interests.

Received 15 March 2006; accepted 8 June 2006.
ArcView software package: a) the average of prediction error (PE), where PE is the average of the difference between the predicted and measured daily PM values at each monitoring site; b) the average of standardized prediction error (SPE), where SPE is the PE divided by the SE of estimation across all sites; and c) root mean square standardized (RMSS), the standard deviation (SD) of all SPEs across all sites. Additionally, we assessed the goodness of fit of each semivariogram by the average of the SEs of the estimations, generated by the kriging procedure, across all 94,135 geocoded addresses. The expectations for a good-fitting semivariogram and kriging model are an average PE and SPE near 0, an RMSS near 1, and a small SE. If RMSS < 1, there is a tendency toward overestimation of the variance; if > 1, there is a tendency toward underestimation (ESRI Inc. 2001). These criteria were consistently used to guide our model selection processes throughout this study (Liao et al. 2005c).

As an alternative to using the automatically calculated semivariogram (calculated using the weighted least-squares method (Gribov et al. 2004; Jian et al. 1996), one can also manually specify the semivariogram parameters to improve the cross-validation parameters in ArcView. We selected six least satisfactory daily semivariograms throughout year 2000 and manually adjusted the semivariogram parameters to obtain the best achievable average RMSS and SPE (RMSS as close to 1 and average SPE as close to 0 as possible). The cross-validation parameters from the weighted least-squares method–calculated semivariograms were then compared to those of the manually adjusted semivariograms.

We performed daily ordinary krings on both the original scale (regular ordinary kriging) and the lognormal scale (lognormal ordinary kriging) (Cressie 1993b; Johnston 2001) for all WHI CT addresses for the year 2000 and compared the cross-validation parameters between the two kriging procedures. Lognormal ordinary kriging was used because it has the ability to eliminate the negative predicted values, which is a problem in ordinary kriging, especially when the source data contain extreme values.

**Results**

**Characteristics of the site-specific daily average PM$_{10}$ and PM$_{2.5}$ concentrations.** During 1994–2003, the number of monitoring sites that provided GIS-readable daily PM$_{10}$ data varied widely (range, 120–1,340). On 17% of days, GIS-readable data were provided by ≥ 400 monitoring sites; on 39% of days, by 200–400 sites; and on 44% of days, by 120–200 sites. The corresponding values for PM$_{2.5}$ during 1999–2003 were 33% of days by ≥ 400 sites and 67% of days by 148–400 sites. Specific to the year 2000, there were averages of 325 PM$_{10}$ and 456 PM$_{2.5}$ monitoring sites operating per day across the contiguous United States, with minima and maxima of 148 and 1,061 sites for PM$_{10}$ and 178 and 1,019 sites for PM$_{2.5}$. As a result, there were 118,791 site-days during 2000 for which we can retrieve measured PM$_{10}$ data and 166,796 site-days for PM$_{2.5}$ data. The mean (± SD) of PM$_{10}$ and PM$_{2.5}$ from these retrievable site-days were 26.29 ± 58.13 and 13.14 ± 8.59 µg/m$^3$, respectively, with medians of 21.33 and 11.20 µg/m$^3$, respectively. A right-skewed distribution of both PM$_{10}$ and PM$_{2.5}$ are evident, especially for PM$_{10}$. Figure 1 illustrates the spatial relationships between the geocoded addresses and the PM monitoring sites on an optimal day and a typical day. The mean distance between each address and its nearest PM monitor was 12.35 km, with an SD of 13.98 km, a median of 7.81 km, and an interquartile range of 10.53 km, and 99th percentile of 68.36 km.

**Comparisons of three widely used spatial models.** Tables 1 and 2 present summary statistics of the cross-validation parameters (PE, SPE, and RMSS) comparing three widely used spatial models (spherical, exponential, Gaussian) for PM$_{10}$ and PM$_{2.5}$, respectively. In general, both average PE and average SPE are very close to 0, with a very narrow range of variation from the 366 daily cross-validations. More specifically, > 95% of average PEs were within ± 2 µg/m$^3$ of measured PM$_{10}$, and ± 0.5 µg/m$^3$ of measured PM$_{2.5}$, an average measurement error that we considered acceptable. In terms of RMSS, we considered > 95% of cross-validations as acceptable, but there were days when RMSS indicated a slight over- or underestimation of the prediction variability. These data support the overall validity of using kriging-based estimation approaches to estimate location-specific PM concentrations across the contiguous United States.

**Comparisons of default and manually adjusted semivariograms.** Table 3 presents the cross-validations and actual kriging estimations from the weighted least-squares mean method calculated semivariogram and manually adjusted semivariogram. For the 6 days when the PE, SPE, or RMSS indicated a less satisfactory default-calculated semivariogram, these three cross-validation parameters could be improved satisfactorily through adjustment of the semivariogram parameters by an operator. However, the application of such “improved” semivariograms to the estimation of PM$_{10}$ concentrations at geocoded locations across the United States did not necessarily provide better estimation of location-specific PM (i.e., smaller SEs). To the contrary, the average SEs from the default semivariograms were smaller than those from manually adjusted semivariograms. Because each average SPE of the default-calculated daily semivariograms was close to 0, and each default-calculated daily semivariogram produced a smaller estimation error, we recommend using the default-calculated semivariogram, even though the RMSS from the default-calculated semivariogram was not fully satisfactory.

**Comparisons of regular versus lognormal ordinary krings.** We applied regular ordinary kriging (spherical model, default-calculated

![Figure 1](image-url). Spatial relationships between the residential locations, PM monitoring sites, and the geographic regions of the study.
daily semivariograms) to estimate daily PM10 concentrations at geocoded addresses (n = 94,135) of WHI CT participants and examination sites in the contiguous United States. We examined the estimated PM10 concentrations and identified 22 days during 2000 when estimated values exceeded the range of observed values. In some cases, the estimated values were negative. The number of addresses affected by this problem ranged from a few on most days to 3.5% of all addresses. This problem was related to skewed PM10 distributions and to small numbers of extreme outlying values or operating sites on some days. We therefore compared regular ordinary kriging and lognormal ordinary kriging anticipating that lognormal kriging would attenuate this problem.

Table 4 lists the 22 days on which regular ordinary kriging yielded estimated PM10 values that were outside the range of measured values. For comparison, the minima and maxima of the measured and estimated PM10 concentrations from both regular and lognormal ordinary krigings are also listed in Table 4. In summary, during 2000, lognormal ordinary kriging effectively reduced the number of problematic days from 22 to 1. Even on this one day, lognormal ordinary kriging yielded a minimum value that was closer to the range of measured data than that from regular ordinary kriging.

Table 5 shows the mean values of cross-validation parameters of daily PM10 semivariograms for both regular ordinary kriging and lognormal ordinary kriging. Cross-validation parameters were within the acceptable range from both regular and lognormal ordinary krigings, except for the 22 "out-of-range" days as defined above. On these out-of-range days, the SPE was well within the acceptable range for both regular and lognormal krigings, but the RMSS was > 1 from both approaches. Even so, for these out-of-range days RMSS from lognormal ordinary kriging was closer to 1 than that from regular ordinary kriging.

We then performed regular and lognormal ordinary kriging to estimate PM10 concentrations at geocoded addresses of WHI CT participants and examination sites, based on year 2000 PM10 data (94,135 locations and 366 days). The mean, SD, median, and maximum of the daily mean SE of the estimated PM10 from the regular ordinary kriging were 27.36, 83.35, 13.93, and 1160.20 µg/m3, respectively. In contrast, those from the lognormal ordinary kriging were 16.29, 6.65, 15.05, and 67.46 µg/m3. Clearly, the distribution of the estimation errors from lognormal ordinary kriging was considerably less skewed and had fewer outlying values than that from regular ordinary kriging. Alternative methods (winorizing extreme PM10 values; using ArcView’s “no-sector” option to search for measured data points from a circle centered around a location that needs of an estimation—i.e., disabling the default “sector” search for measured data points in the four sectors of a circle, reducing the range or nugget) were less effective in estimating predicted values within the range of measured values (data not shown).

Similar to the situation observed in PM10 estimations, lognormal ordinary kriging also effectively eliminated the negative or out-of-range problem that occurred in about 5% of PM1.5 data when using regular ordinary kriging. Other cross-validation parameters were comparable between the lognormal and regular ordinary krigings (data not shown).

**Comparisons between national and regional krigings.** From the 61 days when 900 or more monitoring sites were operating in the year 2000 in the 48 contiguous states, the first of such days from each month was selected for comparisons between ordinary kriging models on a national versus regional scale. National krigings and cross-validations were performed on these 12 selected days using daily site-specific PM10 data. Regional krigings and

### Table 1. Cross-validation summary statistics and semivariogram parameter estimates for PM10 from three different spatial models, year 2000.

| Model       | Days | Mean    | SD      | Median   | 2.5th percentile | 97.5th percentile |
|-------------|------|---------|---------|----------|------------------|------------------|
| PE (µg/m³)  |      |         |         |          |                  |                  |
| Exponential | 366  | 0.2347  | 1.3212  | 0.0294   | −0.6437          | 1.6690           |
| Gaussian    | 366  | −0.1037 | 1.0509  | −0.1216  | −1.1230          | 1.0020           |
| Spherical   | 366  | 0.0629  | 1.1999  | −0.0705  | −0.7914          | 1.4810           |
| RMSS        |      |         |         |          |                  |                  |
| Exponential | 366  | 1.8374  | 1.5431  | 1.1410   | 0.8638           | 6.0240           |
| Gaussian    | 366  | 1.1709  | 0.9891  | 1.0070   | 0.8140           | 2.2660           |
| Spherical   | 366  | 1.2549  | 0.7988  | 1.0270   | 0.8094           | 4.1550           |
| SPE         |      |         |         |          |                  |                  |
| Exponential | 366  | 0.0118  | 0.0330  | 0.0036   | 0.0274           | 0.1058           |
| Gaussian    | 366  | −0.0094 | 0.0333  | −0.0071  | −0.0418          | 0.0284           |
| Spherical   | 366  | −0.0011 | 0.0212  | −0.0034  | −0.0318          | 0.0470           |
| Nugget (µg/m³) |  |         |         |          |                  |                  |
| Exponential | 366  | 2.8373  | 27.8393 | 93.5230  | 0.0000           | 5.3320           |
| Gaussian    | 366  | 4.0861  | 38.7389 | 191.975  | 26.6230          | 7.4862           |
| Spherical   | 366  | 3.5152  | 33.3490 | 142.955  | 0.0000           | 7.1343           |
| Partial sill (µg/m³) |  |         |         |          |                  |                  |
| Exponential | 366  | 7.9574  | 91.5892 | 258.515  | 94.1300          | 23.0070          |
| Gaussian    | 366  | 6.4833  | 73.3159 | 176.240  | 39.4570          | 23.7160          |
| Spherical   | 366  | 6.3742  | 71.0240 | 201.215  | 36.6550          | 22.7360          |
| Range (m)   |      |         |         |          |                  |                  |
| Exponential | 366  | 2.6962  | 2.8321  | 1.392250 | 282.500          | 9.0642           |
| Gaussian    | 366  | 2.1632  | 2.2770  | 1.207050 | 262.460          | 8.9583           |
| Spherical   | 366  | 2.4479  | 2.3753  | 1.424500 | 280.820          | 8.9583           |

*Data daily operating monitoring sites range from 148 to 1,061 sites.*
cross-validations were performed on the same data using the regional map (Figure 1) that divides the U.S. continent into five regions (northwest, southwest, middle north, southeast, and northeast). These five regions were created based on the assumption that different semivariogram parameters would be needed for different geographic areas. In general, for both regional and national krings, the average SPE and RMSS from cross-validations of semivariograms calculated for the 12 selected days were very close to 0 and 1, respectively (Table 6).

**Discussion**

Classical methods often assume that measures are uniformly or randomly distributed. The assumptions are often inappropriate for analyses of environmental measures because values at neighboring locations are rarely independent, particularly over short distances. This form of dependence (spatial autocorrelation) nonetheless makes it possible to interpolate values at unmonitored locations from known values at monitored locations. Krigeing is one such interpolation method originally developed by mining engineers (Krige 1966). It is especially attractive in this setting because it takes the spatial autocorrelation structure function (variogram) into account by considering known values from monitored locations, weighting them with values read from the variogram at corresponding distances, and splitting weights among adjacent locations. The method thereby ensures that interpolations do not depend on monitor density (Legendre and Fortin 1989). By doing so, krigeing yields best linear unbiased estimates, in this setting, of location-specific daily mean ambient PM concentrations and their SEs.

Large-scale population-based epidemiologic investigations of the health effects of ambient air pollution often rely on data collected from a network of air quality monitors maintained by the U.S. EPA—the AQS data (U.S. EPA 1995a, 1995b, 2005). It is revealing to compare krigeing with interpolation methods used in the well-known time-series and cohort studies of PM effects on mortality and cardiovascular disease (Abbey et al. 1991, 1999; Dockery et al. 1993; Katsouyanni et al. 1996, 2001; Miller et al. 2004, 2005; Pope et al. 2004; Samet et al. 2000a, 2000b). These studies uniformly estimated PM exposures using area-based arithmetic averaging or nearest-neighbor imputation—alternative methods that have important limitations (Moore and Carpenter 1999). Such limitations include the assumption of homogeneous exposures within study areas and the inability (or failure) to estimate exposures or associated PEs. For example, when daily exposure was of interest and there were no operating PM monitors with a study area, data pairs (daily PM concentrations, death counts) were unavailable in these studies. In addition, when longer-term (monthly to yearly) exposure was of interest, area aggregated exposures were based on available measurements within a given time frame. If there were five 24-hr measures in a month, for example, the monthly average exposure was calculated as the mean of the five readings. In contrast, our krigeing-based approach estimated daily mean exposures and SEs at geocoded addresses of participants and their examination sites across the contiguous United States that can be readily integrated over time with little influence of missing data. Studies in the geosciences have also found that krigeing provides consistently improved interpolation accuracy over traditional inverse-distance weighting and other, simpler spatial interpolation methods (Zimmerman 1999). Another important advantage of GIS-based estimation over the traditional area-average

---

Table 3. Comparison of estimated PM10 (µg/m3) at 94,135 geocoded addresses of WHI CT participant residences and examination sites using default and manually modified semivariograms.  

| Date       | PE  | RMSS  | SPE  | Default | Modified | Default | Modified | Default | Modified | Default | Modified | Default | Modified | Default | Modified | Default | Modified | Default | Modified | Default | Modified | Default | Modified |
|------------|-----|-------|------|---------|----------|---------|----------|---------|----------|---------|----------|---------|----------|---------|----------|---------|----------|---------|----------|---------|----------|---------|----------|
| 02/16/2000 | 0.0122 | −0.0099 | 5.034 | 1.037 | 0.0470 | 0.0021 |
| 03/05/2000 | 0.1660 | 0.0474 | 5.134 | 1.360 | 0.0469 | 0.0058 |
| 07/15/2000 | 0.5278 | 0.0193 | 5.564 | 1.180 | 0.0674 | −0.0024 |
| 08/07/2000 | 0.5252 | −0.1056 | 6.183 | 1.134 | 0.1417 | −0.0053 |
| 09/19/2000 | 0.7059 | 0.3651 | 4.744 | 1.146 | 0.0963 | 0.0142 |
| 10/29/2000 | 0.4590 | 0.0363 | 4.243 | 1.276 | 0.0780 | 0.0018 |

---

Table 4. Minima and maxima of measured and estimated PM10 (µg/m3) on the 22 days in 2000 when estimated values exceeded the range of measured values.  

| Date       | Minimum measured | Estimated from ordinary krings | Maximum measured | Estimated from ordinary krings |
|------------|------------------|-------------------------------|------------------|-------------------------------|
|            | Regular          | Lognormal                     | Regular          | Lognormal                     |
| 01/11      | 3.80             | −3.135                        | 5.535            |                               |
| 01/15      | 3.00             | 2.106                         | 11.756           |                               |
| 01/16      | 3.00             | 2.102                         | 9.150            |                               |
| 02/13      | 1.00             | −4.006                        | 9.388            |                               |
| 02/28      | 3.00             | −0.005                        | 7.196            |                               |
| 03/06      | 3.68             | −5.270                        | 7.281            |                               |
| 03/11      | 4.00             | 2.945                         | 9.064            |                               |
| 03/18      | 5.29             | 3.179                         | 8.841            |                               |
| 04/08      | 4.00             | −43.540                       | 9.097            |                               |
| 04/16      | 0.14             | −3.759                        | 0.901            |                               |
| 05/04      | 5.65             | −5.768                        | 14.550           |                               |
| 05/09      | 2.00             | −15.362                       | 10.889           |                               |
| 05/10      | 3.00             | −18.598                       | 13.805           |                               |
| 05/14      | 6.00             | 5.472                         | 6.175            |                               |
| 06/07      | 9.13             | −49.164                       | 18.164           |                               |
| 06/10      | 8.00             | 7.456                         | 8.224            |                               |
| 06/15      | 7.22             | 5.282                         | 12.582           |                               |
| 07/04      | 7.00             | 6.946                         | 9.128            |                               |
| 08/02      | 3.00             | −1.224                        | 16.587           |                               |
| 08/17      | 8.22             | 5.296                         | 7.132            |                               |
| 09/08      | 5.00             | 4.244                         | 5.889            |                               |
| 09/30      | 7.00             | 6.074                         | 11.686           |                               |

---

Environmental Health Perspectives • VOLUME 114 | NUMBER 9 | September 2006 1377
approach is the availability of both the location-specific estimated pollutant concentrations and their SEs.

Our goal in this study was to contribute methodologic and practical insights toward standardized, semiautomated GIS approaches to estimation of daily air pollution concentrations and their associated estimation errors. The air pollution data estimated using these approaches will support the Environmental Epidemiology of Arrhythmogenesis in WHI study (Whitsel 2006) examining the cardiac effects of air pollution in 68,133 postmenopausal women 50–79 years of age at baseline in the WHI CT (WHI Study Group 1998). Here we describe our experience resolving several important methodologic and practical issues in adopting a systematic, standardized, and semiautomated kriging approach to estimate daily air pollution concentrations and the associated estimation errors at geocoded addresses across the contiguous United States over 10 years.

We successfully downloaded from AQS the PM$_{10}$ and PM$_{2.5}$ raw data from 1993–2004. We then cleaned, calculated, and reconstructed site-specific daily PM concentration data ready for GIS applications. It is well known that the monitoring sites in AQS are not randomly distributed, which is one of the assumptions in kriging estimation, and the density of the monitoring sites is relatively low given the size of the contiguous United States. However, the AQS is the only currently available nationwide database. Our cross-validation studies suggest that the AQS data can be used as source data for kriging estimation of ambient pollution concentrations at various locations across the 48 contiguous states.

In this study, we performed cross-validation to assess the goodness of fit of various semivariogram and spatial models using four major parameters: the average PE, SPE, RMSS, and SE of estimation. Details can be found elsewhere (Webster and Oliver 2001), but it is worth noting that in addition to using the SE as a measure of the goodness of fit of a kriging model, one could improve the health effects models by incorporating SE in the models to account for the error in the estimation of location-specific PM concentrations. We consider this an important advantage of GIS-based estimation over the traditional area-average approach and are performing studies of using SE in health effects models.

We compared the performance of three widely spatial models (spherical, exponential, Gaussian) for PM$_{10}$ and PM$_{2.5}$ estimations using regular ordinary kriging on a national scale (Tables 1 and 2). In general, the cross-validation parameters suggest that all three models performed fairly well. Overall, the spherical model seemed to perform slightly better, consistent with the observation that the spatial distribution pattern of ambient air pollutants is closest to the assumption of the spherical model. The spherical model has been used most often in modeling spatially distributed data, providing a further rationale for its use in our large-scale population-based study of the health effects of PM. Furthermore, from the perspective of the cross-validation results, both average PE and average SPE are very close to 0, with a very narrow range of variation from the 366 daily cross-validations. These data support the overall validity of using kriging-based estimation approaches to estimate location-specific PM concentrations across the contiguous United States.

We completed an empirical analysis to investigate whether manually adjusting semivariogram parameters improves a) cross-validation parameters and b) estimated PM$_{10}$ concentrations and their SEs (Table 3). From these data, we conclude that manually adjusting semivariogram parameters improves cross-validation parameters. However, the application of such “improved” semivariograms to the estimation of PM$_{10}$ concentrations at geocoded locations across the United States did not necessarily provide better estimation of location-specific PM. Therefore, we recommend using the default-calculated semivariogram.

Semivariograms are sensitive to strong positive skewness. As a result, regular ordinary kriging can yield negative predicted values or values exceeding the range of the source data. Kriging works best if the input data have a normal distribution. One solution is to log-transform the input data—using “lognormal kriging.” In the ArcView software package, performing lognormal kriging is a standard option. This option log-transforms the input data to normalize its distribution and attenuate the impact of very large values. It also back-transforms the estimated values and the “unbiased” SE of the estimation to the original scale (Cressie 1993b; Johnston 2001). Our results comparing lognormal ordinary kriging versus regular-scale ordinary kriging suggest that lognormal ordinary kriging not

### Table 5. Means ± SDs of the cross-validation summary statistics from both ordinary and lognormal krigings, year 2000.

| Date       | Number of days | PM$_{10}$ | PM$_{2.5}$ |
|------------|----------------|-----------|------------|
| All days   | n = 366        | SPE       | RMSS       |
| 03/01      | 0.0106         | 0.0193    | −0.0238    |
| 03/06      | 0.0034         | 0.0320    | −0.0013    |
| 05/03      | 0.0159         | 0.0089    | 0.0456     |
| 04/06      | −0.0015        | 0.0308    | 0.0286     |
| 05/06      | −0.0052        | 0.0284    | −0.0420    |
| 06/06      | −0.0079        | 0.0150    | 0.0105     |
| 07/06      | 0.0031         | 0.0010    | 0.0083     |
| 08/06      | 0.0108         | 0.0220    | −0.0025    |
| 09/06      | 0.0053         | 0.0086    | −0.0013    |
| 10/06      | 0.0055         | 0.0245    | 0.0164     |
| 11/06      | 0.0190         | 0.0565    | 0.0364     |
| 12/06      | 0.0130         | 0.0193    | 0.0379     |
| Mean       | 0.0060         | 0.0190    | 0.0394     |
| Median     | 0.0054         | 0.0193    | 0.0594     |
| SD         | 0.0083         | 0.0150    | 0.0259     |

### Table 6. Comparisons of goodness of fit between national and regional scale krigings of the 12 days studied in 2000.

| Date       | National-scale | Regional-scale |
|------------|----------------|----------------|
| All days   | SPE            | RMSS           |
| 01/01      | 0.0106         | 0.0193         |
| 03/01      | 0.0106         | 0.0193         |
| 05/03      | 0.0159         | 0.0089         |
| 04/06      | −0.0015        | 0.0308         |
| 05/06      | −0.0052        | 0.0284         |
| 06/06      | −0.0079        | 0.0150         |
| 07/06      | 0.0031         | 0.0010         |
| 08/06      | 0.0108         | 0.0220         |
| 09/06      | 0.0053         | 0.0086         |
| 10/06      | 0.0055         | 0.0245         |
| 11/06      | 0.0190         | 0.0565         |
| 12/06      | 0.0130         | 0.0193         |
| Mean       | 0.0060         | 0.0190         |
| Median     | 0.0054         | 0.0193         |
| SD         | 0.0083         | 0.0150         |

Abbreviations: Natl, national-scale kriging; MN, kriging in middle north region; NE, kriging in northeast region; NW, kriging in northwest region; SE, kriging in southeast region; SW, kriging in southwest region.
of cross-validation results, both performed on PM concentrations within the range of the measured data for the days regular ordinary kriging yielded negative or “out of range” PM estimations, but also yielded a smaller average SE than did regular ordinary kriging and estimations. Therefore, our results support the use of lognormal ordinary kriging as an acceptable solution to the problem commonly posed by positively skewed distributions of environmental data.

Our comparisons of national- versus regional-scale kriging indicate that, in terms of cross-validation results, both performed similarly. However, such comparisons are based on krigings using the source data from optimal days (when > 900 sites across the country were reporting data), which account for only 17% of all days in a year. Therefore, there is additional justification for using national-scale kriging: Usually, there were very few operating sites within a region. On typical days—when only about 200 monitoring sites were operating—ability to derive stable and meaningful semivariograms was greatly impaired. Regional kriging also poses problems for estimation at locations near regional borders. For example, at locations within Washington State but near the Washington–Idaho border, regional kriging is based solely on PM10 concentrations in the “Washington/Oregon, Northern California” region. It is not based on PM10 concentrations measured immediately across the border in Idaho, despite the real possibility that they would have the largest weights in national-scale kriging estimation. For all these reasons, we recommend national-scale kriging.

Considering the number of study participants and the length of study period (1994–2003) for the Environmental Epidemiology of Arrhythmogenesis in WHI study, development of an automated procedure enabling large-scale daily krigings and semivariogram cross-validations was critical. In this study, we decided to use ArcView for predicting individuals’ PM exposure concentrations because of the flexibility it offers for automation. Because ArcView GIS relies on either the weighted least-squares method or visual adjustment to create semivariograms, we did not compare the relative performance of semivariograms generated using alternative methods such as maximum likelihood and restricted maximum likelihood. For generating semivariograms, we compared only three popular spatial models (spherical, exponential, and Gaussian). Our results, however, do not invalidate alternative spatial models (e.g., power). In the end, we selected the spherical model for our study because it is the most studied model, and its assumption pertaining to the spatial correlation of data is probably closest to our pollutant data. Furthermore, the spherical model seemed to perform as well as or slightly better than the remaining models in terms of cross-validation parameters.

We chose ordinary kriging instead of universal or simple kriging for several reasons. First, the assumption for simple kriging of a known mean concentration on any given day across space is not practical for our data. Although it may seem more appropriate because of the “varying mean” concentration across the contiguous U.S. assumption, universal kriging requires a predetermined set of “exploratory variables” to explain the varying means. The candidates, many of which are spatial variables, include emissions, land use, population, road network distribution, altitude, rainfall, latitude, climatology, and other quality data. Denby et al. (2005) recently recommended a method that uses measured concentration data in combination with some “exploratory variables” as suggested above. However, their approach may not be feasible for a national-scale study such as ours, because little guiding information is available as to how to identify a set of widely acceptable variables that can be applied to the entire nation. Moreover, even if we could identify a set of exploratory variables, we do not know the

Appendix 1. WHI Institutions and Investigators

| WHI Program Office, National Heart, Lung, and Blood Institute, Bethesda, MD | Barbara Alving, Jacques Rossouw, Shari Ludlam, Linda Potten, Joan McGowan, Leslie Ford, Nancy Geller |
| -- | -- |
| Clinical Coordinating Centers | Ross Prentice, Garnet Anderson, Andrea LaCroix, Charles L. Kooperberg, Ruth E. Patterson, Anne McTiernan, Shirley Barstoford, Sally Shumaker, Evan Stein, Steven Cummings |
| Fred Hutchinson Cancer Research Center, Seattle, WA | Sylvia Wasserman-Smith, Jennifer Hays, JoAnn Manson |
| Wake Forest University School of Medicine, Winston-Salem, NC | AnnLouise R. Assaf, Lawrence Phillips, Judith Hsia |
| Medical Research Labs, Highland Heights, KY | Rowan Chlebowski, Evelyn Whillock, Bette Caan |
| University of California—San Francisco, San Francisco, CA | Jane Morley Knitchen, Barbara V. Howard, Linda Van Hom, Rebecca Jackson, Henry Black |
| Albert Einstein College of Medicine, Bronx, NY | Marcia L. Stefanick, Dorothy Lane, Cora E. Lewis, Tamara Bassford, Jean Wachtowski-Wende, John Robbins, F. Allan Hubbell, Howard Judd, Robert D. Langer, Margery Gass, Mariam Limacher, David Curb, Robert Wallace, Judith Ockene, Norman Lasser, Mary Jo O’Sullivan, Karen Margolis, Robert Brunner, Gerardo Heiss, Lewis Kuller, Karen C. Johnson, Robert Bryska, Gloria E. Sarto, Denise Bonds, Susan Hendrix |
| Baylor College of Medicine, Houston, TX | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| Brigham and Women’s Hospital, Harvard Medical School, Boston, MA | Shari Ludlam, Linda Potten, Joan McGowan, Leslie Ford, Nancy Geller |
| Brown University, Providence, RI | Ross Prentice, Garnet Anderson, Andrea LaCroix, Charles L. Kooperberg, Ruth E. Patterson, Anne McTiernan, Shirley Barstoford, Sally Shumaker, Evan Stein, Steven Cummings |
| Emory University, Atlanta, GA | Sylvia Wasserman-Smith, Jennifer Hays, JoAnn Manson |
| George Washington University Medical Center, Washington, DC | AnnLouise R. Assaf, Lawrence Phillips, Judith Hsia |
| Harbor-UCLA Research and Education Institute, Torrance, CA | Rowan Chlebowski, Evelyn Whillock, Bette Caan |
| Kaiser Permanente Center for Health Research, Portland, OR | Jane Morley Knitchen, Barbara V. Howard, Linda Van Hom, Rebecca Jackson, Henry Black |
| Kaiser Permanente Division of Research, Oakland, CA | Marcia L. Stefanick, Dorothy Lane, Cora E. Lewis, Tamara Bassford, Jean Wachtowski-Wende, John Robbins, F. Allan Hubbell, Howard Judd, Robert D. Langer, Margery Gass, Mariam Limacher, David Curb, Robert Wallace, Judith Ockene, Norman Lasser, Mary Jo O’Sullivan, Karen Margolis, Robert Brunner, Gerardo Heiss, Lewis Kuller, Karen C. Johnson, Robert Bryska, Gloria E. Sarto, Denise Bonds, Susan Hendrix |
| Medical College of Wisconsin, Milwaukee, WI | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| MedStar Research Institute/Howard University, Washington, DC | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| Northwestern University, Chicago/Evanston, IL | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| Ohio State University, Columbus, OH | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| Rush Medical Center, Chicago, IL | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| Stanford Prevention Research Center, Stanford, CA | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| State University of New York–Stony Brook, Stony Brook, NY | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| University of Alabama at Birmingham, Birmingham, AL | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| University of Arizona, Tucson/Phoenix, AZ | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| University at Buffalo, Buffalo, NY | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| University of California–Los Angeles, Los Angeles, CA | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| University of California–San Diego, La Jolla/Chula Vista, CA | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| University of Cincinnati, Cincinnati, OH | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| University of Florida, Gainesville/Jacksonville, FL | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| University of Hawaii, Honolulu, HI | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| University of Iowa, Iowa City/Davenport, IA | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| University of Massachusetts/Fallon Clinic, Worcester, MA | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| University of Medicine and Dentistry of New Jersey, Newark, NJ | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| University of Miami, Miami, FL | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| University of Minnesota, Minneapolis, MN | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| University of Nevada, Reno, NV | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| University of North Carolina–Chapel Hill, Chapel Hill, NC | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| University of Pittsburgh, Pittsburgh, PA | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| University of Tennessee, Memphis, TN | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| University of Texas Health Science Center, San Antonio, TX | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| University of Wisconsin, Madison, WI | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| Wake Forest University School of Medicine, Winston-Salem, NC | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
| Wayne State University School of Medicine/Hutzel Hospital, Detroit, MI | Barbara Grady, Anne McTiernan, Anne Schairer, Leslie Ford, Joan McGowan, Leslie Ford, Nancy Geller |
forms or shapes of their independent and joint relationships to the air pollution measures. Further studies that involve large-scale national data using universal kriging are still needed. In this study, we empirically tested whether the non-constant mean assumption for universal kriging was needed; we performed five regional ordinary krigings so that different parts of the country would assume a different mean PM concentration. Our data suggested that regional and national ordinary kriging performed similarly. Therefore, our data indirectly validated and supported the use of national ordinary kriging.

Although the primary objective of our study is to assess the short-term relationship between PM and cardiac responses, the proposed kriging method also enables us to calculate the long-term cumulative exposure of an individual by taking into account the change of his or her residences over time, because the WHI study recorded the residential location history over 10 years. Nevertheless, from the environmental perspective, an inherited limitation of the kriging-based approach is that the estimates of the PM concentrations will provide only surrogates, or the best guesses, of the true exposure levels at the locations of interest. Thus, the accuracy of the estimations depends highly on the quality of the measured data and their spatial correlation. Even if the estimations were made with a high level of confidence, they cannot be directly interpreted as the true individual-level exposures. However, to correlate individual level cardiac responses with a surrogate of location-specific exposure, our approach represents one of the best available methods for a large-scale population-based study.

In summary, our investigation of GIS approaches for estimating daily mean geocoded location-specific air pollutant concentrations and their SES supports the use of a spherical model to perform lognormal ordinary kriging on a national scale. Our findings also support the use of default-generated semivariograms (estimated using the weighted least-squares method) without visual adjustment. We developed a semiautomated program to access and execute ArcView to implement these approaches for large-scale daily kriging estimations and semivariogram cross-validations. Detailed information about this program can be obtained on request.

REFERENCES

Abby DE, Moore J, Petersen F, Beeson L. 1991. Estimating cumulative ambient concentrations of air pollutants: description and precision of methods used for an epidemiological study. Arch Environ Health 46:281–287.

Abby DE, Nishino N, McDonnell WF, Burchette RJ, Knutsen SF, Lawrence Beeson W, et al. 1999. Long-term inhalable particles and other air pollutants related to mortality in non-smokers. Am J Respir Crit Care Med 159:373–382.

Cressie NAC. 1993a. Geostatistics. In: Statistics for Spatial Data (Cressie NAC, ed). New York:John Wiley & Sons, 29–104.

Cressie NAC. 1993b. Spatial prediction and kriging. In: Statistics for Spatial Data (Cressie NAC, ed). New York:John Wiley & Sons, 105–209.

Davis JC. 2002. Analysis of sequences of data. In: Statistics and Data Analysis in Geography (Davis JC, ed). 3rd ed. New York:John Wiley & Sons, 159–292.

Denby B, Walker SE, Horâlek DJ, Eben K, Fiala J. 2005. Interpolation and Assimilation Methods for European Scale Air Quality Assessment and Mapping. Part I: Review and Recommendations. European Topic Centre on Air and Climate Change Technical Paper 2005/7. Bilthoven, the Netherlands:European Topic Centre on Air and Climate Change.

Dockery DW, Pope CA III, Xu X, Spengler JD, Ware JH, Fay ME, et al. 1993. An association between air pollution and mortality in six U.S. cities. N Engl J Med 329:1753–1759.

ESRI Inc. 2001. Using analytic tools when generating surfaces. In: Geostatistical Analyst Extension. Redlands, CA:ESRI Inc., 128–167.

Gribov A, Krivaruchko K, Hoef JMV. 2004. Modeling the semivariogram: new approach, methods comparison and case study. In: Stochastic Modeling and Geostatistics—Principles, Methods and Case Studies (Yarus JM, Chambers RL, eds). Vol 2, 2nd Ed, New York:John Wiley & Sons, 149–192.

Katsouyanni K, Touloumi G, Samoli E, Gryparis A, Le Tertre A, Katsouyanni K, Schwartz J, Spix C, Touloumi G, Zmirou D, et al. 1996. Short term effects of air pollution and mortality: results from 29 European cities within the APHEA2 project. Epidemiology 12:521–531.

Johnston K. 2001. Lognormal linear kriging. In: Using ArcGIS Geostatistical Analyst (ESRI, ed). Redlands, CA:ESRI Press, 247–273.

Katsouyanni K, Schwartz J, Spix C, Touliou G, Zmirou D, Zanobetti A, et al. 1996. Short term effects of air pollution on health: a European approach using epidemiologic time series data: the APHEA protocol. J Epidemiol Community Health 50(suppl 1):S12–S18.

Kriging DG. 1966. Two-dimensional weighted moving average [Abstract]. Circulation 111:237.

Legendre P, Fortin M-J. 1989. Spatial pattern and ecological analysis. Vegetatio 80:107–138.

Liao D, Yuan Y, Whitsett EA, Zheng Z, Heiss G, Chinchilli VM, et al. 2004. Association of higher levels of ambient criteria pollutants with impaired cardiac autonomic control: a population-based study. Am J Epidemiol 159:768–777.

Liao D, Heiss G, Chinchilli VM, Duan Y, Folsom AR, Lin H, et al. 2005a. Association of criteria pollutants with plasma hemostatic/inflammatory markers—a population-based study. J Expo Anal Environ Epidemiol 15:319–328.

Liao D, Peuquet DJ, Duan Y, Whitsett EA, Dou J, Smith RL, et al. 2005b. Estimation of residential-level ambient PM concentrations from the U.S. EPA’s air quality monitoring database [Abstract]. Epidemiology 16(5):S27–S28.

Liao D, Peuquet DJ, Duan Y, dou J, Smith RL, Whitsett EA, et al. 2005c. GIS approaches for estimation of residential-level ambient PM concentrations [Abstract]. Epidemiology 16(5):S28.

Miller KA, Siscovick DS, Shepard L, Anderson GL, Kaufman JD. 2004. Air pollution and cardiovascular disease events in the women’s health initiative observational (WHI-DS) study [Abstract]. Circulation 109:E189.

Miller KA, Siscovick DS, Shepard L, Sheppard K, Anderson GL, Kaufman JD. 2005. Effect of traditional risk factors on the association of air pollution and incident cardiovascular disease in the women’s health initiative observational study (WHI-OS) [Abstract]. Circulation 111:E228–E229.

Moore DA, Carpenter TE. 1999. Spatial analytical methods and geographic information systems: use in health research and epidemiology. Epidemiol Rev 21:143–161.

Mtza PE, Burnett RT, Thurston GD, Thun MJ, Calle EE, Krewski D, et al. 2004. Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of dis ease. Circulation 109:71–77.

Samet JM, Dominici F, Curriero FC, Coursac I, Zeger SL. 2000a. Fine particulate air pollution and mortality in 20 U.S. cities, 1987–1994. N Engl J Med 343:1742–1749.

Jain X, Olea RA, Yu YS. 1996. Semivariogram modeling by weighted least squares. Comput Geosci 22:387–397.

Johnston K. 2001. Lognormal linear kriging. In: Using ArcGIS Geostatistical Analyst (ESRI, ed). Redlands, CA:ESRI Press, 247–273.

Liao et al. | 1380 | VOLUME 114 | NUMBER 9 | September 2006 | Environmental Health Perspectives