Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Environmental Pollution 125 (2003) 369–383
ENVIRONMENTAL POLLUTION
www.elsevier.com/locate/envpol

A critical examination of ozone mapping from a spatial-scale perspective

Jeremy E. Diem*

Department of Anthropology and Geography, Georgia State University, Atlanta, GA 30303-3083, USA

Received 28 October 2002; accepted 7 March 2003

"Capsule": *The purpose of the review is to encourage researchers to consider spatial scale in their work.*

Abstract

Following the establishment of point measurements of ground-level ozone concentrations have been attempts by many researchers to develop ozone surfaces. This paper offers a critique of ozone-mapping endeavors, while also empirically exploring the operational scale of ground-level ozone. The following issues are discussed: aspects of spatial scale; the spatial complexity of ground-level ozone concentrations; and the problems of previous attempts at ozone mapping. Most ozone-mapping studies are beset with at least one of the following core problems: spatial-scale violations; an improper evaluation of surfaces; inaccurate surfaces; and the inappropriate use of surfaces in certain analyses. The major recommendations to researchers are to acknowledge spatial scale (especially operational scale), understand the prerequisites of surface-generating techniques, and to evaluate the resultant ozone surface properly.

Keywords: Ozone; Spatial scale; Operational scale; Model evaluation; Atmospheric pollution

1. Introduction

Ozone's harmful impact on crops, forests, and human health qualifies it as a serious air pollutant in many countries. In the United States, for example, ambient ozone concentrations have been regulated since the early 1970s by the US Environmental Protection Agency. At the present time, ozone is measured at thousands of stationary monitors scattered across the nation, most of which are located in urban areas. Since point measurements are limited in their applicability across Cartesian space, the mapping of ozone levels is a critical procedure. The resulting surfaces are often used to inform decisions regarding the protection of public health and welfare from elevated ozone levels.[1] The generation of surfaces requires an explicit consideration of spatial scale, but recognition of spatial scale is lacking in the ozone-mapping literature.

The aim of this paper is to present spatial scale in the context of ozone mapping in order to initiate increased consideration of spatial scale by researchers in the future. Therefore, this paper aims to: (1) explore spatial scale; (2) speculate about the spatial complexity of ground-level ozone concentrations; (3) discuss the problems of previous attempts at ozone mapping; and (4) provide recommendation for future ozone-mapping projects. This paper reviews the ozone-mapping literature and presents original research to illustrate the earlier objectives.[2]

2. The meanings of spatial scale

Spatial scale is the spatial equivalent of temporal scale. For example, most researchers studying atmospheric phenomena consider spatial scale and temporal scale simultaneously, thus spatio-temporal processes are categorized as follows: microscale (up to tens of meters; up to several hours), mesoscale (tens of kilometers; several hours to several days), synoptic scale (hundreds to thousands of kilometers; several days to several weeks),

* Tel.: +1-404-651-1759; fax: +1-404-651-3235.
E-mail address: gegjed@langate.gsu.edu (J. E. Diem).

[1] In this paper, a surface refers to a continuous surface, which can be treated as the boundary of a three-dimensional figure having height values (Z-values) distributed without interruption.

[2] Every effort has been made to provide a comprehensive review of published literature presenting to varying degrees ozone-mapping techniques and applications. Even so, it is probably not a fully comprehensive review, thus the absence of one or more studies should be anticipated.

0269-7491/03/$ - see front matter
doi:10.1016/S0269-7491(03)00110-6

and global scale (thousands of kilometers; several weeks to several years). Spatial scale has four different meanings: geographic scale, measurement scale, operational scale, and cartographic scale (Cao and Lam, 1997).[3] The remainder of this paper focuses on geographic scale, measurement scale, and operational scale; cartographic scale is not examined, because it has minimal relevance to the generation of ozone surfaces.

2.1. Geographic scale

Geographic scale is the spatial extent of a study region. For instance, the development of ozone surfaces for the Atlanta metropolitan area and the North American continent would be considered small- and large-scale studies, respectively. Ozone surfaces have been developed for geographic scales ranging from approximately 32–8,000,000 km^2 (Table 1). The importance of geographical scale results from its connection to measurement scale and operational scale, especially in the context of surface generation.

2.2. Measurement scale

Measurement scale is the sampling interval used in a study. In fact, in the case of point measurements (e.g. ozone samples), it is more appropriate to refer to the sampling interval rather than measurement scale. Within nearly all domains, ozone monitors are not located at fixed intervals. For example, the distance between ozone monitors in the Atlanta metropolitan area is highly variable with nearest-neighbor distances ranging from approximately 8 km to nearly 60 km (from Tolbert et al., 2000). The density of ozone monitors, which is 13 monitors per 16,000 km^2, is equivalent to having a monitor located every 35 km throughout the area (Table 1). The sampling interval changes with a change in the geographic scale, especially when the geographic scale is originally a city (e.g. City of Atlanta) and is increased subsequently to include surrounding areas with fewer monitors.

Resulting from differences in the number of ozone monitors and the aforementioned range in geographic scales, the sampling intervals for the ozone-mapping studies range from approximately 1.5 km to more than 400 km (Table 1). In the context of ozone mapping, a small sampling interval is preferable for two reasons: it facilitates accurate spatial interpolation (refer to Section 2.3.3) and it is essential for assessing operational scale.

[3] Cartographic scale pertains to the representative fraction of a map (e.g. 1/10,000,000 is a small-scale map), and is thus related directly to spatial resolution [i.e. smallest distinguishable part in an object (Tobler, 1988)], which in turn is associated with operational scale.

2.3. Operational scale

2.3.1. Overview of operational scale and semivariograms

Operational scale or "scale of action" refers to the spatial extent at which a particular phenomenon (e.g. ground-level ozone) operates. Operational scale and spatial complexity are inversely related, thus a phenomenon has a small operational scale if its surface is dominated by small-scale (geographic) variations.

An essential tool for examining operational scales is an experimental semivariogram, which is a plot of semivariance against spatial lag.[4] Most semivariograms have a nugget, range, and sill (Fig. 1). The nugget is a function of measurement errors and differences in values among locations separated by distances much shorter than the sample spacing (Burrough and McDonnell, 1998). The sill is the part of the semivariogram having the most semivariance, which indicates that positive spatial autocorrelation among the sample points has disappeared.[5] Finally, the range is the spatial lag at which the sill is reached. Within the range, positive spatial autocorrelation increases with an increase in spatial lag, while outside the range spatial autocorrelation is negligible and there is little change in spatial autocorrelation with an increase in spatial lag.

Semivariograms are appropriate for identifying the range of spatial scales within which the variable is spatially dependent (Bian and Walsh, 1993), and operational scales are represented by "break points" in the semivariogram (Lam and Quattrochi, 1992). The range is a major "break point," thus the range is equivalent to one of the operational scales.

For many phenomena the requirement of values for many spatial locations prohibits the construction of a robust semivariogram that can be used to determine operational scale. Burroughs and McDonnell (1998) note that in order to produce even a stable semivariogram a sampling network should have at least 50–100 points. There are an infinite number of pairs at each spatial lag (or distance class), thus the sample pairs within each distance class should be a reasonable representation of a population of pairs. The central limit theorem implies that at least 100 pairs should be used to estimate the semivariance for each distance class (Griffith and Lane, 1999), while Cressie (1991) recommends a minimum of 30 pairs per distance class. Spatial modeling can be used to overcome the earlier sample-based limitation, and the next section provides an example of modeled ozone and operational scale.

[4] The semivariance equation is as follows:

$$\gamma = \frac{1}{2n_{\text{pairs}}} * \sum_{i=1}^{i=n} (x_i - x_{i+h})^2$$

where n is the number of pairs at spatial lag h, x_i is the value at location i, and x_{i+h} is the value at a location within spatial lag h of x_i.

[5] Positive spatial autocorrelation occurs when values at nearby points are more similar than are values at distant points.

Table 1
Characteristics of published studies that have spatially interpolated/predicted ambient ozone levels. Relative error is calculated by dividing the RMSE (root mean squared error) by the average observed value; those values are presented in the manuscripts either in tabular or graphical form. If RMSE is not available, MAE (mean absolute error) is used instead. MAE yields an underestimate of relative error

Author(s)	Location	Geographic scale	Number of monitors	Sampling interval	Mapping method	Spatial resolution	Relative error (%)
Adams et al. (1985)	Conterminous USA	~8,000,000	[a]	[a]	K	[a]	[a]
Lefohn et al. (1987)	Conterminous USA	~8,000,000	[a]	[a]	OK	~50	22[b]
Lefohn et al. (1988)	Southeastern USA	~4,000,000	[a]	[a]	OK	~50	20[b]
Abbey et al. (1991a)	California, USA	~424,000	126	[a]	IDW	[a]	[a]
Abbey et al. (1991b)	California, USA	~424,000	126	[a]	IDW	[a]	[a]
McKendry (1993)	Montreal, Quebec, Canada	~2000	9	~15	C	[a]	[a]
Bower et al. (1994)	United Kingdom	~245,000	18	~120	DW	25	[a]
Brown et al. (1994)	Ontario, Canada	~1,070,000	7	~400	BI	[a]	[a]
Casado et al. (1994)	Southeastern USA	~700,000	29	~30	OK	8	[a]
Guttorp et al. (1994)	Sacramento, CA, USA	~10,600	17	~25	STM	[a]	[a]
Loibl et al. (1994)	Austria	~84,000	114	~27	NLR, K	1	19[b,c]
Rosenbaum et al. (1994)	Conterminous USA	~8,000,000	565	~120	PGO	80	[a]
Westenbarger and Frisvold (1994)	Eastern USA	~3,120,000	[a]	[a]	K	~50	[a]
Brown et al. (1995)	United Kingdom	~245,000 km^2	17	~120 km	LIR	1	[a]
Duddek et al. (1995)	Ontario, Canada	~1,070,000	21	~225	BI	[a]	[a]
Fowler et al. (1995)	United Kingdom	~245,000	17	~120	LIR	1	[a]
Korc (1996)	Los Angeles, CA, USA	~48,000	38	~35 km	IDW	10 km	[a]
Loibl and Smidt (1996)	Austria	~84,000	130	~25 km	NLR, K	[a]	[a]
Liu and Rossini (1996)	Toronto, Ontario, Canada	~22,500	19	~34	K	[a]	54[b]
McNair et al. (1996)	Los Angeles, CA, USA	~25,000	37	~25	IDW	[a]	36[b]
Pauly and Drüeke (1996)	Trier, Germany	~300	20	~4	UK	[a]	13[b]
Carroll et al. (1997)	Houston, TX, USA	~4600	12	~20	STM	[a]	[a]
de Leeuw and van Zantvoort (1997)	The Netherlands	~34,000	38	~28	IDW	5	22%[b]
Georgopoulos et al. (1997)	Mid-Atlantic USA	~75,000	38	~45	OK	5	[a]
Hogsett et al. (1997)	Eastern USA	~4,700,000	294	~125	OA	20	[a]
Lefohn et al. (1997)	Southeastern USA	~750,000	[a]	[a]	K	~50	[a]
Liu et al. (1997)	San Diego, CA, USA	~32	13	~1.5	K, IDW	[a]	[a]
Godzik (1997)	Kraków, Poland	~3254	18	~13	K	[a]	[a]
Phillips et al. (1997)	Southeastern USA	~870,000	235	~60	CK, IDW, K	20 km	24[b,d]
Christakos and Vyas (1998a)	Eastern USA	~2,800,000	1228	~48	S/TRF	[a]	1.4[e]
Christakos and Vyas (1998b)	Eastern USA	~2,800,000	1228	~48	S/TRF	25 km	[a]
Meiring et al. (1998)	Sacramento, CA, USA	~21,000	32	~25 km	STM	12	[a]
Mulholland et al. (1998)	Atlanta, GA, USA	~16,000	10	~40	UK	3	[a]
Nikiforov et al. (1998)	Conterminous USA	~8,000,000	1112	~85 km	LR	[a]	13[f]
Sun et al. (1998)	Ontario, Canada	~1,070,000	21	~225	BI	[a]	[a]
Zidek et al. (1998)	Southern Ontario, Canada	[a]	22	[a]	BI	[a]	[a]
Hopkins et al. (1999)	Houston, TX, USA	~50,000	20	~50	IDW, K	[a]	[a]
Royle and Berlinger (1999)	Midwestern USA	~980,0000	147	~80	HSM	~50	[a]
Duc et al. (2000)	Sydney, Australia	~4000 km^2	13	~17	K	[a]	[a]
Kuik et al. (2000)	The Netherlands	~34,000	[a]	[a]	IDW	[a]	[a]
Pissimanis et al. (2000)	Athens, Greece	~1,750	9	~14 km	[a]	[a]	[a]
Tolbert et al. (2000)	Atlanta, GA, USA	~16,000	10	~40	UK	3	[a]
Laurence et al. (2001)	Eastern USA	~2,800,000	3	~305	[a]	10	[a]
Le et al. (2001)	SW British Columbia, Canada	~4800	23	~14	BI	[a]	[a]
Lee and Hogsett (2001)	Western USA	~3,420,000	204	~130	LOR	2	25[b,d]
Schichtel and Husar (2001)	Conterminous USA	~8,000,000	1415	~75	IDW	40	[a]
Varns et al. (2001)	Dallas, TX, USA	~25,000	36	~26	IDW	[a]	[a]
Bytnerowicz et al. (2002)	Carpathian Mtns., Europe	~158,000	32	~70	CK	[a]	20[f]
Coyle et al. (2002)	United Kingdom	~245,000	80	~55	S, LIR	1 km	[a]
Diem and Comrie (2002a,b)	Tucson, AZ, USA	~11,000	8	~37	MLR	0.5	7[g]

BI = Bayesian Interpolation; C = Contouring; CK = Cokriging; D = Distance Weighting; HSM = Hierarchical Spatial Model; IDW = Inverse-Distance Weighting; IDW^2 = Inverse-Distance Weighting Squared; K = Kriging; LIR = Linear Regression; LOR = Loess Regression; MLR = Multiple Linear Regression; NLR = Nonlinear Regression; OA = Overlay Analysis; OK = Ordinary Kriging; PGO = Point-to-Grid Overlay; S = Spline; S/TRF = Spatiotemporal Random Field; STM = Spatio-Temporal Model; and UK = Universal Kriging.

[a] Not reported or cannot be determined from information presented in the article.
[b] Cross validated.
[c] Based on values at 10 sites; unknown if those sites were selected at random.
[d] Used MAE instead of RMSE.
[e] Based on values at four sites; unknown if those sites were selected at random.
[f] Not cross validated.
[g] Not a true cross validation.

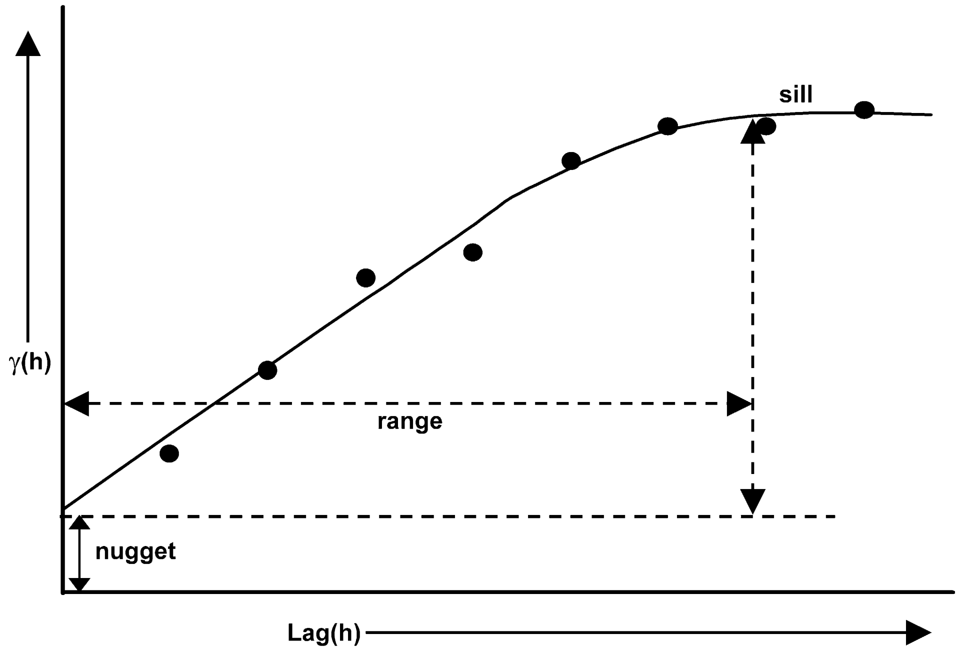

Fig. 1. Idealized semivariogram showing the nugget, range, and sill.

2.3.2. Operational scale of ground-level ozone

As with most other spatial phenomena and certainly with other atmospheric pollutants, the operational scale of ground-level ozone is unknown. An exceptionally dense network of monitors (i.e. sampling interval on the order of 1 km) is needed to assess operational scale. As can be gleaned from Table 1, sufficiently dense networks are not available anywhere, especially for networks of continuous ozone monitors. As noted earlier, preliminary assessments of operational scale can be achieved with modeled data, and the chosen data for this study come from a spatial-modeling study described in detail in Diem and Comrie (2002a). In that study, ozone concentrations are estimated at a spatial resolution of 0.5 km throughout the Tucson, Arizona region. With an overall relative error of 7%, the surfaces are accurate enough to allow further examinations of the estimated ozone concentrations. Although over 700 summer days are considered in Diem and Comrie (2002a), data from only 5 days are analyzed in this study.[6]

For this investigation, the Tucson region is divided into three subregions: upwind, source-intensive, and downwind. The subregions are delineated to enable statistical analyses across the region, which, in turn, increase the robustness of the results and potential for applicability to other regions. The area of interest is centered on the City of Tucson, which contains a bulk of the reliable, predicted ozone concentrations (Diem and Comrie, 2002a). The subregions reflect their geographical relationship to the city. Since the source-intensive subregion is situated over the city, it contains a large proportion of the Tucson region's ozone precursor emissions (Diem and Comrie, 2002b). Northwest of the city is the upwind subregion, while east of the city is the downwind subregion.

Semivariograms, which are constructed for each cluster/subregion combination, are used as the primary tool for assessing operational scale. Each of the 15 combinations has at least 456 sample points, thereby easily enabling the construction of stable semivariograms. In addition, correlograms are constructed and used subsequently to determine the sampling interval at which positive spatial autocorrelation is no longer statistically significant ($\alpha = 0.05$).[7] The correlogram-derived information is used to corroborate findings from the semivariogram examinations. Semivariance and spatial

[6] Each day epitomizes one of the five ozone clusters discussed in Diem and Comrie (2002a). The clusters have the following characteristics: clusters 1, 2, and 3 occur most frequently in mid- to late-summer and have relatively high ozone concentrations; and clusters 4 and 5 occur most frequently in early-summer and have relatively low ozone concentrations. Each day examined in this study also has relatively low prediction errors at the core ozone monitoring sites. The five representative days are as follows: 2 July 1999 (Cluster 1); 26 August 1995 (Cluster 2); 14 August 1998 (Cluster 3); 12 June 1998 (Cluster 4); and 18 June 1997 (Cluster 5).

[7] A correlogram is a plot of spatial autocorrelation against spatial lag. The correlogram equation is as follows:

$$I = \frac{n_{\text{individuals}}}{n_{\text{pairs}}} * \frac{n_{\text{pairs}} * \sum_{i=1}^{i=n}\sum_{j=1}^{j=n}(x_i - \overline{x})(x_j - \overline{x})}{\sum_{i=1}^{i=n}(x_i - \overline{x})}$$

where I is Moran's I, n is the number of locations, n_{pairs} is the number of pairs within a certain spatial lag, x_i is the value at location i, x_j is the value at location j, and $\overline{x}$ is the global mean. Moran's I values are bounded by -1 and $+1$ with -1 indicating extremely-strong, negative, spatial autocorrelation and $+1$ indicating extremely-strong, positive, spatial autocorrelation.

autocorrelation are calculated for each 0.5-km interval up to a spatial lag of 12 km. The resultant semivariograms and correlograms are shown in Figs. 2 and 3.

The semivariograms and correlograms illustrate quantitatively the strong spatial dependence among the predicted ozone concentrations. Each semivariogram exhibits multiple "break points", with the first point representing theoretically the operational scale for a given subregion/cluster combination. Each correlogram has a point at which the spatial autocorrelation coefficient (i.e. Moran's I) becomes statistically equivalent to zero. This point is the distance at which positive spatial autocorrelation ceases, thus it also represents the operational scale.

The semivariogram and correlogram analyses provide converging results, with the operational scale values ranging from 1.5 to 8.0 km (Table 2). The overall operational scale, which is weighted by the frequency of occurrence of the five clusters, is between 5.0 and 6.0 km. These values are not direct artifacts of the subjectivity-laden process of selecting predictor variables for the multiple linear regression (MLR) modeling (Diem and Comrie, 2002a). If the spatial influences of the predictor variables are weighted by their respective standardized coefficients and if the resultant average value equals the operational scale, then the operational scale values may not be valid. As it stands, the average value is 8.3 km, thus the operational scale values are not biased greatly by the initial selection of predictor variables.

It is not the intent of this paper to discuss each of the estimated operational scale values. All values have been presented to illustrate the robust nature of the resulting average value (~5.5 km). Nevertheless, to provide further clarification of operational scale assessment, attention is focused on the semivariograms and correlograms of Cluster 1. This cluster exemplifies the typical spatial dependencies of the upwind, source-intensive, and downwind subregions. Analyses of the semivariograms (Fig. 2a–c) reveal "break points" at 6.0, 5.0, and 5.5 km, respectively, for the three subregions. Concerning the correlograms, statistically-significant positive spatial autocorrelation does not exist after 7.0, 6.0, and 5.5 km for the subregions. Therefore, based on the earlier information the average operational scale for Cluster 1 is approximately 6 km.

Although the estimation of operational scale is based on analyses of modeled data and should thus be treated with some skepticism, the results do agree with findings from other studies. Pauly and Drüeke (1996) construct a semivariogram using daily ozone concentrations measured at 20 ozone passive samplers in the city of Trier, Germany. The range of the semivariogram is approximately 6 km. These results, however, are not especially strong because the sampling network consisted of only 20 sample points. McNair et al. (1996) note that for the Los Angeles basin significant variability in observed pollutant concentrations does occur at spatial lags comparable to and smaller than the airshed model's 5-km grid cells. For example, two ozone monitoring sites in the Long Beach area located 4.8 km apart had ozone measurements that differed by up to 50%. In addition, significant differences existed between several pairs of monitors located within 10 km of each other in the Los Angeles basin (McNair et al., 1996). Based on the earlier information, it appears that the operational scale of ground-level ozone in a typical metropolitan area is less than 10 km. This spatial complexity is controlled by local sources of ozone precursors (especially nitrogen oxides), topography, micro-meteorology (e.g. urban canyon effect), and varying rates of ozone deposition. In remote, forested areas, which are typically devoid of significant, spatially distributed sources (e.g. motor vehicles) of nitrogen oxides, one would expect the operational scale of ground-level ozone to be much larger than the scale in a metropolitan area.

Table 2
Estimated operational scale for each subregion/cluster combination

Subregion/cluster	Operational scale (km)
Upwind/1	6.0
Source-Intensive/1	5.5
Downwind/1	5.5
Upwind/2	6.0
Source-Intensive/2	8.0
Downwind/2	8.0
Upwind/3	5.0
Source-Intensive/3	7.0
Downwind/3	5.5
Upwind/4	6.5
Source-Intensive/4	4.0
Downwind/4	5.0
Upwind/5	1.5
Source-Intensive/5	6.0
Downwind/5	6.0
Overall	5.6

2.3.3. Surface generation

Various techniques produce surfaces of a phenomenon by estimating/predicting values at unsampled locations based on measurements at sample points. Common surface-generating techniques include inverse distance weighting (IDW), kriging (i.e. ordinary kriging, universal kriging, and cokriging), and regression (e.g. MLR). In addition to each method having its own set of drawbacks (Lam, 1983; Diem and Comrie, 2002a), the proper employment of any method requires careful consideration of spatial scale, particularly operational scale.

Kriging is the most popular technique used to create ozone surfaces. It is employed in 19 of the 50 ozone-mapping studies listed in Table 1. Kriging estimates values at unsampled locations using weights reflecting

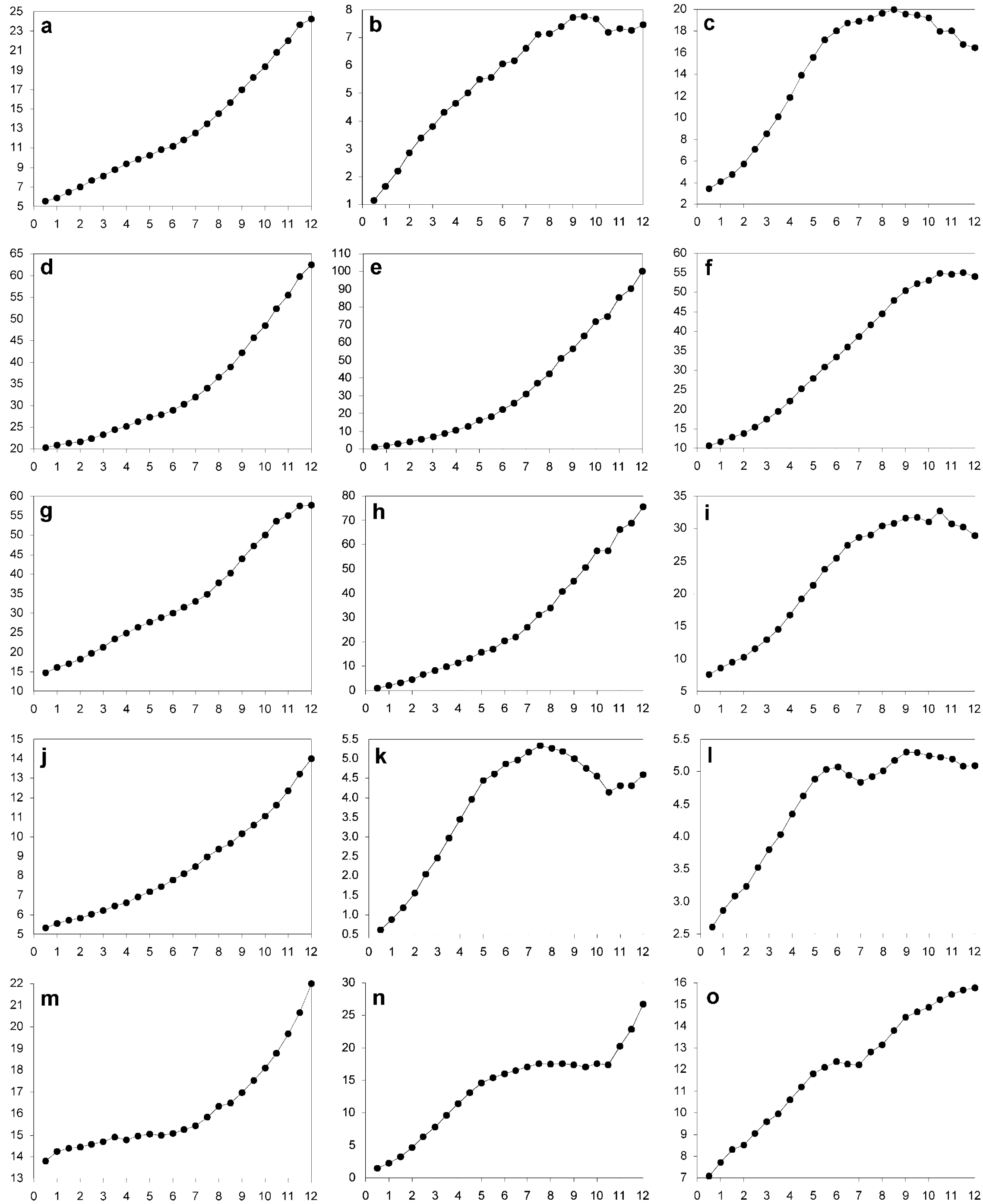

Fig. 2. Semivariograms for the following subregion/cluster combinations: (a) upwind/1; (b) source-intensive/1; (c) downwind/1; (d) upwind/2; (e) source-intensive/2; (f) downwind/2; (g) upwind/3; (h) source-intensive/3; (i) downwind/3; (j) upwind/4; (k) source-intensive/4; (l) downwind/4; (m) upwind/5; (n) source-intensive/5; and (o) downwind/5. Semivariance (ppb) is along the y-axis and spatial lag (km) is along the x-axis.

the correlation between data at two sample locations or between a sample location and the location to be estimated (Myers, 1991). The success of kriging in any mapping exercise depends greatly on the estimation and modeling of the semivariogram (Burroughs and McDonnell, 1998). Not only should the semivariogram be constructed preferably with information from at least 100 pairs of observations per distance class, but it also should exhibit low semivariances (i.e. strong spatial autocorrelation) at small spatial lags. Concomitant with the strong spatial autocorrelation is the presence of a range, which, as noted previously, represents the operational scale. Only spatial lags exhibiting spatial dependence should be considered during kriging; therefore, the operational scale represents the appropriate search radius to be used during kriging.

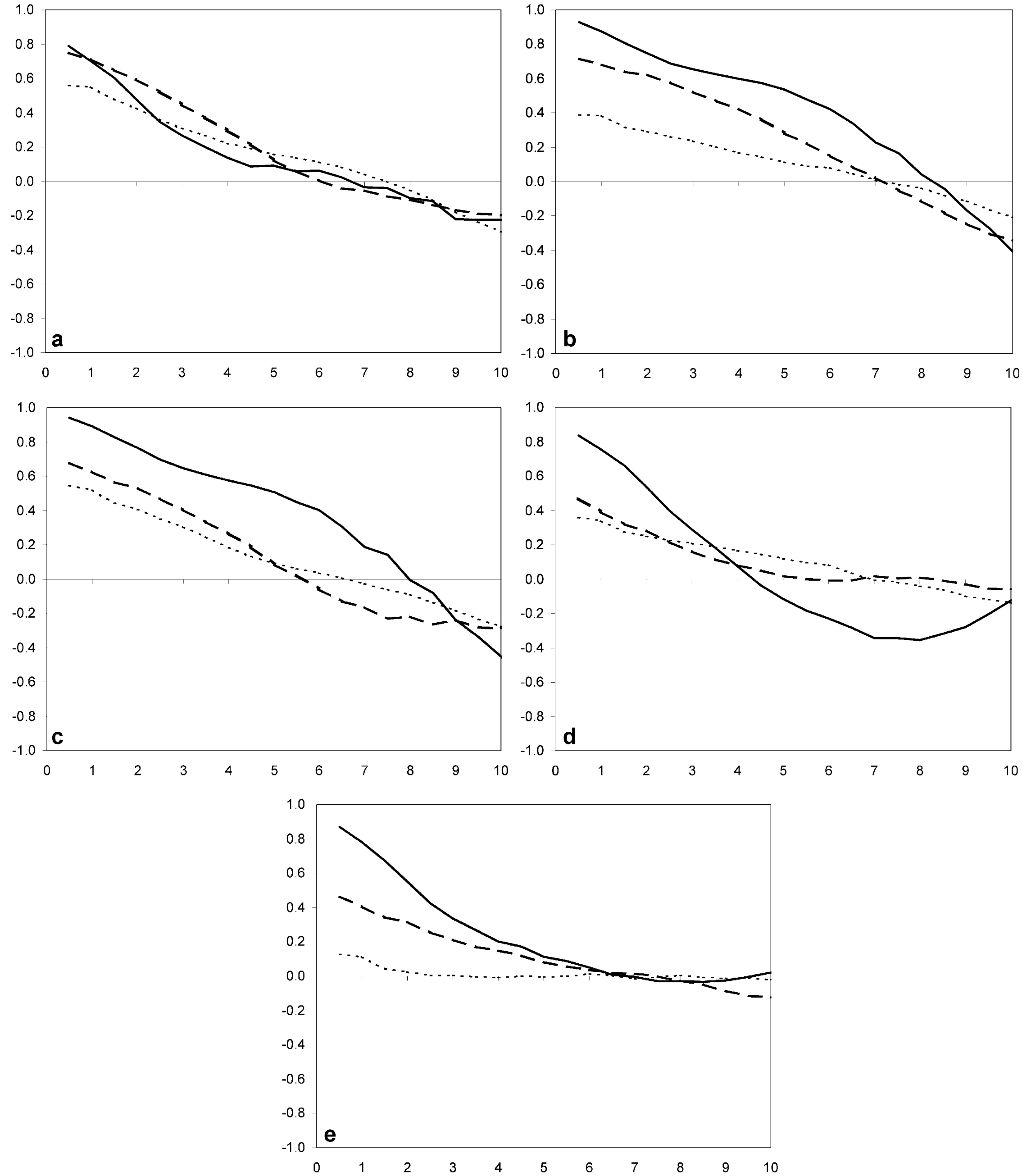

Fig. 3. Correlograms for the upwind (······), source-intensive (——), and downwind (– –) subregions for (a) Cluster 1, (b) Cluster 2, (c) Cluster 3, (d) Cluster 4, and (e) Cluster 5. Moran's *I* is along the *y*-axis and spatial lag (km) is along the *x*-axis.

2.3.4. Operational scale and the optimal sampling interval

The operational scale represents the largest sampling interval that can be used to ascertain the spatial complexity of a surface. For example, assuming that 5.5 km is a valid estimate for the operational scale of ozone and that most of the Tucson region's population exists within a 950 km^2 area, then at least 30 monitors are needed to capture sufficiently the spatial complexity of ozone concentrations in the heavily populated portion of the region. Theoretically, within the 5.5-km interval there is considerable redundancy among measured ozone concentrations. Creating an ozone surface via kriging would be difficult with the above sampling scheme, because many more monitors are needed to construct a semivariogram. If a semivariogram is constructed with information from 30 monitors in the Tucson area, the semivariogram should exhibit negligible spatial dependence.

An alternative mapping method is MLR. Diem and Comrie (2002a) have shown that MLR is a suitable spatial-prediction technique provided continuous surfaces of variables associated with ambient ozone levels are available. The operational scale of ozone relates readily to MLR. Since minimal redundancy should be present in any MLR analysis, the smallest sampling interval for MLR modeling is the operational scale. For the Tucson region, a sampling interval of 5.5 km would be optimal for the prediction of ozone concentrations using MLR.

If spatial-interpolation techniques, such as kriging, are to be used to produce an ozone surface, an optimal sampling interval, which is smaller or equal to the operational scale, needs to be determined. It is well known that a decrease in the sampling interval results in an increase in the number of data points and thus an increased likelihood of producing a more spatially complex (and accurate) surface. The surface will have an increasing magnitude of small-scale variations in ozone concentrations. It is not practical, however, to sample ozone concentrations at extremely small intervals, for ozone monitors are expensive to operate. The costs of materials, sampler maintenance and retrieval in the field, and laboratory analyses make even ozone passive samplers costly. Therefore, the optimum sampling interval is the largest interval that produces a surface with the required minimum interpolation error (Burroughs and McDonnell, 1998). The optimal sampling interval can be determined using results from an intensive field survey. For example, ozone passive samplers could be located 1 km apart on a regular grid covering 100 km^2. The samplers would collect daily ozone levels, thereby enabling the construction of a semivariogram. The sampling network would yield 4950 total pairs of samplers, and over 300 and 100 pairs would be available at the 1 and 10 km lags, respectively. An extremely stable semivariogram would be possible. The optimal sampling interval can be determined from a stable semivariogram using a procedure described in Atkinson et al. (1990), which is as follows:

- Compute the maximum kriging variance for sampling intervals at 1-km increments from 1 to 10 km. The maximum kriging variance is an output of the kriging process, and, in this example, it would occur at locations farthest from the samplers.
- Determine the maximum allowable kriging variance and then select the largest sampling interval having a kriging variance less than or equal to this value. One might decide that the maximum allowable kriging variance is 10% of the average measured ozone concentration.

The earlier method is just one of the many possible ways of assessing optimal sampling intervals. All methods require a dense sampling network and the knowledge of a desired minimum accuracy of the resultant surface.

To summarize, knowledge of the operational scale of ground-level ozone can help determine the optimal sampling interval. In addition, the optimal sampling interval represents the highest possible spatial resolution of the resultant ozone surface. For nearly all studies, it is ideal to produce an ozone surface with a relatively high spatial resolution.

3. Problems associated with ozone mapping

A review of the ozone-mapping literature reveals many problems in the development, evaluation, and use of ozone surfaces. The problems are ubiquitous and can be placed into the following categories:

- Spatial-scale violations.
- Improper evaluation of surfaces.
- Inaccurate surfaces.
- Inappropriate use of surfaces in certain analyses.

Critical mapping-related problems conducive to severe negative consequences are discussed in the following subsections.

3.1. Spatial-scale violations

The five principal ways in which many published studies have committed spatial-scale violations are as follows: (1) disregard the linkages among geographic scale, sample size, and spatial resolution; (2) disregard spatial autocorrelation issues; (3) improper integration of data with multiple spatial resolutions; and (4) treating surfaces as ‘‘reality’’.

3.1.1. Disregard the linkages among geographic scale, sample size, and spatial resolution

Studies that disregard the linkages among geographic scale, sample size, and spatial resolution fall into two categories: studies that employ a spatial-interpolation technique despite having an ozone monitoring network that is not large enough to support the technique; and studies that develop ozone surfaces with spatial resolutions higher than is appropriate. Nearly every study has committed some aspect of this violation. As can be discerned through the examination of Table 1, the number of ozone monitors within any given geographic domain varies substantially. All studies that use kriging to develop an ozone surface despite having a small number of sample points commit this first violation. Multiple studies (Pauly and Drüeke, 1996; Godzik, 1997; Liu et al., 1997; Mulholland et al., 1998; Hopkins et al., 1999; Duc et al., 2000; Bytnerowicz et al., 2002) employ geostatistical techniques even though the semivariogram is null owing to its creation from an inadequate number of sample points. For example, Mulholland et al. (1998)

and Tolbert al. (2000) produce a kriged surface based on a semivariogram developed using information from just 10 to 13 monitors in the Atlanta metropolitan area.

It is also common for an ozone surface to be produced at a falsely high spatial resolution. Kriging is especially vulnerable to this violation if the modeled semivariogram is based on extrapolations. Such an extrapolation occurs between the smallest sampling interval and the origin of the semivariogram. This violation is committed by Loibl et al. (1994), Mulholland et al. (1998), and Tolbert et al. (2000). Loibl et al. (1994) achieve 1-km resolution with 114 monitors in an approximately 80,000 km^2 domain using nonlinear regression and kriging; however, it is doubtful that there is even a single pair of monitors within one kilometer of each other. Mulholland et al. (1998) and Tolbert et al. (2000) achieve 3-km resolution with only 10–13 monitors in a 26,000 km^2 domain. As mentioned previously, the shortest distance between monitors is approximately 8 km. This violation may be more prevalent than is described here, since most studies fail to note the spatial resolution of the resultant ozone surface (Table 1).

3.1.2. Disregard spatial-autocorrelation issues

Those studies that employ IDW and kriging are particularly vulnerable to spatial-autocorrelation issues. IDW, which assigns more weight to nearby points than to distant points (Myers, 1991), requires spatial autocorrelation. Kriging also depends on spatial autocorrelation; hence, as stated previously, a critical component of kriging is the semivariogram. A useful semivariogram cannot be developed without the presence of spatial autocorrelation; the degree of spatial autocorrelation determines how successful spatial interpolation will be (Griffith and Lane, 1999).

Spatial autocorrelation is not mentioned in most IDW and kriging studies despite its central role. Those studies that do mention semivariograms acknowledge spatial autocorrelation implicitly. Semivariograms and correlograms are needed to assess the efficacy of surface-generating techniques, such as IDW and kriging, whose success can be predicted somewhat by the presence or absence of spatial autocorrelation among measured ozone concentrations.

3.1.3. Improper integration of data with multiple spatial resolutions and unknown accuracies

Several studies (Loibl et al., 1994; Fowler et al., 1995; Loibl and Smidt, 1996; Pauly and Drüeke, 1996; Hogsett et al., 1997; Phillips et al., 1997; Lee and Hogsett, 2001; Bytnerowicz et al., 2002; Coyle et al., 2002; Diem and Comrie, 2002a) develop ozone surfaces using ancillary data, such as digital elevation models (DEMs), pollutant emissions surfaces, and meteorological surfaces. The two major problems that stem from the use of ancillary surfaces are inappropriate "downscaling" of the data and possible errors in the surfaces.

An example of inappropriate "downscaling" is presented in Hogsett et al. (1997) and Phillips et al. (1997). In both studies, an index of ozone-exposure potential is constructed using county-resolution data of anthropogenic nitrogen oxides (NO_x) emissions data. The studies simply resample the county-resolution NO_x data to 20-km resolution to equal the resolution of surfaces such as temperature, wind direction, and elevation.

None of the studies that employ ancillary surfaces attach any accuracy value to those surfaces. Even though Hogsett et al. (1997) and Phillips et al. (1997) note that meteorological surfaces are created using a form of IDW, an evaluation of those surfaces apparently is not conducted. The emissions surfaces used in Hogsett et al. (1997), Phillips et al. (1997), Coyle et al. (2002), and Diem and Comrie (2002a) probably have the most errors of all the ancillary surfaces, yet the possibility of error-laden emissions surfaces is never questioned in the literature. Diem and Comrie (2002b) discuss the development of spatially resolved databases of pollutant emissions, and a brief glimpse of the data and methods reveals a high probability for grid cells having exceedingly inaccurate estimates of emissions.

The inadequacies of the ancillary surfaces are propagated into the ozone surfaces when the ancillary and derived surfaces are eventually used in spatial applications. For example, Phillips et al. (1997) use their derived index as the covariate in a cokriging procedure despite the fact that cokriging in this instance requires an accurate, spatially continuous variable that is highly correlated with ozone concentrations. The accuracy of the index should have been assessed before using it in a cokriging procedure (refer to Section 3.2 for a discussion of evaluation metrics). Bytnerowicz et al. (2002) use invalid ancillary data when attempting to interpolate ozone concentrations throughout the Carpathian Mountains. The researchers use cokriging with elevation as the covariate even though "elevation does not necessarily affect [ozone] concentrations in the mountainous terrain (Bytnerowicz et al., 2002: 23)." Finally, Diem and Comrie (2002a) employ up to 10 derived surfaces, which are mostly emissions-related, as predictor variables in multiple linear regression equations used to predict ozone concentrations. This is a risky practice, for there is undoubtedly much error present in those predictor surfaces. The technique in this instance, however, is not affected greatly by those errors and thus the resultant ozone surfaces have an acceptable level of accuracy (refer to Section 3.2 and Table 1).

3.1.4. Treating surfaces as "reality"

Another problem associated with the application of surface-generating techniques, especially spatial interpolation (e.g. IDW and kriging), is that the resulting ozone surface is an abstraction of "reality" yet the surface is sometimes considered to be a true reflection of "reality." Spatial-interpolation techniques do not capture the spatial complexity of ozone surfaces. In addition to the prominent inaccuracies of ozone surfaces as discussed in Section 3.3, ozone surfaces produced with spatial-interpolation techniques are considerably "smoother" than the true (but unknown) surfaces. The structure of the ozone-monitoring network has a major influence on one's ability to develop a "realistic" surface. For example, Bower et al. (1994: 123) note that the present ozone monitoring network in the United Kingdom "has not been optimised to reconstruct in detail the national concentration field for ozone." The limitations of monitoring networks and interpolation techniques is not recognized universally, for Westenbarger and Frisvold (1994: 2898) state that their results "provide a detailed picture of the magnitude and distribution of acid precipitation and ozone pollution in the eastern United States".

3.2. Improper evaluation of surfaces

Many ozone-mapping studies provide insufficient metrics for the evaluation of results from ozone mapping. This evaluation "is the process of examining and appraising the performance by comparing the model's concentration estimates to measured air quality data (Fox, 1981: 600)." Some researchers do not even appear to be aware of the evaluation process, for they do not provide any evaluation metrics (refer to Table 1). The dearth of proper evaluation metrics in the literature may result from the researchers' unfamiliarity with error reporting, or it may be a preëmptive measure that enables avoiding discussion of accuracy. On a positive note, nearly all researchers who do attempt to assess accuracy do so with cross-validation procedures.[8] Common metrics reported in the ozone-mapping literature include mean biased error (MBE), mean absolute error (MAE), root mean squared error (RMSE), correlation coefficient (r), and coefficient of determination (r^2). MBE, which is the average value of the residuals, indicates the degree of over- or under-prediction. Ideally, MBE should be approximately equal to zero. MAE and RMSE are similar in that they are measures of overall error; RMSE is always larger than MAE owing to its squaring of large residuals. The final two statistics, r and r^2, should not be part of an array of model performance metrics, because the magnitudes of r and r^2 are not related consistently to the accuracy of prediction (Willmott, 1982).

[8] Cross-validation is the process of obtaining predictions at a monitor that are not influenced by observations at that monitor. Cross-validation produces predicted values that are independent of observed values, which then enables a more realistic assessment of the accuracy of a surface.

Every study should provide several core evaluation metrics to facilitate a comparison among ozone surfaces and thus a comparison of interpolation and prediction techniques. Willmott (1982) suggests reporting the following statistics: $\bar{P}$, $\bar{O}$, s_p, s_o, a, b, RMSE, $RMSE_s$, $RMSE_u$, and d. $\bar{O}$ is the mean of the observed values. $\bar{P}$ is the mean of the predicted values. Values for s_o and s_p are the standard deviations of the observed and predicted values, respectively. The intercept and slope of the least-squares regression line ($\hat{P}_i = a + bO_i$) are represented by a and b, respectively. RMSE can be disaggregated into systematic ($RMSE_s$) and unsystematic ($RMSE_u$) components. In addition, the number of cases n as well as the relative error $RMSE/\bar{O}$ should be reported. An accurate surface has the following characteristics:

- $\bar{P} \approx \bar{O}$;
- s_p approaches s_o;
- relative error approaches 0;
- $RMSE_u$ approaches RMSE (i.e. most of the difference between predicted and observed values is not derived from the mapping method); and
- d approaches 1.

Surfaces having statistics that deviate substantially from the earlier guidelines should be treated as invalid surfaces.

3.3. Inaccurate surfaces

Most mapping studies that do report evaluation metrics have relatively inaccurate ozone surfaces. Relative error can be calculated for 13 of the 50 studies. In this instance, the relative error does not enable a sound cross-comparison of results, because occasionally MAE is reported instead of RMSE or the number of cases is extremely small or both. Calculating relative error using MAE results in a smaller value than if RMSE were used, while a small number of cases do not provide an adequate sample of the population of observation–prediction pairs. Therefore, excluding Christakos and Vyas (1998a) and Loibl et al. (1994), which provide observed and predicted values for only 4 and 10 cases, respectively, the range in relative error is from 7 to 54%. The relative error of 7% reported in Diem and Comrie (2002a) may be a slight underestimate, since the value is not derived from a true spatio-temporal cross-validation. Conversely, the relative error of 54% reported in Liu and Rossini (1996) may be a substantial overestimate; the predicted values are based on interpolation using values at continuous ozone monitors and the observed values are from passive samplers.

Surfaces created using IDW or kriging tend to have a relative error exceeding 20%, thereby providing some evidence of the inappropriateness of these techniques for the spatial interpolation of ozone levels. A possible reason for the low-quality surfaces resulting from kriging is that semivariograms do not receive enough attention from researchers. Semivariograms encapsulate a substantial amount of spatial-scale information. Disregarding a semivariogram signifies the mistreatment of spatial scale.

3.4. Inappropriate use of surfaces in certain analyses

The problems of ozone mapping are magnified when the resultant surfaces are used in further analyses, such as estimating the impacts of ozone exposure. Examples of such analyses are presented in Adams et al. (1985), Abbey et al. (1991b), Westenbarger and Frisvold (1994), Brown et al. (1995), Duddek et al. (1995), Fowler et al. (1995), Korc (1996), Carroll et al. (1997), de Leeuw and van Zantvoort (1997), Georgopoulos et al. (1997), Hogsett et al. (1997), Lefohn et al. (1997), Christakos and Vyas (1998b), Coyle et al. (2002), Diem and Comrie (2002a), Laurence et al. (2001), Loibl and Smidt (1996), Mulholland et al. (1998), Zidek et al. (1998), Kuik et al. (2000), and Tolbert et al. (2000). Most of these studies have not considered the inherent problems of ozone mapping and the subsequent use of values from the resultant ozone surfaces. The two factors noticeable in those studies that can lead to slightly misleading and possibly erroneous conclusions are as follows: (1) the accuracy of ozone surfaces is low or unknown; and (2) the disregard of the problems associated with areal units. Those two factors are discussed below.

3.4.1. Accuracy of ozone surfaces is low or unknown

A major problem with most ozone exposure and response studies is that they are used to determine the risk to crops, forests, and people based on ozone surfaces with poor or unknown accuracies. Of the 22 studies noted above, only de Leeuw and van Zantvoort (1997) and Diem and Comrie (2002a) reveal the accuracies of their respective ozone surfaces. This is troubling, because results from such studies may have a major influence on air quality policy decisions. If the accuracy of an ozone surface is not reported, one is forced to treat any subsequent results with a considerable amount of doubt. In fact, the results from de Leeuw and van Zantvoort (1997) may also be highly questionable owing to the low accuracy (i.e. $>20\%$ relative error) of the ozone surfaces, while the true cross-validated accuracies of the ozone surfaces presented in Diem and Comrie (2002a) are unknown.

3.4.2. Disregard of the problems associated with areal units

If ozone surfaces are employed in exposure–response studies, the importance of the spatial resolution of the surface increases. The operational scale should guide the spatial resolution, for, ideally, the largest possible areal unit of an ozone surface is equivalent to the operational scale of ozone. Major problems can arise by not treating the characteristics of areal units as important determinants of the results from a study. This is part of the Modifiable Areal Unit Problem (MAUP). When the MAUP is present, conclusions spawned from the linkage of ozone surfaces to other surfaces are questionable based solely on the nature of areal units. The MAUP consists of the scale effect and the zoning effect: the scale effect refers to the inconsistency of analytical or statistical results derived from data representing different levels of spatial partitioning for the same area; and the zoning effect refers to the variability of analytical or statistical results derived from data for the same region, but partitioned in different ways, with the number of areal units in different partitioning schemes being the same (Wong, 1996). Previous research (e.g. Openshaw and Taylor, 1979; Fotheringham and Wong, 1991) on the MAUP indicates that correlation coefficients increase with increased aggregation; hence, the association between ozone levels and hypothesized responses in any domain should decrease in significance with a decrease in size of the areal units. Conversely, the validity of results decreases with a decrease in spatial resolution of the surfaces involved. Therefore, one way to minimize the MAUP is to use data from the individual or the most disaggregated level (Wong, 1996).

The MAUP is inherently present during estimates of forest-biomass loss (Hogsett et al., 1997), forest exposure (Loibl and Smidt, 1996; Lefohn et al., 1997), tree growth (Laurence et al., 2001), estimates of the areal coverage of high ozone concentrations (Coyle et al., 2002; De Leeuw and van Zantvoort, 1997; Fowler et al., 1995; Mulholland et al., 1998), estimates of agricultural yield (Adams et al., 1985; Westenbarger and Frisvold, 1994; Brown et al., 1995; Kuik et al., 2000), estimates of population exposure (Abbey et al., 1991b; Carroll et al., 1997; Georgopoulos et al., 1997; Korc, 1997; Mulholland et al., 1998; Diem and Comrie, 2002a), and estimates of associations between ozone levels and health impacts (Duddek et al., 1995; Mulholland et al., 1998; Zidek et al., 1998; Tolbert et al., 2000). In the earlier studies, ozone surfaces are created with a specified spatial resolution, which is not always reported. Consequently, the researchers assume that there is no spatial variation in ozone concentrations within the areal units. This assumption can lead to erroneous results if the areal unit is relatively large with respect to the operational scale of the variable. For example, if the spatial resolution of an ozone surface is 50 km, then the results of the analysis rest on the assumption that there is no significant variation among ozone concentrations in each 2500 km^2 unit. Based on results presented in this paper concerning the operational scale of ground-level ozone, all studies that attempt to relate ozone concentrations in a surface with a spatial resolution greater than 10 km probably will yield problematic conclusions. The operational scales of the associated variables (e.g. human population) are equally important. Results from the earlier studies would all change, sometimes dramatically, with a change in the spatial resolutions of the variables, especially ozone, as well as changes in the spatial partitioning of a domain (e.g. using census blocks instead of zip-code polygons).

The MAUP is an especially important factor to consider when assessing the results of studies attempting to establish causal relationships between ambient ozone concentrations and health problems. With respect to ozone surfaces, spurious correlations are likely to result between ozone concentrations at the areal-unit level and various human-health problems. In fact, personal ozone exposure differs dramatically from ozone levels measured at stationary, outdoor monitors (Liu et al., 1997), and this difference is postulated to be controlled by micro-environmental factors, such as those present within automobiles and buildings (Liu et al., 1995). Even an extremely accurate high-resolution ozone surface, which is not a common entity, may not adequately reflect individual-level ozone exposure.

Examples of spurious correlations exist in Duddek et al. (2000), Mulholland et al. (1998), and Tolbert et al. (1999). Scale-related problems are never mentioned in the three papers, yet, in all three studies, an attempt is made to establish a causal link between ambient ozone concentrations and respiratory ailments. The following problems plague each study even before the actual exposure–response analysis is conducted: (1) the ozone surface has an unknown accuracy; and (2) the final resolution of the ozone surface is much larger than the suspected operational scale of ozone. The spatial resolution of the eventual ozone surface in Duddek et al. (1995) is approximately 225 km; the size of the areal units may be over 1500 times larger than the appropriate size. Although the 10-km spatial resolution of the eventual ozone surfaces used in Mulholland et al. (1998) and Tolbert et al. (2000) is relatively fine, the surfaces' overall accuracies are probably poor resulting from a misuse of kriging (refer to Sections 3.1 and 3.3). For all three studies, the researchers assume that all variables (e.g. ozone, respiratory morbidity, adolescent population, etc.) do not vary significantly within each areal unit, even though in the case of ozone in Duddek et al. (1995) this assumption does not hold. Duddek et al. (1995) do not find a significant relationship between respiratory morbidity and air pollution, but even if they did the relationship could be challenged easily. The other two studies consider their results more compelling. Mulholland et al. (1998) determine that a positive ozone-asthma association exists in the form of a 4% increase in the emergency-room presentation rate per 20-ppb (parts per billion) increase in ambient ozone concentrations, while Tolbert et al. (2000: 808) state that their results "add to the body of evidence that supports an association of air pollution with exacerbation of asthma." Results presented in these studies are dubious; the results would be considerably more valid if analyses would have been conducted with individual-level data. The studies do provide baseline information for future research exploring the MAUP in the context of ozone exposure and health responses.

4. Recommendations

The development of a meaningful ozone surface is a difficult task. It comprises multiple steps, each of which requires knowledge of various assumptions and rules. What follows are some recommendations to researchers who are embarking on an ozone-mapping project. The following recommendations reflect the findings presented in this paper:

- Acknowledge spatial scale (i.e. geographic scale, sampling interval, and operational scale).
- Understand the requirements and limitations of the various surface-generating techniques.
- Select a surface-generating technique suitable for the task at hand.
  - Only use kriging if a robust semivariogram can be developed.
  - If the scarcity of ozone monitors prohibits the use of a technique, investigate multivariate statistical techniques that incorporate ancillary data.
- Provide the spatial resolution of the ozone surface.
- Present the overall accuracy of the ozone surface.
  - The following evaluation statistics should be provided: $\bar{P}$, $\bar{O}$, s_p, s_o, a, b, RMSE, RMSE/$\bar{O}$, RMSE_s, RMSE_u, and d.
- Include a scale bar on the map so viewers can relate map distance to actual distance (i.e. cartographic scale).
- Address and, if possible, minimize the MAUP when relating the ozone surface to other variables.
- In exposure–response studies, use data specific to individuals.
- If no appropriate surface-generating techniques are available, then forego the development of a surface and use a more suitable form of cartographic display, such as a proportional-symbol map (Fig. 4).

To repeat partially what is provided earlier, the strongest recommendation that can be made to researchers is to understand the prerequisites of a sur-

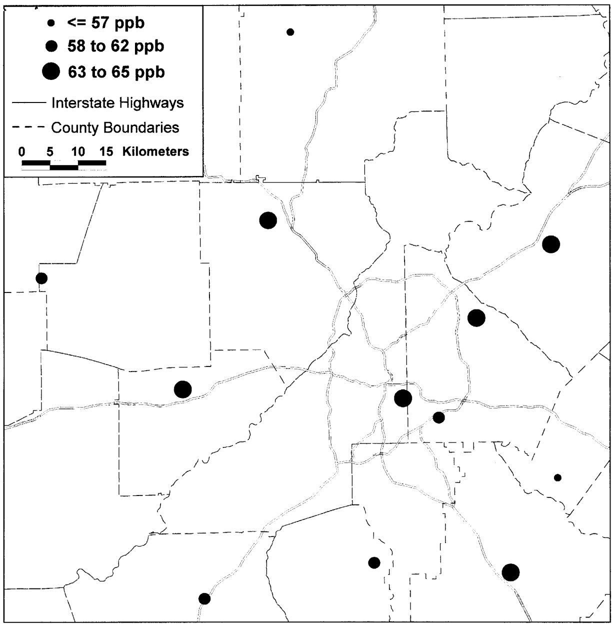

Fig. 4. Proportional-symbol map of average hourly concentrations in the Atlanta metropolitan area in 2000.

face-generating technique before employing it to produce an ozone surface.

5. Conclusions

This paper has examined the ozone-mapping literature from a spatial-scale perspective, and several problem areas have been revealed. It should be emphasized that despite the critical nature of this paper all attempts at creating ozone surfaces represent positive advances in the understanding of spatial variations in ground-level ozone concentrations. Information found in the ozone-mapping literature is needed for understanding the role of spatial scale in the appropriateness of surface-generating techniques.

Of the several types of spatial scale pertinent to ozone mapping, this paper has focused on the operational scale of ozone. Operational scale and spatial complexity are inversely related. Empirical results presented in this paper and results from other studies reveal that the operational scale of ground-level ozone in metropolitan areas is most likely less than 10 km and possibly as small as 5 km. The analysis presented in this paper involved examining the spatial dependence of ozone concentrations as predicted by a regression-based modeling approach. The assumed operational scale of ground-level ozone implies that the ozone landscape has substantial relief and thus is far more complex than is reflected in the work of most ozone researchers.

Almost every ozone-mapping study is plagued with one or more major problems, which include spatial-scale violations, an improper evaluation of surfaces, inaccurate surfaces, and the inappropriate use of surfaces in certain analyses. Spatial-scale violations are ubiquitous and consist of the following: (1) disregard the linkages among geographic scale, sample size, and spatial resolution; (2) disregard spatial-autocorrelation issues; (3) improper integration of data with multiple spatial resolutions; and (4) treating surfaces as "reality." Few studies report any accuracy-assessment statistics, and even fewer accurate ozone surfaces are produced. Nearly all the surfaces are inappropriate for further analyses (e.g. estimates of population exposure) owing to the accuracies of surfaces being low or unknown, the assumption of ozone homogeneity within areal units, and the MAUP.

Researchers involved in ozone-mapping projects should be fully aware of the influence of spatial scale (geographic, measurement, and operational) on the mapping methodology and results. Surface-generating techniques should be selected carefully, and the chosen technique should not be in conflict with any aspect of spatial scale. Ozone surfaces should be created only if it is technically sound to do so, and once it is created the accuracy of the surface should be evaluated using a standard suite of evaluation metrics. An enhanced awareness of the nature of ozone surfaces is needed if the predicted ozone concentrations are to be related to other variables. Therefore, researchers should be aware of the MAUP so that potentially spurious results are not considered legitimate.

Ozone mapping is important, for ozone maps can influence decisions concerning air-quality policy, which, in turn, affect the attitudes and behaviors of the general public. Unfortunately, maps in the form of surfaces have an immense amount of power, and one can be easily tempted to treat the surface as a perfect representation of reality. An example of the potential power of ozone surfaces involves the US Environmental Protection Agency's Ozone Mapping Project. This project, which is part of the agency's EMPACT (Environmental Monitoring for Public Access and Community Tracking) program, provides the general public with real-time ozone maps for various domains via the Internet. The goal of the project is to create maps "that provide communities with real-time information about ozone pollution in an easy-to-understand pictorial format (US EPA, 1999: 2)." This kind of project is important for public awareness of ozone; however, the methods used to produce those maps require the scrutiny outlined in this paper.

Kriging and IDW are the two most common spatial-interpolation techniques employed in the project, and the paucity of ozone measurements, especially when considered at a small geographic scale, such as a metropolitan area, precludes the employment of these techniques. The accuracy of the surfaces is not presented; therefore, users are led to believe that the maps are an adequate representation of the "true" ozone surface and that the uncertainties in ozone estimates do not vary across a domain. In some metropolitan areas, ozone maps are shown in newspapers and on television. As a result, the maps are seen by hundreds of thousands to millions of people. An alternative is to provide a proportional-symbol map of ozone concentrations (refer to Fig. 4). Because this map only provides monitor-specific ozone concentrations and is thus not an ozone surface, it does not provide false information concerning ozone concentrations at monitor-less locations. For instance, viewers of the map will not be led to believe that a specific ozone concentration exists at their place of residence.

Finally, this paper has addressed the need for a rigorous assessment of the operational scale of ground-level ozone. Knowledge of operational scale can enhance greatly the practice of ozone mapping. Optimal ozone-monitoring networks can be developed with the operational scale as a reference sampling interval. Operational scale also represents the largest allowable areal unit of an ozone surface. The acquisition of operational-scale information for ozone is a future research endeavor that would yield unprecedented benefits for the ozone-research community.

References

Abbey, D.E., Moore, J., Petersen, F., Beeson, W.L., 1991a. Estimating cumulative ambient concentrations of air pollutants: description and precision of methods used for an epidemiological study. Archives of Environmental Health 46 (5), 281–287.

Abbey, D.E., Mills, P.K., Peterson, F.F., Beeson, W.L., 1991b. Long-term ambient concentrations of total suspended particulates and oxidants as related to incidence of chronic disease in California Seventh-Day Adventists. Environmental Health Perspectives 94, 43–50.

Adams, R.M., Hamilton, S.A., McCarl, B.A., 1985. An assessment of the economic effects of ozone on US agriculture. Journal of the Air Pollution Control Association 35 (9), 938–943.

Atkinson, P.M., Curran, P.J., Webster, R., 1990. Sampling remotely sensed imagery for storage, retrieval, and reconstruction. The Professional Geographer 42 (3), 345–353.

Bian, L., Walsh, S.J., 1993. Scale dependencies of vegetation and topography in a mountainous environment of Montana. The Professional Geographer 45 (1), 1–11.

Bower, J.S., Stevenson, K.J., Broughton, G.F.J., Lampert, J.E., Sweeney, B.P., Wilken, J., 1994. Assessing recent surface ozone concentrations in the UK. Atmospheric Environment 28 (1), 115–128.

Brown, P.J., Le, N.D., Zidek, J.V., 1994. Multivariate spatial interpolation and exposure to air pollutants. The Canadian Journal of Statistics 22 (4), 489–509.

Brown, M., Cox, R., Bull, K.R., Dykr, H., Sanders, G., Fowler, D., Smith, R., Ashmore, M.R., 1995. Quantifying the fine scale (1 km×1 km) exposure, dose and effects of ozone: Part 2. Estimating yield losses for agricultural crops. Water, Air and Soil Pollution 85 (3), 1485–1490.

Bytnerowicz, A.B., Godzik, B., W. Fraczek, W., Grodzinska, K, Krywult, M, Badea, O., Barancok, P., Blum, O., Cerny, M., Godzik, S., Mankovska, B., Manning, W., Moravcik, P., Musselman, R., Oszlanyi, J., Postelnicu, D., Szdzuj, J., Varsavova, M., Zotal, M., 2002. Distribution of ozone and other air pollutants in forests of the Carpathian Mountains of central Europe. Environmental Pollution 116 (1), 3–25.

Burrough, P.A., McDonnell, R.A., 1998. Principles of Geographic Information Systems. Oxford University Press, New York.

Cao, C., Lam, N.S., 1997. Understanding the scale and resolution effects in remote sensing. In: Quatrocchi, D.A., Goodchild, M.F. (Eds.), Scale in Remote Sensing and GIS. Lewis Publishers, Boca Raton, FL, pp. 57–72.

Carroll, R.J., Chen, R., George, E.I., Li, T.H., Newton, H.J., Schmiediche, H., Wang, N., 1997. Ozone exposure and population density in Harris County, Texas. Journal of the American Statistical Association 92 (438), 392–404.

Casado, L.S., Rouhani, S., Cardelino, C.A., Ferrier, A.J., 1994. Geostatistical analysis and visualization of hourly ozone data. Atmospheric Environment 28 (12), 2105–2118.

Christakos, G., Vyas, V.M., 1998a. A composite space/time approach to studying ozone distribution over eastern United States. Atmospheric Environment 32 (16), 2845–2857.

Christakos, G., Vyas, V.M., 1998b. A novel method for studying population health impacts of spatiotemporal ozone distribution. Social Science and Medicine 47 (8), 1051–1166.

Coyle, M., Smith, R.I., Stedman, J.R., Weston, K.J., Fowler, D., 2002. Quantifying the spatial distribution of surface ozone concentration in the UK. Atmospheric Environment 36 (6), 1013–1024.

Cressie, N., 1991. Statistics for Spatial Data. Wiley, New York.

de Leeuw, F.A.A.M., Van Zantvoort, E.G., 1997. Mapping of exceedances of ozone critical levels for crops and forest trees in the Netherlands: preliminary results. Environmental Pollution 96 (1), 89–98.

Diem, J.E., Comrie, A.C., 2002a. Predictive mapping of air pollution involving sparse spatial observations. Environmental Pollution 119(1), 99-117.

Diem, J.E., Comrie, A.C., 2002b. Spatial allocation of anthropogenic pollution emissions: application to ozone pollution management. Journal of Environmental Management 63 (4), 425–447.

Duc, H., Shannon, I., Azzi, M., 2000. Spatial distribution characteristics of some air pollutants in Sydney. Mathematics and Computers in Simulation 54 (1–3), 1–21.

Duddek, C., Le, N.D., Zidek, J.V., Burnett, R.T., 1995. Multivariate imputation in cross-sectional analysis of health effects associated with air pollution. Environmental and Ecological Statistics 2, 191–212.

Fotheringham, A.S., Wong, D.W.S., 1991. The Modifiable Areal Unit Problem in multivariate statistical analysis. Environment and Planning A 23 (7), 1025–1044.

Fowler, D., Smith, R, Coyle, M., Weston, K.J., Davies, T.D., Ashmore, M.R., Brown, M., 1995. Quantifying the fine scale (1 km×1 km) exposure and effects of ozone: Part 1. Methodology and applications for effects on forests. Water, Air and Soil Pollution 85 (3), 1479–1484.

Fox, D.G., 1981. Judging air quality model performance. Bulletin of the American Meteorological Society 62 (5), 599–609.

Georgopoulos, P.G., Purushothaman, V., Chiou, R., 1997. Comparative evaluation of methods for estimating potential human exposure to ozone: photochemical modeling and ambient monitoring. Journal of Exposure Analysis and Environmental Epidemiology 7 (2), 191–215.

Godzik, B., 1997. Ground level ozone concentrations in the Kraków region, southern Poland. Environmental Pollution 98 (3), 273–280.

Griffith, D.A., Layne, L.J., 1999. A Casebook for Spatial Statistical Data Analysis. Oxford University Press, New York.

Guttorp, P., Meiring, W., Sampson, P.D., 1994. A space-time analysis of ground-level ozone data. Environmetrics 5 (3), 241–254.

Hogsett, W.E., Weber, J.E., Tingey, D., Herstrom, A., Lee, E.H., Laurence, J.A., 1997. Environmental auditing: an approach for characterizing tropospheric ozone risk to forests. Environmental Management 21 (1), 105–120.

Hopkins, L.P., Ensor, K.B., Rifai, H.S., 1999. Empirical evaluation of ambient ozone interpolation procedures to support exposure models. Journal of the Air and Waste Management Association 49 (7), 839–846.

Korc, M.E., 1996. A socioeconomic assessment of human exposure to ozone in the South Coast Air Basin of California. Journal of the Air and Waste Management Association 46 (6), 547–557.

Kuik, O.J., Helming, J.F.M., Dorland, C., Spaninks, F.A., 2000. The conomic benefits to agriculture of a reduction of low-level ozone pollution in The Netherlands. European Review of Agricultural Economics 27 (1), 75–90.

Lam, N.S., 1983. Spatial interpolation methods: a review. The American Cartographer 10 (2), 129–149.

Lam, N.S., Quattrochi, D.A., 1992. On the issue of scale, resolution, and fractal analysis in the mapping sciences. The Professional Geographer 44 (1), 88–97.

Laurence, J.A., Retzlaff, W.A., Kern, J.S., Lee, E.H., Hogsett, W.E., Weinstein, D.A., 2001. Predicting the regional impact of ozone and precipitation on the growth of loblolly pine and yellow-poplar using linked TREGRO and ZELIG models. Forest Ecology and Management 146,1–3 247-263.

Le, N.D., Sun, L., Zidek, J.V., 2001. Spatial prediction and temporal backcasting for environmental fields having monotone data patterns. The Canadian Journal of Statistics 29 (4), 529–554.

Lee, E.H., Hogsett, W.E., 2001. Interpolation of temperature and non-urban ozone exposure at high spatial resolution over the western United States. Climate Research 18 (3), 163–179.

Lefohn, A.S., Knudsen, H.P., Logan, J.A., Simpson, J., Bhumralkar, C., 1987. An evaluation of the kriging method to predict 7-h seasonal mean ozone concentrations for estimating crop losses. Journal of Air Pollution Control Association 37 (5), 595–602.

Lefohn, A.S., Knudsen, H.P., McEvoy, L.R., 1988. The use of kriging to estimate monthly ozone exposure parameters for the southeastern United States. Environmental Pollution 53 (1–4), 27–42.

Lefohn, A.S., Jackson, W., Shadwick, D.S., Knudsen, H.P., 1997. Effect of surface ozone exposures on vegetation grown in the southeastern Appalachian Mountains: identification of possible areas of concern. Atmospheric Environment 31 (11), 1695–1708.

Liu, L.J.S., Koutrakis, P., Leech, J., Broder, I., 1995. Assessment of ozone exposures in the greater metropolitan Toronto area. Journal of the Air and Waste Management Association 45 (4), 223–234.

Liu, L.J.S., Rossini, A.J., 1996. Use of kriging models to predict 12-hour mean ozone concentrations in metropolitan Toronto—a pilot study. Environment International 22 (6), 677–692.

Liu, L.J.S., Delfino, R., Koutrakis, P., 1997. Ozone exposure assessment in a southern California community. Environmental Health Perspectives 105 (1), 58–65.

Loibl, W., Winiwarter, W., Kopsca, A., Zueger, J., Bauman, R., 1994. Estimating the spatial distribution of ozone concentrations in complex terrain. Atmospheric Environment 28 (16), 2557–2566.

Loibl, W., Smidt, S., 1996. Areas of potential ozone risk for selected tree species. Environmental Science and Pollution Research 3 (4), 213–217.

McKendry, I.G., 1993. Ground-level ozone in Montreal, Canada. Atmospheric Environment 27B (1), 93–103.

McNair, L.A., Harley, R.A., Russell, A.G., 1996. Spatial inhomogeneity in pollutant concentrations, and their implications for air quality model evaluation. Atmospheric Environment 30 (24), 4291–4301.

Meiring, W., Guttorp, P., Sampson, P.D., 1998. Space-time estimation of grid-cell hourly ozone levels for assessment of a deterministic model. Environmental and Ecological Statistics 5 (3), 197–222.

Mulholland, J.A., Butler, A.J., Wilkinson, J.G., Russell, A.G., 1998. Temporal and spatial distributions of ozone in Atlanta: regulatory and epidemiologic implications. Journal of the Air and Waste Management Association 48 (5), 418–426.

Myers, D.E., 1991. Interpolation and estimation with spatially located data. Chemometrics and Intelligent Laboratory Systems 11, 209–228.

Nikiforov, S.Y., Aggarwal, M., Nadas, A., Kinney, P.L., 1998. Methods for spatial interpolation of long-term ozone concentrations. Journal of Exposure Analysis and Environmental Epidemiology 8 (4), 465–481.

Openshaw, S., Taylor, P.J., 1979. A million or so correlation coefficients: three experiments on the modifiable areal unit problem. In: Wrigley, N. (Ed.), Statistical Applications in the Spatial Sciences. Pion, London, pp. 127–144.

Pauly, M., Drüeke, M., 1996. Raumliche modellierung der ozonimissionen im mesoscale. Gefahrstiffe—Reinhaltung der Luft 56 (6), 225–230.

Phillips, D.L., Tingey, D.T., Lee, E.H., Herstrom, A.A., Hogsett, W.E., 1997. Use of auxiliary data for spatal interpolation of ozone exposure in southeastern forests. Environmetrics 8 (1), 43–61.

Pissimanis, D.K., Notaridou, V.A., Kaltsoundis, N.A., Viglas, P.S., 2000. On the spatial distribution of the daily maximum hourly ozone concentrations in the Athens basin in summer. Theoretical and Applied Climatology 65 (1–2), 49–62.

Rosenbaum, B.J., Strickland, T.C., McDowell, M.K., 1994. Mapping critical levels of ozone, sulfur dioxide for crops, forests, and natural vegetation in the United States. Water, Air, and Soil Pollution 74 (3–4), 307–319.

Royle, J.A., Berlinger, L.M., 1999. A hierarchical approach to multivariate spatial modeling and prediction. Journal of Agricultural, Biological, and Environmental Statistics 4 (1), 29–56.

Schichtel, B.A., Husar, R.B., 2001. Eastern North American transport climatology during high- and low-ozone days. Atmospheric Environment 35 (6), 1029–1038.

Sun, W., Le, N.D., Zidek, J.V., Burnett, R., 1998. Assessment of a Bayesian multivariate interpolation approach for health impact studies. Environmetrics 9 (5), 565–586.

Tobler, W.R., 1988. Resolution, resampling, and all that. In: Mounsey, H., Tomlinson, R. (Eds.), Building Databases for Global Science. Taylor and Francis, London, pp. 129–137.

Tolbert, P.E., Mulholland, J.A., MacIntosh, D.L., Xu, F., Daniels, D., Devin, O.J., Carlin, B.P., Klein, M., Dorley, J., Butler, A.J., Nordenberg, D.F., Frumkin, H., Ryan, P.B., White, M.C., 2000. Air quality and pediatric emergency room visits for asthma in Atlanta, Georgia. American Journal of Epidemiology 151 (8), 798–810.

US Environmental Protection Agency, 1999. Ozone Monitoring, Mapping, and Public Outreach: Delivering Real-Time Ozone Information to Your Community (EPA/625/R-99/007). US EPA National Risk Management Laboratory, Office of Research and Development, Cincinnati, Ohio.

Varns, J.L., Mulik, J.D., Sather, M.E., Glen, G., Smith, L., Stallings, C., 2001. Passive ozone network of Dallas: A modeling opportunity with community involvement. 1. Environmental Science and Technology 35 (5), 845–855.

Westenbarger, D.A., Frisvold, G.B., 1994. Agricultural exposure to ozone and acid precipitation. Atmospheric Environment 28 (18), 2895–2907.

Willmott, C.J., 1982. Some comments on the evaluation of model performance. Bulletin of the American Meteorological Society 63 (11), 1309–1313.

Wong, D., 1996. Aggregation effects in geo-referenced data. In: Arlinghaus, S.L. (Ed.), Practical Handbook of Spatial Statistics. CRC Press, Boca Raton, pp. 83–106.

Zidek, J.V., White, R., Le, N.D., Sun, W., Burnett, R.T., 1998. Imputing unmeasured explanatory variables in environmental epidemiology with application to health impact analysis of air pollution. Environmental and Ecological Statistics 5 (2), 99–115.