ELSEVIER

Statistical comparisons of methods for interpolating the output of a numerical air quality model

Xiaofeng Shao[a,*], Michael Stein[a], Jason Ching[b,1]

[a]*Department of Statistics, University of Chicago, 5734 S. University Avenue, Chicago, IL 60637, USA*
[b]*Atmospheric Sciences Modeling Division, Air Resources Laboratory, NOAA, RTP, NC 27711, USA*

Abstract

This paper compares Models-3/Community Multiscale Air Quality (CMAQ) outputs at multiple resolutions by interpolating from coarse resolution to fine resolution and analyzing the interpolation difference. Spatial variograms provide a convenient way to investigate the spatial character of interpolation differences and, importantly, to distinguish between naive (nearest neighbor) interpolation and bilinear interpolation, which takes a weighted average of four neighboring cells. For example, when the higher resolution is three times the lower, the variogram of the difference between naive interpolation of the lower resolution output and the higher resolution output shows a depression at every third lag. This phenomenon is related to the blocky nature of naive interpolation and demonstrates the inferiority of naive interpolation to bilinear interpolation in a way that pixelwise comparisons cannot. Theoretical investigations show when one can expect to observe this periodic depression in the variogram of interpolation differences. Naive interpolation is in fact used widely in a number of settings; our results suggest that it should be routinely replaced by bilinear interpolation.
© 2006 Elsevier B.V. All rights reserved.

1. Introduction

Numerical air quality models play a fundamental role in the scientific study and regulation of air pollution. These models are generally computationally intensive, especially if run at high spatial resolution. Thus, it is natural to compare model output at various resolutions and to describe how the results differ. This problem bears some resemblance to the use of observations at multiple spatial and/or temporal scales, for which tree-structured models have been proposed

---

* Corresponding author.

  E-mail addresses: shao@galton.uchicago.edu (X. Shao), stein@galton.uchicago.edu (M. Stein), ching.jason@epa.gov (J. Ching).

to capture the multiscale features; see Huang and Cressie (2001), Huang et al. (2002) and Johannesson and Cressie (2004) for some recent developments. However, deterministic model output at multiple scales has certain features (e.g., low noise, fine grids are a perfect partition of coarse grids, no missing values) that are unlikely to hold for observational data. This paper examines some simple statistical methods for comparing model output at multiple resolutions.

The model we study here is the Models-3/Community Multiscale Air Quality, or CMAQ (cf. Byun and Ching, 1999). CMAQ is an air quality simulation model that the US Environmental Protection Agency uses as a basis for implementing the national ambient air quality standards and is a tool for performing risk-based assessments and for developing environmental management strategies. It is also the model used for air quality forecasting by EPA and NOAA. One of the major motivations for running CMAQ at fine resolution is to facilitate the investigation of sub-grid variability (Ching et al., 2005). It is thus very important to assess the difference between the outputs at different resolutions. A natural idea for assessing what is lost by not having high-resolution output would be to interpolate from coarse resolution to fine resolution and look at the interpolation difference. Then the question arises: What interpolation scheme is best, or at least preferred, and is there a sensible statistical way by which we can tell the differences among various interpolation methods? Here, we mainly consider naive (nearest neighbor) interpolation and bilinear interpolation, which is a natural generalization of linear interpolation to two dimensions (see Section 2.2 for more details). In practice, both naive interpolation and bilinear interpolation are simple and inexpensive, and have been widely used in various fields such as digital image processing (Pratt, 1991), geostatistics (Wartenberg et al., 1991; Legates and Deliberty, 1993), remote sensing (Gupta, 2002) and ecological modeling (Matsushita et al., 2004). Some advantages and disadvantages have been discussed in Gupta (2002) and statistical comparisons with other interpolation methods have been addressed by Soufflet et al. (1991) and Satherley et al. (1996) in EEG spectral topography.

In this paper, we demonstrate that the spatial variogram and space–time variogram of interpolation differences provide meaningful and informative ways to investigate the difference of the model at two different resolutions and to distinguish between naive interpolation and bilinear interpolation. In particular, these variograms highlight anomalies of naive interpolation and show that it is distinctly inferior to bilinear interpolation at recovering higher resolution model output. When looking at the variogram of interpolation differences with a factor of three difference in resolution for CMAQ output, we find that a depression occurs at every third lag in the variogram of naive interpolation difference, but not for bilinear interpolation difference. This "three-lag" phenomenon is something we discovered early in our study of CMAQ output at multiple resolutions and a major motivation of this paper is to explain this phenomenon. By a variogram decomposition, we find that the main contribution to the three-lag phenomenon is from the variogram of the difference between bilinear interpolation and naive interpolation. For a one-dimensional intrinsic random process Z (see Cressie, 1993 for definition) with variogram γ_Z observed regularly with spacing 1, let Y be the difference of naive and linear interpolation of Z at spacing $\frac{1}{3}$. For any random process X, define each lag for X to be the distance between neighboring observations in X, for example, lag 1 for Z is distance 1, lag 1 for Y is distance $\frac{1}{3}$. Although Y is not an intrinsic random function, its empirical variogram is still a useful descriptor of its behavior. Our analytic calculations show that for this empirical variogram at lags $3q - 1$, $3q$ and $3q + 1$ (or distances $q - \frac{1}{3}$, q and $q + \frac{1}{3}$), the limiting expectations as the number of observations increases depends on γ_Z only at lags 1, $q - 1$, q and $q + 1$ and, is lower at lag $3q$ than at lags $3q - 1$ and $3q + 1$ if and only if $\gamma_Z(q+1) + \gamma_Z(q-1) > 2\gamma_Z(q)$. Thus, roughly speaking, we might expect a particularly strong depression/jump at lag $3q$ in the empirical variogram of Y if γ_Z is convex/concave at q. When $q = 1$, a third lag depression/jump is expected if Z is smoother/rougher than a Brownian motion, for which $\gamma(2) = 2\gamma(1)$. We obtain an analogous but much more complicated result for the two-dimensional case. Further, the variogram of bilinear interpolation difference allows us to find features that a direct visualization of the difference would not provide, such as moderate anisotropy, more local variation along E–W and NW–SE than NE–SW and N–S directions etc; see Section 2.3 for more details.

Naive interpolation is used explicitly or implicitly in a number of circumstances. For example, the way CMAQ is currently run is that when simulating at high resolution, the boundary and initial conditions are provided by naive interpolation of low resolution output over larger domains (cf. Byun and Ching, 1999, Chapter 13, 13.4.1). Furthermore, in the comparison of model output with monitoring data, it is common practice to compare the output of the cell containing the monitoring site to the observed value at that site, which is effectively using naive interpolation. We argue that bilinear interpolation is preferred on both practical and theoretical grounds. Although the improvement may be small in some of these situations, it is hard to see when bilinear interpolation would be inferior to naive interpolation.

The plan of the article is as follows. Section 2 describes the CMAQ output and investigates the interpolation difference for naive interpolation, bilinear interpolation and thin plate spline interpolation, showing the three-lag phenomenon for the empirical variogram of naive interpolation difference. Section 3 provides a theoretical explanation for the phenomenon. Section 4 summarizes the findings and provides some discussions and conclusions. Mathematical details are given in the appendices.

2. Comparison of CMAQ output at different resolutions

2.1. The data

A set of model simulations at multiple scales using CMAQ has been prepared and is used in this study. For our study, we sequentially modeled at 36, 12 and 4 km grid sizes using progressively smaller model domains to set up the fine scale modeling at 1.33 km grid size for the model domain and study area centered on the Philadelphia metropolitan area. The fine scale modeling required the introduction of additional vertical layers inside the urban canopy to incorporate an urban canopy parameterization for urban application (cf. Lacser and Otte, 2002). CMAQ model runs at 1.33 km resolution are unusual and provide a good opportunity for comparing output at different resolutions. Our case study focused on simulating the meteorology and chemistry for the 20-h period 1 am–8 pm (local time), on each of two days, July 12 and 14, 1995, days characterized by a stagnating high pressure synoptic regime, abundant sunshine, and abnormally high temperatures that dominated the eastern United States and southern Canada, and thus favorable to the progressive production of ozone and transport of precursor pollutants. Additional discussion of the meteorology and chemistry of this case can be found in Berman et al. (1999). The simulation details can be found in Otte et al. (2004).

In our study, we take a subset of the outputs from the top three resolutions with the domain matched to the 1.33 km domain, which correspond to 11×11, 33×33, 99×99 cells, respectively. In the following analysis, we mostly focus on the 4 and 1.33 km resolution output for the species NO_2 (ppm), NO (ppm) and CO (ppm). We denote by Z_{33} and Z_{99} the 4 and 1.33 km resolution output. Note that each value in the CMAQ output should be considered a grid square average. Define $A(Z_{99})$ to be the aggregation of Z_{99} to the 33×33 grid, i.e. each value of $A(Z_{99})$ is the average of Z_{99} over nine corresponding cells. Fig. 1 shows that Z_{33} is not just $A(Z_{99})$. For NO_2 at 1 am on July 14th, there is quite a lot of variation in Z_{33} given $A(Z_{99})$, especially when the magnitude of $A(Z_{99})$ gets large. There is also a tendency for Z_{33} to be larger than $A(Z_{99})$.
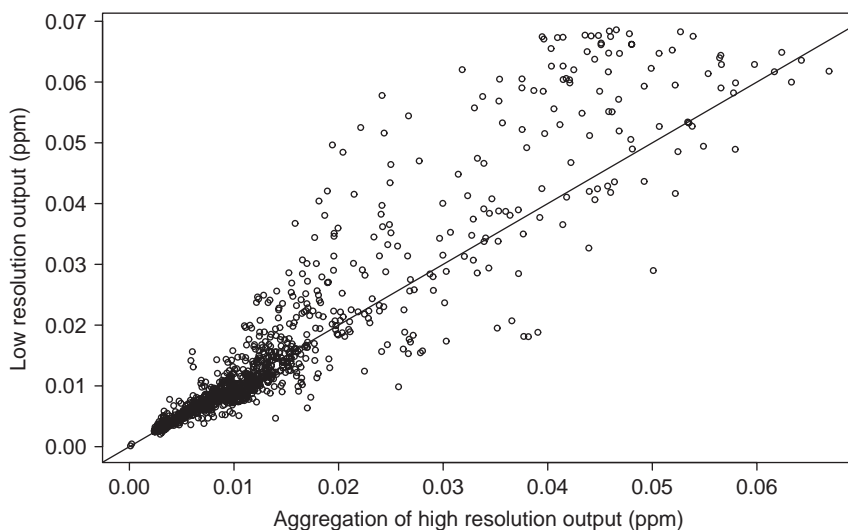

Fig. 1. The scatterplot of low resolution output (33×33) versus the aggregation of high resolution output (99×99) for NO_2 at 1 am (local time) on July 14th.
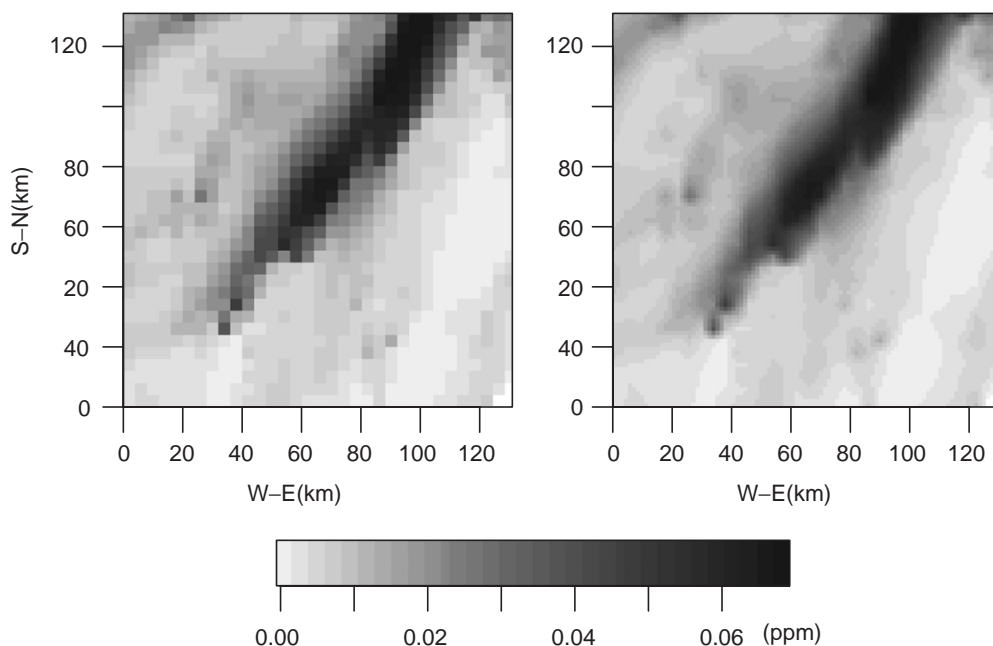
Fig. 2. The picture on the left depicts naive interpolated values of 4 km resolution NO_2 at 1 am (local time) on July 14th. The one on the right-hand side corresponds to bilinear interpolated values of 4 km resolution NO_2 at the same time.

2.2. Interpolations

High resolution runs of CMAQ can be highly demanding computationally, so a natural question to ask is how different the lower resolution output is from the higher resolution output. To address this question, we need a simple and effective way to compare the output at different resolutions. One way is to interpolate from 33×33 to 99×99 and look at the interpolation difference, i.e. $Z_{99} - \hat{Z}_{99}$, where $\hat{Z}_{99}$ is the interpolated value from Z_{33}. A naive way of doing interpolation is to assign each value of the 33×33 grid to the nine corresponding cells in the 99×99 grid. We call this interpolation scheme naive interpolation, but it is sometimes called the nearest neighbor method (see Gupta, 2002). However, naive interpolation produces sudden jumps between the cells of the coarser grid. By using a smoother interpolation scheme such as bilinear interpolation, we can remove these jumps. Bilinear interpolation is a two-dimensional analog of linear interpolation, for which a weighted average of the four nearest pixels is used to determine the interpolated value (see Gupta, 2002). To do bilinear interpolation from 33×33 to 99×99, one can, for example, do one-dimensional linear interpolation horizontally along each row of observations and then do linear interpolation along the vertical direction using the horizontally interpolated results. Due to the linearity of bilinear interpolation, one gets the same result by doing one-dimensional linear interpolation vertically along each column of observations and then doing linear interpolation along the horizontal direction using the vertically interpolated results. In the literature, the disadvantages of naive interpolation have been recognized. For example, it results in a disjointed or blocky image appearance, whereas bilinear interpolation is smoother and appears more natural (see Fig. 2). In the following section, we show that the stepped effect of naive interpolation produces anomalies in the spatial variograms and space–time variograms of the interpolation differences, which are removed by using bilinear interpolation.

We compare these two interpolations in several ways. As an overall measure, for each hour we compare the sum of squares (SS) of the differences after interpolation. In Fig. 3(a), we plot the relative improvement (in percentage) of bilinear interpolation over naive interpolation for NO_2, NO and CO on both days. In general, bilinear interpolation does a better job than naive interpolation for all species. For NO from 1 am to 4 am July 14th and 12 pm July 12th, a common feature of interpolation differences is that there are a small number of relatively large values that dominate the SS; see Fig. 3(b) for the quantile plot of absolute interpolation differences versus the quantiles of the absolute values of a standard normal variable for NO at 2 am on July 14th. Bilinear interpolation does a little bit worse than naive interpolation at these hotspots, which leads to the slight overall superiority of naive interpolation at these times for NO.
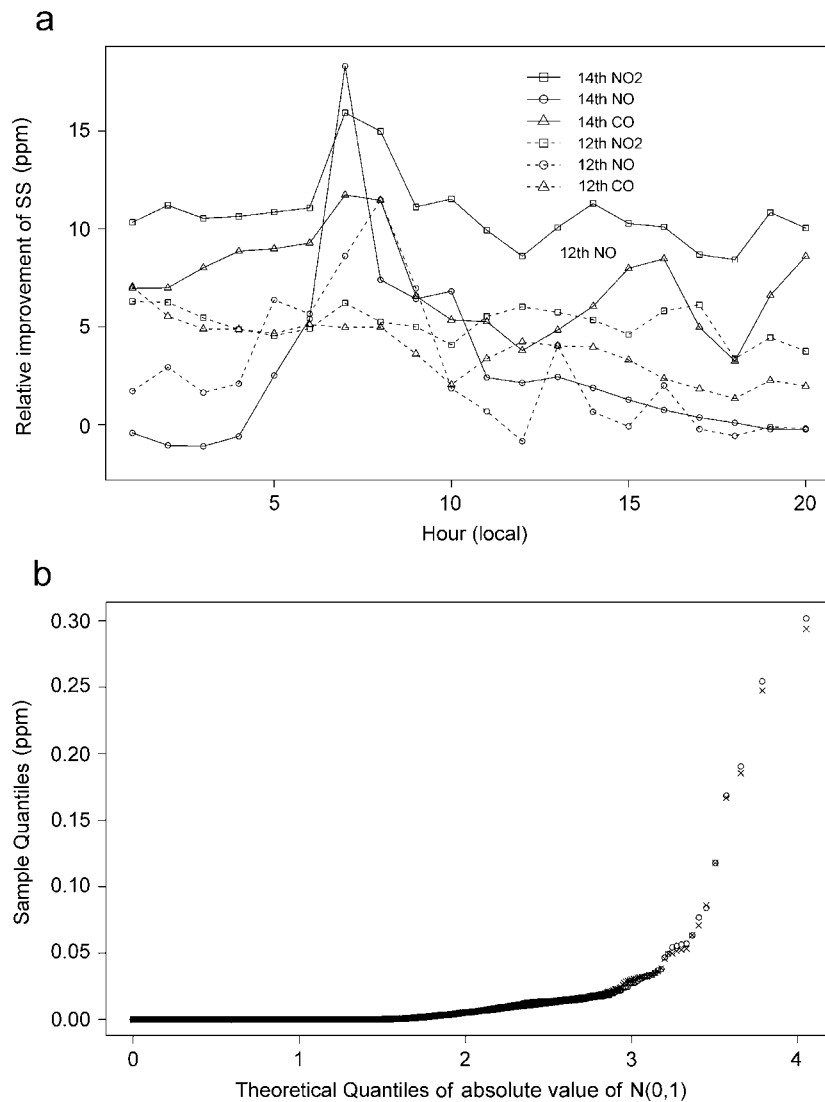
a

b

Fig. 3. (a) Relative improvement (in percentage) of bilinear interpolation over naive interpolation for NO_2, NO and CO from 1 am to 8 pm (local time) on July 14th and 12th in terms of sum of squares (SS) of interpolation differences, i.e. $Z_{99} - \hat{Z}_{99}$, where $\hat{Z}_{99}$ is the interpolated value from Z_{33}. (b) Quantile plot of absolute interpolation differences versus the quantiles of $|N(0, 1)|$ for NO at 2 am on July 14th (bilinear interpolation: $\circ$, naive interpolation: $\times$).

We also tried thin plate spline interpolation, which produces similar results as bilinear interpolation. Let $TPS(Z_{33})$, $B(Z_{33})$ and $N(Z_{33})$ denote the interpolated result by thin plate spline interpolation, bilinear interpolation, and naive interpolation, respectively. The first three rows of Table 1 compare the SS of interpolation differences $Z_{99} - TPS(Z_{33})$, $Z_{99} - B(Z_{33})$ and $Z_{99} - N(Z_{33})$ of NO_2 from 1 am to 6 am on July 14th. Bilinear interpolation is performs slightly better than thin plate spline interpolation, while naive interpolation is constantly the worst among the three. This slight superiority of bilinear interpolation over thin plate spline interpolation was observed for all 20 h on both days. Compared to naive and bilinear interpolation, thin plate spline interpolation is much more computationally expensive. For example, it takes 77.01 s of processing time on a PC with dual AMD Athlon MP 1800+ processor to interpolate from 33×33 to 99×99 grid in R using "tps" (69.92 s) and "predict" (7.09 s) commands, while it takes 0.00327 s for bilinear interpolation and 0.00205 s for naive interpolation.

There have been a number of studies comparing methods of interpolation (Tabios and Salas, 1985; Phillips et al., 1992; Dirks et al., 1998) when the observations equal the quantity of interest, or are the quantity of interest plus

Table 1
Sums of squares (SS) of interpolation differences $Z_{99} - \text{TPS}(Z_{33})$, $Z_{99} - B(Z_{33})$ and $Z_{99} - N(Z_{33})$ for NO_2 from 1 am to 6 am on July 14th

Hour (am)	1	2	3	4	5	6
$SS(Z_{99} - N(Z_{33}))$	0.501	0.443	0.481	0.510	0.482	0.356
$SS(Z_{99} - B(Z_{33}))$	0.450	0.393	0.430	0.456	0.430	0.316
$SS(Z_{99} - \text{TPS}(Z_{33}))$	0.463	0.401	0.435	0.463	0.438	0.317
$SS(\text{TPS}(Z_{33}) - B(Z_{33}))$	0.00328	0.00341	0.00363	0.00374	0.00351	0.00298
$SS(N(Z_{33}) - B(Z_{33}))$	0.0417	0.0442	0.0470	0.0477	0.0454	0.0390
$SS(N(Z_{33}) - \text{TPS}(Z_{33}))$	0.0413	0.0438	0.0466	0.0474	0.0450	0.0385
$\overline{SS}(Z_{99} - Z_{33})$	0.457	0.398	0.438	0.478	0.451	0.325

In the last row, $\overline{SS}(Z_{99} - Z_{33})$ corresponds to the sum of squares of differences at central pixels (within a 3×3 grid) multiplied by 9.

independent measurement errors. Our situation is rather different in that our "errors", the differences between the low and high resolution output, are only to a minor degree due to lack of resolution in the low-resolution results, but are mainly due to the fact that running the model at lower resolution changes the chemistry and dynamics of the model in a fundamental way. To illustrate the idea, we decompose the total error as follows:

$$Z_{99} - N(Z_{33}) = [Z_{99} - \text{Interpolation}(Z_{33})] + [\text{Interpolation}(Z_{33}) - N(Z_{33})]$$
$$= \text{intrinsic error} + \text{interpolation error}.$$

Rows 4–6 of Table 1 show that the intrinsic error is about 10 times as large as the interpolation error in terms of SS for both bilinear interpolation and thin plate spline interpolation.

Another way to measure the intrinsic error is to consider only the central pixels of each 3×3 grid in the high resolution, since all three interpolation methods give the corresponding value of Z_{33} at these pixels. Thus, looking at the sum of squared differences at just these central pixels provides an approximate lower bound on the performance of any interpolation scheme (that is, any scheme that does not seek to smooth the low resolution output). In particular, if these central pixels can be viewed as a representative sample of all pixels, 9 times the sum of squared differences for these pixels should provide a good measure of how well any interpolation scheme could hope to do. The last row of Table 1 gives these values, which are in fact slightly *higher* than the sum of squared differences at all pixels when using bilinear interpolation. This result suggests that the intrinsic error could not be reduced substantially by adopting a more complicated interpolation scheme.

Thin plate spline interpolation and bilinear interpolation are both special cases of kriging; see Cressie (1993, pp. 181–182) for a discussion about the connection between kriging and thin plate spline interpolation. For example, when the low resolution process is a Brownian sheet lattice (cf. Toth, 2004), the kriging procedure is equivalent to bilinear interpolation. The implied covariance functions behind bilinear and thin plate spline interpolation are quite different in terms of both smoothness of the corresponding process and isotropy and we are not claiming either is necessarily a good model for the low resolution output. The fact that two such different covariance models give such similar results here is further evidence that many interpolation schemes that smoothly fill in values between central pixels are likely to give similar results and that no interpolation methods could do much better than bilinear interpolation.

2.3. Variogram of interpolation difference

Comparison of numerical model output to data plays a central role in evaluation of such models. Standard methods focus on marginal characteristics of differences between model output and observations such as mean squared error and fractional bias (Irwin, 1997). Jun and Stein (2004) show that spatial and spatial-temporal variograms can provide useful insights into model errors that cannot be captured in marginal statistics. We contend that variograms are also a useful data analytic tool for comparing output of different models. For example, the empirical variograms of the output from different models for the same underlying process can be compared based on all of the usual characteristics one considers for spatial variograms such as sill, range, isotropy and local behavior, as well as characteristics of spatial-temporal variation such as asymmetry (Jun and Stein, 2004; Stein, 2005). In this work, the focus is on the variogram of the difference of two processes. The sill is then a measure of the overall differences between the models. However,

variograms with the same sill could have very different behavior at short lags and these differences can be informative. For example, a variogram of the difference between two models that has a large sill but very small values at short lags would indicate that the two model outputs have local fluctuations that are very similar, even if the actual levels of the output can be quite different.

For an intrinsic random function $Z(s), s \in S \subseteq R^2$, the theoretical variogram γ is defined by $\gamma(h) = \frac{1}{2}E\{Z(s) - Z(s+h)\}^2$. The (empirical) variogram is

$$V(h) = \frac{1}{2N(h)} \sum \{Z(s) - Z(s+h)\}^2,$$

where h is a two dimensional vector, the sum is over locations s for which $Z(s)$ and $Z(s+h)$ are both observed and $N(h)$ is the number of such pairs (Cressie, 1993). In our calculation of the variogram, we assign each value to the center of its corresponding cell, so the distance between any two cells is the distance between the centers of those two cells. Zhao and Wall (2004) investigate the impact on the estimation of the variogram when the distance between cells is redefined to be the distance between two points each of which is randomly chosen from within its corresponding cell, but identifying each cell with its center is more natural for output from a numerical model like CMAQ. Since we have 99×99 points in space, the variogram we obtained by averaging over space should be fairly stable. Fig. 4 shows some typical variogram plots. The variogram for bilinear interpolation is below the variogram for naive interpolation at all lags along each direction for all hours. Here each lag is the distance between neighboring cells at 1.33 km resolution. For example, along the N–S and E–W directions, each lag corresponds to 1.33 km physical distance, while along the NW–SE and NE–SW directions it is $1.33 \times \sqrt{2}$ km. Especially for small lags, the relative reduction of the variogram of bilinear interpolation over naive interpolation is substantial, which means that bilinear interpolation captures the local fluctuations in the higher resolution output much better than naive interpolation. For example, at the shortest lag in the NE–SW direction, the variogram of the interpolation difference is about 60% smaller for bilinear interpolation than for naive interpolation. Thus the variogram provides an informative alternative to the pixelwise comparison in Fig. 2 for distinguishing bilinear interpolation and naive interpolation. Another noticeable feature is that the variogram along the NE–SW direction is depressed every three lags for naive interpolation but not for bilinear interpolation. Along the E–W and N–S directions, this "three-lag" phenomenon is less obvious but still apparent. This anomaly, which is partly due to the discontinuous nature of naive interpolation, provides evidence that bilinear interpolation is distinctly superior to naive interpolation in this setting. From Fig. 4, it is seen that the variation is larger along the NW–SE than the NE–SW directions while the third lag depression is more obvious along the NE–SW than the NW–SE directions. We think that both of these phenomenon are related to the trend along the NE–SW direction as seen from Fig. 2. Part of the explanation will be provided in the next section; see Fig. 6.

Fig. 5(a) shows results for two interpolations from 12 km resolution to 1.33 km resolution. The format is the same as Fig. 4(a). We now see that the depression occurs at the 9th and 18th lags along the NE–SW direction. This nine-lag phenomenon, together with the three-lag phenomenon suggests that the lag at which the variogram is depressed should be a multiple of the factor of which the fine and coarse resolution are different in size. Figs. 5(c) and 5(d) show the variogram of interpolation differences from 12 km resolution to 4 km resolution. Again, an obvious third lag depression is spotted along NE–SW, N–S and E–W directions. This indicates that the three-lag phenomenon is not due to the urban canopy parameterization used in the 1.33 km resolution model run.

For each time point, we did two interpolations from Z_{33} to Z_{99} and obtained the space–time variograms of interpolation differences. For a space–time process $Z(s, t), s \in S$, stands for spatial location, $t = 1, \ldots, T$ stands for time. Given spatial lag h and time lag l,

$$V(h, l) = \frac{1}{2N(h, l)} \sum \{Z(s, t) - Z(s+h, t+l)\}^2,$$

where the sum is over all (s, t) for which $\{Z(s, t), Z(s+h, t+l)\}$ is observed and $N(h, l)$ is the number of such (s, t). Fig. 5(b) shows space–time variograms (along the NE–SW and NW–SE directions) of the interpolation differences of NO_2 with time lag 1. Again, an obvious "three-lag" phenomenon along the NE–SW direction is spotted for naive interpolation. This is also found in the space–time variograms for time lags 2 and 3 (results not shown). Finally, let us mention that the spatial variogram and space–time variogram for thin plate spline interpolation difference are very close to those for bilinear interpolation difference since these two interpolations produce very similar results. Moreover, we also observed the "three-lag" phenomenon for NO_2 on the log scale.
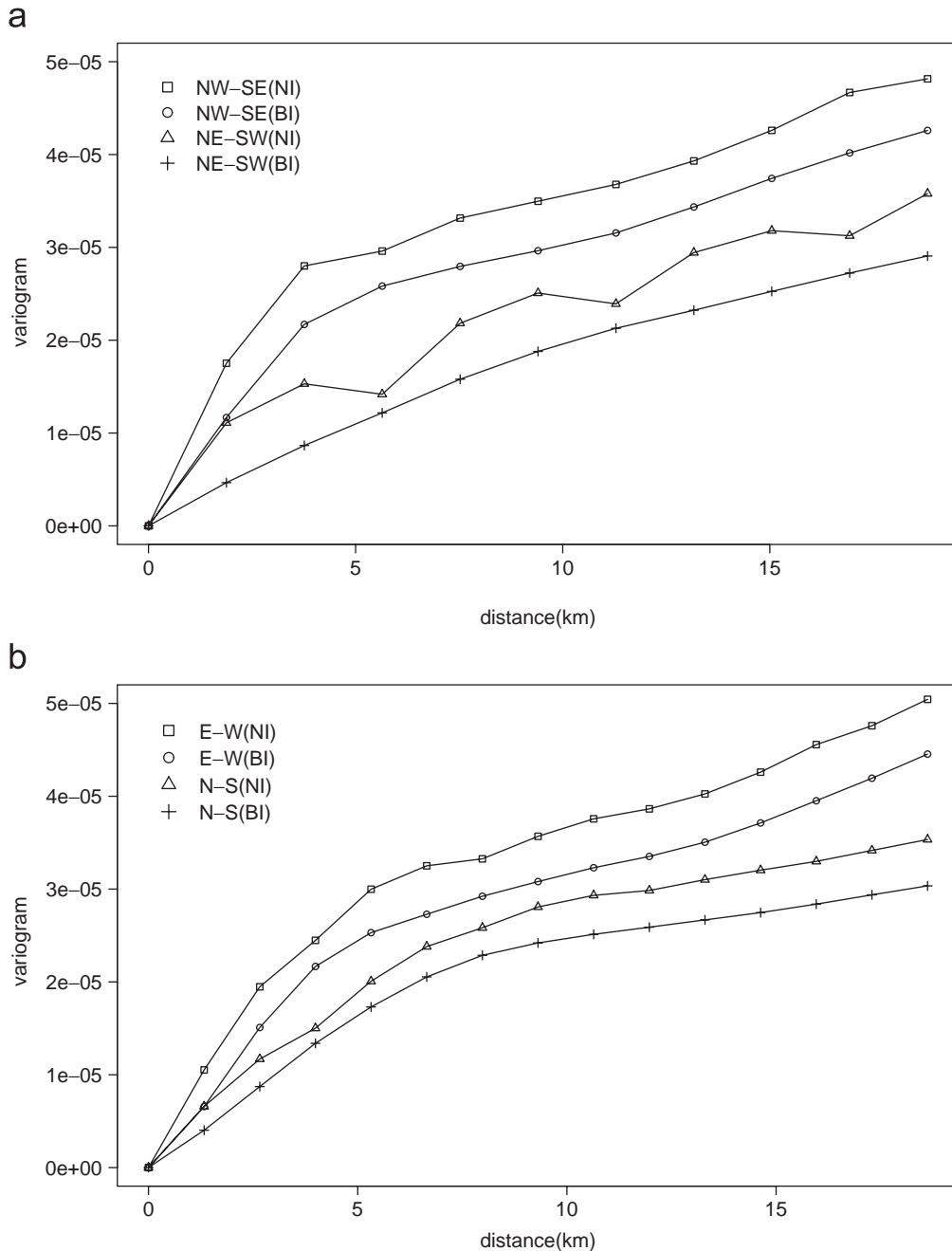
Fig. 4. Spatial variograms of naive interpolation and bilinear interpolation differences from 4 km resolution to 1.33 km resolution for NO_2 at 1 am (local time) on July 14th. (a) NE–SW and NW–SE directions. (b) N–S and E–W directions. (c) All four directions for bilinear interpolation difference.

Fig. 4(c) shows the variogram of bilinear interpolation difference along four directions for NO_2 at 1 am on July 14th. The difference exhibits moderate anisotropic features, with more variation along NW–SE and E–W directions, especially at short distances. Along NE–SW direction, the variability is the smallest at small lags and the variogram increases steadily and does not level off at even 70 km (the whole domain is about 131 km × 131 km), which is presumably due to the trend along that direction. The ranges along E–W, N–S and NW–SE directions are within 30–45 km. It is also noted that NW–SE and E–W directions have very similar variograms and the variograms are locally concave at small lags along four directions.
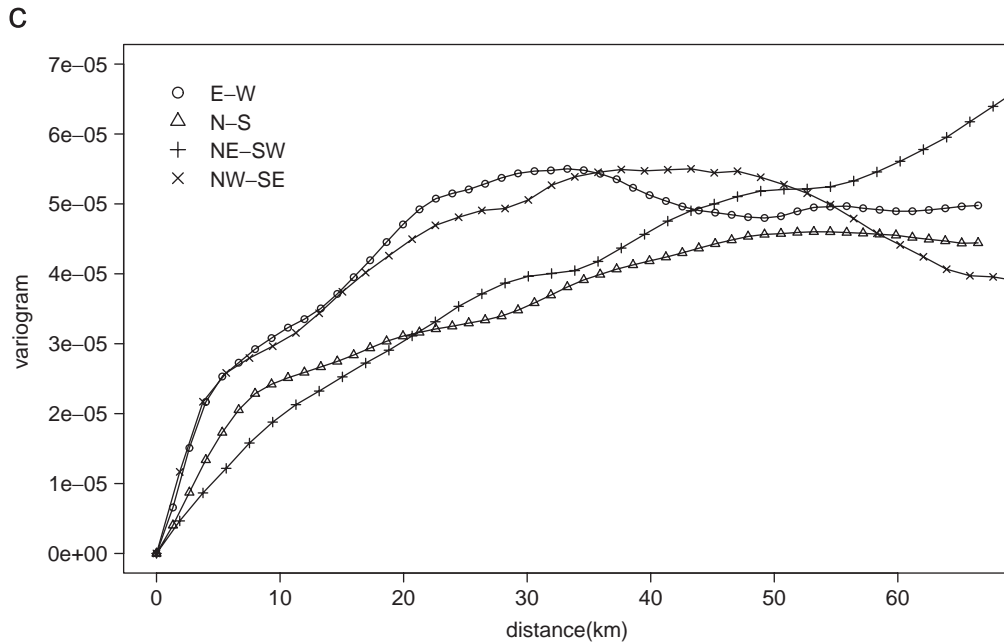
Fig. 4. (*continued*).

3. Three-lag phenomenon

In this section, we will explore further the three-lag phenomenon seen in the spatial variogram of naive interpolation differences. We would like to understand under what circumstances the three-lag phenomenon might be observed. The following variogram decomposition reveals the three-lag phenomenon in a clear way. Note that

$$Z_{99} - N(Z_{33}) = Z_{99} - B(Z_{33}) + B(Z_{33}) - N(Z_{33}),$$

where $Z_{99} - B(Z_{33})$ is bilinear interpolation difference and $B(Z_{33}) - N(Z_{33})$ is the difference between bilinear and naive interpolation. Empirically, we observe that the variogram for $Z_{99} - N(Z_{33})$ is approximately the sum of the variograms for $Z_{99} - B(Z_{33})$ and $B(Z_{33}) - N(Z_{33})$, i.e.

$$V(k_1, k_2; Z_{99} - N(Z_{33})) \approx V(k_1, k_2; Z_{99} - B(Z_{33})) + V(k_1, k_2; B(Z_{33}) - N(Z_{33})), \tag{1}$$

where $V(k_1, k_2; W)$ denotes the empirical variogram for the process W at lag (k_1, k_2). The approximation (1) would hold exactly in expectation if $Z_{99} - B(Z_{33})$ and $B(Z_{33}) - N(Z_{33})$ were uncorrelated.

For NO$_2$ at 1 am July 14th and various values of (k_1, k_2), Table 2 shows the percentage of the relative approximation error

$$R_1 = 100 \times \frac{V(k_1, k_2; Z_{99} - N(Z_{33})) - [V(k_1, k_2; Z_{99} - B(Z_{33})) + V(k_1, k_2; B(Z_{33}) - N(Z_{33}))]}{V(k_1, k_2; Z_{99} - N(Z_{33}))} \tag{2}$$

and the relative contribution from the variogram of the difference between bilinear interpolation and naive interpolation

$$R_2 = 100 \times \frac{V(k_1, k_2; B(Z_{33}) - N(Z_{33}))}{V(k_1, k_2; Z_{99} - N(Z_{33}))}. \tag{3}$$

The approximation (1) holds reasonably well with the relative error ranging from 0.1% to 6.8%. For R_2, the largest value occurs at the shortest lags and the relatively small magnitude for the third lag is partly due to the three-lag depression in the variogram of $B(Z_{33}) - N(Z_{33})$. Qualitatively similar results hold for other hours on July 14th and 12th.

Fig. 6 shows the variogram of $B(Z_{33}) - N(Z_{33})$ along four directions. Interestingly, the variogram at the third lag falls substantially below the variogram at the second and fourth lags. Along the N–S and E–W directions, the three-lag phenomenon is much more apparent than that observed in the variogram plot of $Z_{99} - N(Z_{33})$. The strongest depression occurs at the third lag along the NE–SW direction, which matches what we observed in the variogram of $Z_{99} - N(Z_{33})$. Visually, the three-lag phenomenon is amplified in the variogram of $B(Z_{33}) - N(Z_{33})$ by taking out a smooth part from the variogram for the naive interpolation difference.
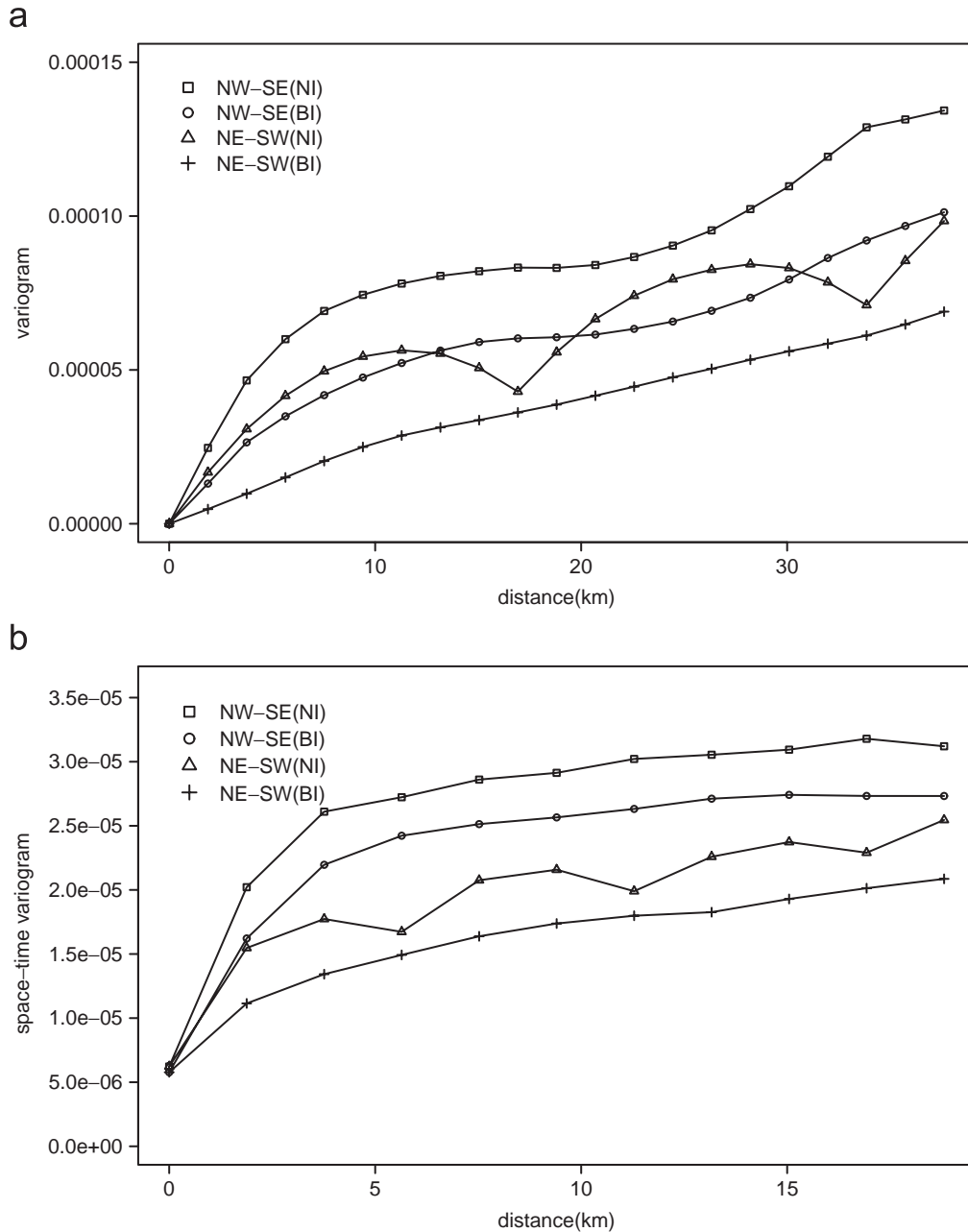

Fig. 5. Spatial and space–time variograms for naive interpolation and bilinear interpolation differences for NO_2 at 1 am (local time) on July 14th. (a) Spatial variogram from 12 km resolution to 1.33 km resolution along the NE–SW and NW–SE directions. (b) Space–time variogram with time lag 1 from 4 km resolution to 1.33 km resolution along the NE–SW and NW–SE directions. (c) Spatial variogram from 12 km resolution to 4 km resolution along the NE–SW and NW–SE directions. (d) Spatial variogram from 12 km resolution to 4 km resolution along the N–S and E–W directions.
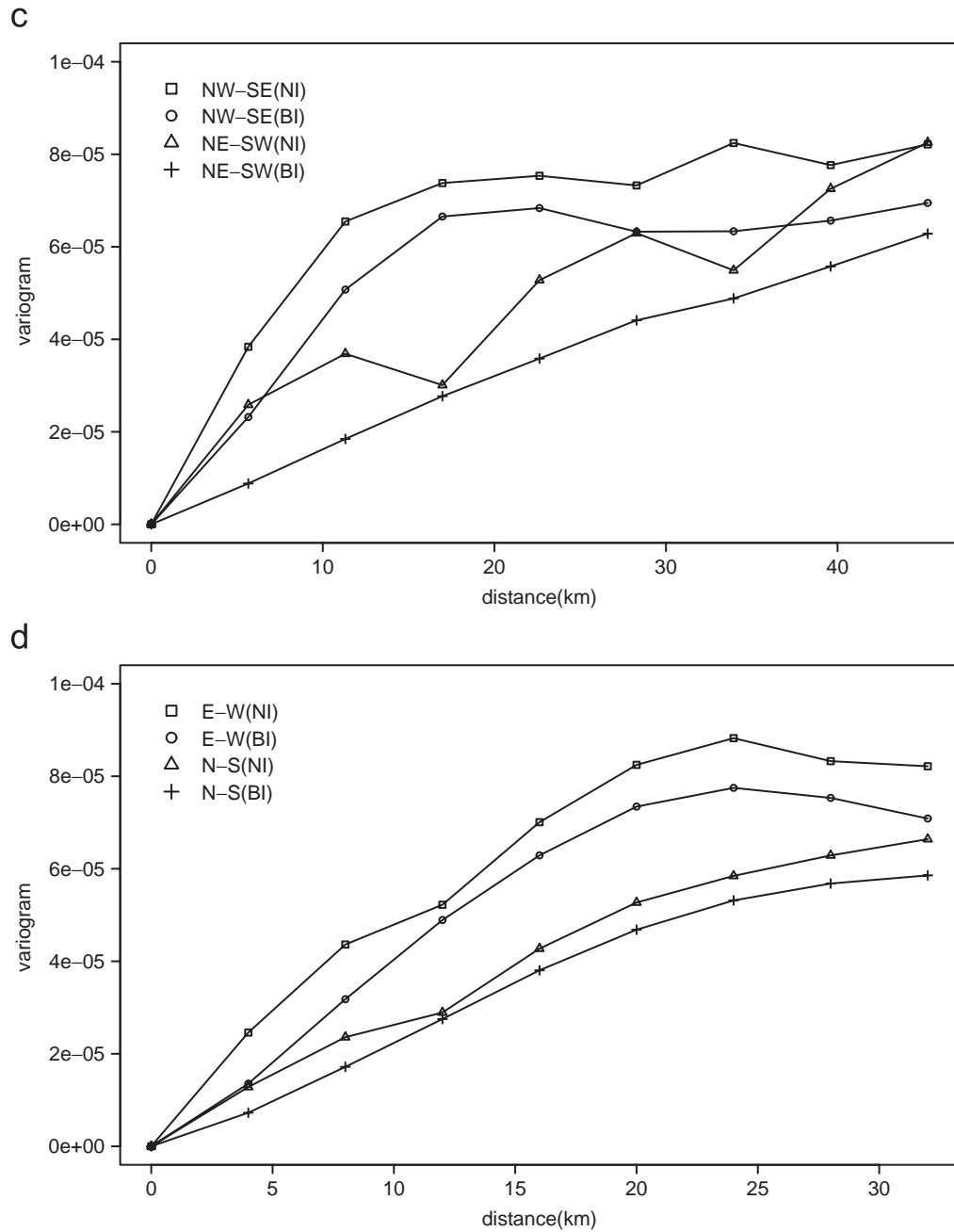
c

d

Fig. 5. (*continued*).

Ideally, we would like to know for what kind of processes Z_{33} and Z_{99} we would see the three-lag phenomenon. To answer this question requires a complete specification of the process at two different scales. Modeling the joint relationship of Z_{33} and Z_{99} is very difficult. However, since it appears that the main contribution to three-lag phenomenon seen in the variogram of $Z_{99} - N(Z_{33})$ is from the variogram of $B(Z_{33}) - N(Z_{33})$, we can ask the simpler question: for what kind of Z_{33} can we see the three-lag phenomenon in the variogram plot of $B(Z_{33}) - N(Z_{33})$?

Let us suppose that Z_L is an intrinsic random field with theoretical variogram γ_L. In our example, $Z_L = Z_{33}$. Analytically, we can express the expectation of the empirical variogram of $B(Z_L) - N(Z_L)$ in terms of γ_L. A more

Table 2
Relative approximation error R_1 (Eq. (2)) and contribution R_2 (Eq. (3)) from the variogram of $B(Z_{33}) - N(Z_{33})$ (in percentage) for the variograms along four directions

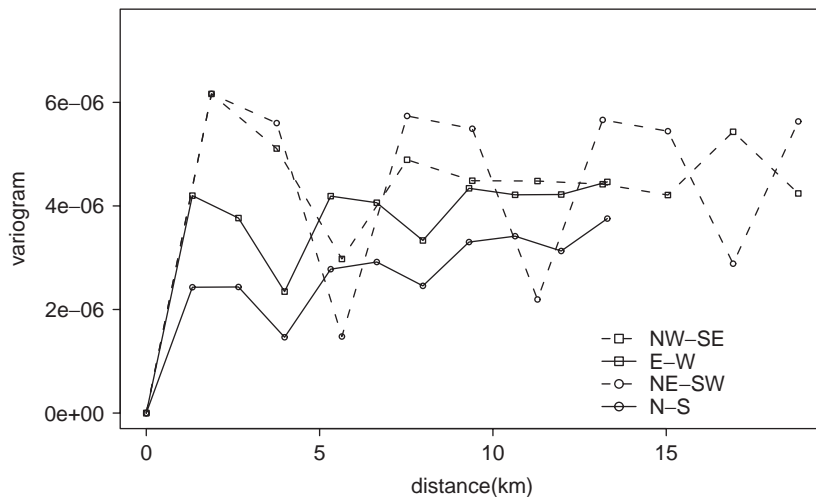| k	N–S R_1	$(k, 0)$ R_2	E–W R_1	$(0, k)$ R_2	NE–SW R_1	(k, k) R_2	NW–SE R_1	$(k, -k)$ R_2
1	2.1	36.9	−2.4	39.9	2.5	55.5	−1.8	35.2
2	4.5	20.8	3.2	19.3	6.8	36.6	4.3	18.2
3	1.1	9.8	1.9	9.6	3.7	10.4	2.7	10.1
4	−0.1	13.8	1.6	14.0	1.3	26.3	1.0	14.8

Fig. 6. Spatial variogram of the difference between bilinear interpolation and naive interpolation for NO_2 4 km resolution at 1 am (local time) on July 14th.

detailed analysis is done for one-dimensional processes in the appendices. The message is that the three-lag phenomenon is related to the sign and magnitude of $2\gamma_L(q) - \gamma_L(q-1) - \gamma_L(q+1)$, $q \geqslant 1$ (see Appendix A). For the empirical variogram V_{3L} of $B(Z_L) - N(Z_L)$ (see Appendix A for definition), the $3q$th lag depression appears when $\min\{V_{3L}(3q-1), V_{3L}(3q+1)\} > V_{3L}(3q)$. Our calculations show that $E\{V_{3L}(3q-1)\} \approx E\{V_{3L}(3q+1)\}$ when the sample size is large and, furthermore, $E\{V_{3L}(3q-1) - V_{3L}(3q)\}$ is roughly a positive constant times $\gamma_L(q-1) - 2\gamma_L(q) + \gamma_L(q+1)$. Thus, the magnitude of the third lag depression depends on $\gamma_L(2) - 2\gamma_L(1)$, which is 0 for Brownian motion and is, roughly speaking, positive for smoother processes. On the other hand, if the low resolution process is less smooth than Brownian motion, i.e. $\gamma_L(2) < 2\gamma_L(1)$, we expect to see a third lag jump in V_{3L} for large L. In general, the $3q$th lag depression/jump occurs in the mean of V_{3L} for large L when γ_L is convex/concave at q; see Fig. 7 for an illustration.

In the two-dimensional case, the limiting values for the expectation of V_{3L} at lags 2–4 along the E–W direction are given in Appendix B. Interestingly,

$$E\{V_{3L}(0, 2)\} = E\{V_{3L}(0, 3)\} = E\{V_{3L}(0, 4)\} \tag{4}$$

in the limit when the low resolution process has theoretical variogram $\gamma(k_1, k_2) \propto |k_1| + |k_2|$, the variogram for a process that is a sum of uncorrelated Brownian motions in each coordinate. To see why, let $Z_L(i, j) = W_1(i, j) + W_2(i, j)$, where $W_1(i, j) = W_1(i)$ and $W_2(i, j) = W_2(j)$ are two uncorrelated Brownian motion. Then by linearity of bilinear
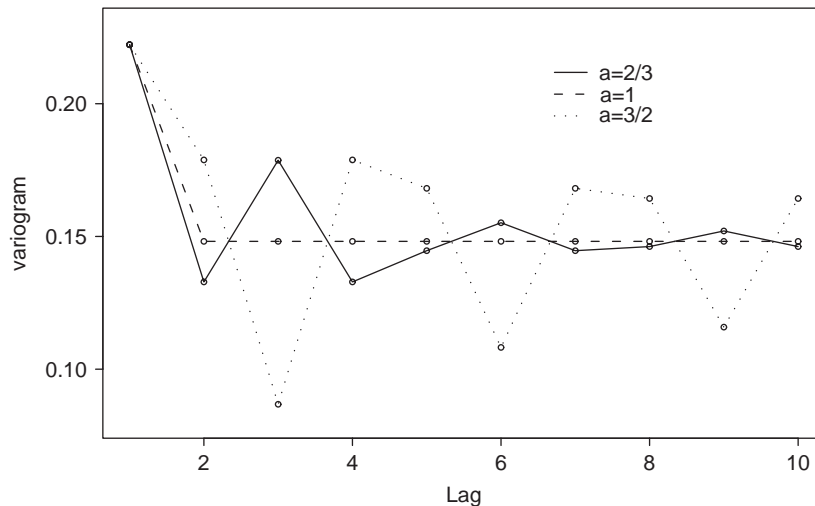
Fig. 7. Limiting value of expected variogram of the difference between linear interpolation and naive interpolation for the processes with various smoothness, $\gamma(h) = h^a$, $a = \frac{2}{3}, 1, \frac{3}{2}$.

and naive interpolation, we have the following decomposition:

$$B(Z_L(i, j)) - N(Z_L(i, j)) = B(W_1(i, j)) - N(W_1(i, j)) + B(W_2(i, j)) - N(W_2(i, j)).$$

Thus for any positive integer k,

$$E\{V_{3L}(0, k)\} = E[V\{0, k; B(W_1(i, j)) - N(W_1(i, j))\}] + E[V\{0, k; B(W_2(i, j)) - N(W_2(i, j))\}]$$

due to the fact that $W_1(\cdot)$ and $W_2(\cdot)$ are uncorrelated. So relation (4) follows from our one-dimensional result. To verify the consistency of our theoretical results with the empirical values, we plug in the variogram values for the low resolution process (i.e. NO_2 4 km resolution at 1 am July 14th) and calculate the limiting value for $E\{V_{3L}(0, 2)\}$, $E\{V_{3L}(0, 3)\}$ and $E\{V_{3L}(0, 4)\}$, which are 3.86×10^{-6}, 2.24×10^{-6}, 4.19×10^{-6}, matching our empirical values (3.77×10^{-6}, 2.35×10^{-6}, 4.19×10^{-6}) very well (see Fig. 6), both sets of values showing a strong depression at lag 3.

4. Discussion and conclusions

In this paper, the advantage of bilinear interpolation over naive interpolation for comparing model output at different resolutions is demonstrated by looking at spatial variograms and space–time variograms of interpolation differences. The three-lag phenomenon is observed in the variogram of naive interpolation difference and is well explained by the variogram decomposition (1). Our analytic calculations show that if the low resolution process is a one-dimensional intrinsic random function with variogram γ_L, when the number of observations of the low resolution process is large enough, the $3q$th lag depression can be expected to show up in the variogram plot of the difference between bilinear interpolation and naive interpolation, provided that $\gamma_L(q + 1) + \gamma_L(q - 1) > 2\gamma_L(q)$. The magnitude of this depression is proportional to $\gamma_L(q + 1) - 2\gamma_L(q) + \gamma_L(q - 1)$. Roughly speaking, we expect to see a third lag depression when the process is smoother than Brownian motion, for which $\gamma_L(2) = 2\gamma_L(1)$. We observe the third lag depression along all directions (see Figs. 4(a) and (b)) since CMAQ model output tends to be smoother than Brownian motion. In another related work where the resolution differs by a factor of four (cf. Shao and Stein, 2006), we observe a four-lag phenomenon in the variogram of naive interpolation differences, which further suggests that this depression in the variogram is an artifact of using naive interpolation. In summary, there is no point in using naive interpolation since we can remove this depression phenomenon by bilinear interpolation with only trivial additional computation.

The work reported on here is a part of a larger effort to study sub-grid variability in numerical models, a topic of major interest in the modeling community (Ching et al., 2005). Efforts to study sub-grid variability in the modeling community generally ignore the spatial and temporal dependencies in the processes under study. Our plan for studying sub-grid variability in a way that accounts for this dependence is to obtain a statistical description for the conditional distribution of high resolution output given lower resolution output. A reasonable approach to doing this is to model the differences between high and low resolution output. The results here show that modeling the bilinear interpolation difference is much easier than modeling naive interpolation difference, since it would be awkward to model a process with a third-lag depressed spatial variogram. Furthermore, more sophisticated interpolation schemes, such as thin plate spline interpolation, may not do any better than bilinear interpolation and, thus, may not be worth the extra computational effort. The features we observed from the variogram of bilinear interpolation difference will provide some guidance in doing statistical modeling of this difference. One conclusion we draw from the present work that has immediate value for numerical modelers is that when running CMAQ or similar models at high resolution, the boundary and initial conditions should be obtained using bilinear interpolation of low resolution output rather than the presently used naive interpolation.

Acknowledgments

Appendix A. Analytical calculations for three-lag phenomenon (1D)

Suppose we have a one-dimensional intrinsic random process $Z(t)$, t real, and the low resolution process is modeled as an equidistant sample from Z with resolution 1 and size L, i.e. $Z(i)$, $i = 1, \ldots, L$ with theoretical variogram $\gamma(\cdot)$, where $\gamma(s) = \frac{1}{2}E\{Z(x) - Z(x+s)\}^2$. We do the interpolation to a finer resolution $1/p$, where p is a positive integer. Let $D_{pL} := B(Z) - N(Z)$ be the difference between linear interpolation and naive interpolation with the empirical variogram denoted by V_{pL}, where at lag s, $V_{pL}(s) := (1/2(pL - s))\sum_{j=1}^{pL-s}[D_{pL}(j/p) - D_{pL}\{(j+s)/p\}]^2$. In the following two cases, we investigate the pth lag depression for p odd and even, respectively.

Case 1. p is odd, $p \geqslant 3$. We have

$$D_{pL}(i) = \begin{cases} \dfrac{p+1-2i}{2p}\{Z(1) - Z(2)\}, & i \leqslant \dfrac{p-1}{2}, \\[2ex] \left(i + \dfrac{p-1}{2} - pL\right)\{Z(L) - Z(L-1)\}, & i \geqslant pL - (p-3)/2, \\[2ex] \dfrac{p+1-2p(i/p - \lfloor i/p \rfloor)}{2p}\left\{Z\left(\left\lfloor \dfrac{i + (p-1)/2}{p} \right\rfloor\right)\right. \\[2ex] \quad \left. - Z\left(\left\lfloor \dfrac{i + (p-1)/2}{p} \right\rfloor + 1\right)\right\}, & \dfrac{p-1}{2p} < \dfrac{i}{p} < L - \dfrac{p-3}{2p}, \ \dfrac{i}{p} \notin \mathscr{Z}, \\[2ex] \dfrac{1-p}{2p}\left\{Z\left(\left\lfloor \dfrac{i + (p-1)/2}{p} \right\rfloor\right)\right. \\[2ex] \quad \left. - Z\left(\left\lfloor \dfrac{i + (p-1)/2}{p} \right\rfloor + 1\right)\right\}, & i/p \in \mathscr{Z}, \ 1 \leqslant i/p \leqslant L - 1, \end{cases}$$

where $\lfloor a \rfloor$ is the integer part of a and $\mathscr{Z}$ is the set of all the integers. To investigate the pth lag depression, we first find the expression for $E\{V_{pL}(p-1)\}$, $E\{V_{pL}(p)\}$ as well as $E\{V_{pL}(p+1)\}$ and consider the limiting value when $L \to \infty$. For the convenience of calculation, we ignore the terms from the two boundaries, i.e. only $D_{pL}(k)$,

$k = (p+1)/2, \ldots, (pL - (p-1)/2)$ are used. The contribution of these border terms is easily shown to be asymptotically negligible. Let

$$S_L^{(1)} := \sum_{i=1}^{L-2} \{Z(i) + Z(i+2) - 2Z(i+1)\}^2, \quad S_L^{(2)} := \sum_{i=1}^{L-2} \{Z(i) - Z(i+2)\}^2$$

and $\quad S_L^{(3)} := \sum_{j=0}^{(p-3)/2} \sum_{i=1}^{L-2} \{jZ(i) + (j+1)Z(i+2) - (2j+1)Z(i+1)\}^2$

$$+ \sum_{j=0}^{(p-3)/2} \sum_{i=1}^{L-2} \{(j+1)Z(i) + jZ(i+2) - (2j+1)Z(i+1)\}^2.$$

Then we have

$$V_{pL}(p-1) = \frac{1}{2(pL - p + 1)p^2} [(p-1)^2 S_L^{(2)}/4 + S_L^{(3)} + \{Z(1) - Z(2)\}^2 + \{Z(L-1) - Z(L)\}^2],$$

$$V_{pL}(p) = \frac{p^2 - 1}{24(pL - p)p} S_L^{(1)} \quad \text{and} \quad V_{pL}(p+1) = \frac{1}{2(pL - p + 1)p^2} \{(p-1)^2 S_L^{(2)}/4 + S_L^{(3)}\}.$$

Let $G_{p,\gamma} := \sum_{j=0}^{(p-3)/2} \{(2j+1)^2 \gamma(1) - (j^2 + j)\gamma(2)\}$. Simple calculations show that $ES_L^{(1)} = 2(L-2)(4\gamma(1) - \gamma(2))$, $ES_L^{(2)} = 2(L-2)\gamma(2)$ and $ES_L^{(3)} = 4(L-2)G_{p,\gamma}$. For $V_{pL}(k), k = p - 1, p, p + 1$, we take expectations and let $L \to \infty$,

$$E\{V_{pL}(p-1)\} = \frac{(p-1)^2(L-2)\gamma(2)/4 + 2\gamma(1) + 2(L-2)G_{p,\gamma}}{(pL - p + 1)p^2}$$

$$\longrightarrow \frac{p^2 - 3p + 2}{3p^2}\gamma(1) - \frac{p^2 - 6p + 5}{12p^2}\gamma(2),$$

$$E\{V_{pL}(p)\} = \frac{(p^2 - 1)(L-2)\{4\gamma(1) - \gamma(2)\}}{12(pL - p)p} \longrightarrow \frac{p^2 - 1}{3p^2}\gamma(1) - \frac{p^2 - 1}{12p^2}\gamma(2)$$

and $\quad E\{V_{pL}(p+1)\} = \frac{(p-1)^2(L-2)\gamma(2)/4 + 2(L-2)G_{p,\gamma}}{(pL - p + 1)p^2}$

$$\longrightarrow \frac{p^2 - 3p + 2}{3p^2}\gamma(1) - \frac{p^2 - 6p + 5}{12p^2}\gamma(2).$$

Interestingly, $E\{V_{pL}(p-1)\} > E\{V_{pL}(p)\}$ in the limit if and only if $2\gamma(1) < \gamma(2)$, independent of p. Also note that the limiting values for $E\{V_{pL}(p-1)\}$ and $E\{V_{pL}(p+1)\}$ are the same.

Case 2. p is even, $p \geqslant 4$. The calculations are similar to the odd case and are omitted. We only show the limiting values as $L \to \infty$.

$$E\{V_{pL}(p-1)\} \to \frac{p^3 - 3p^2 + 2p + 3}{3p^3}\gamma(1) - \frac{p^3 - 6p^2 + 5p + 3}{12p^3}\gamma(2),$$

$$E\{V_{pL}(p)\} \to \frac{p^2 - 1}{12p^2}\{4\gamma(1) - \gamma(2)\},$$

$$E\{V_{pL}(p+1)\} \to \frac{4p^3 - 12p^2 + 8p + 9}{12p^3}\gamma(1) - \frac{p^3 - 6p^2 + 5p + 9}{12p^3}\gamma(2) + \frac{1}{4p^3}\gamma(3).$$

The pth lag depression shows up in the limit if

$$E\{V_{pL}(p-1)\} > E\{V_{pL}(p)\} \equiv 4(p^2 - p - 1)\gamma(1) < (2p^2 - 2p - 1)\gamma(2)$$

and

$$E\{V_{pL}(p+1)\} > E\{V_{pL}(p)\} \equiv (12p^2 - 12p - 9)\gamma(1) - (6p^2 - 6p - 9)\gamma(2) - 3\gamma(3) < 0.$$

In the case $p = 4$, the fourth lag depression appears in the limit if $1.913\gamma(1) < \gamma(2)$ and $2.143\gamma(1) - 0.048\gamma(3) < \gamma(2)$.

Finally, when $p = 3$, we investigate the variogram of $B(Z) - N(Z)$ at lags $3q - 1$, $3q$ and $3q + 1$ where $q \geqslant 2$. Similar to the above procedure, we first find the general expression of $V_{3L}(3q - 1)$, $V_{3L}(3q)$ as well as $V_{3L}(3q + 1)$, then take the expectation and find its limiting value as $L \to \infty$. We have

$$E\{V_{3L}(3q - 1)\} \to \{4\gamma(1) + \gamma(q - 1) - 2\gamma(q) + \gamma(q + 1)\}/27,$$

$$E\{V_{3L}(3q)\} \to 2\{2\gamma(1) + 2\gamma(q) - \gamma(q + 1) - \gamma(q - 1)\}/27,$$

$$E\{V_{3L}(3q + 1)\} \to \{4\gamma(1) + \gamma(q - 1) - 2\gamma(q) + \gamma(q + 1)\}/27.$$

Again in the limit $E\{V_{3L}(3q - 1)\} = E\{V_{3L}(3q + 1)\}$. The $3q$th lag depression occurs in the limit when $\gamma(q - 1) + \gamma(q + 1) > 2\gamma(q)$.

Appendix B. Analytical calculations for three-lag phenomenon (2D)

Now we consider a two-dimensional low resolution process Z, which is an equidistant sample from a two-dimensional intrinsic random process $Z(s)$, $s \in \mathbb{R}^2$, i.e. $Z(i, j)$, $i, j = 1, \ldots, L$. Denote the theoretical variogram of Z by $\gamma(\cdot)$, where $\gamma(s_1, s_2) = \frac{1}{2}E\{Z(i, j) - Z(i + s_1, j + s_2)\}^2$ at lag (s_1, s_2). Let D_{3L} be the difference between bilinear interpolation and naive interpolation of Z observed at (i, j), $i, j = 1, \ldots, L$, i.e. $D_{3L} = B(Z) - N(Z)$ with its variogram denoted by V_{3L}, where at lag (s_1, s_2),

$$V_{3L}(s_1, s_2) := \frac{1}{2(3L - s_1)(3L - s_2)} \sum_{j=1}^{3L-s_1} \sum_{k=1}^{3L-s_2} [D_{3L}(j/3, k/3) - D_{3L}\{(j + s_1)/3, (k + s_2)/3\}]^2.$$

Again, we express the expectation of V_{3L} in terms of γ at various lags and find its limiting value as $L \to \infty$. Here we only show the results for the horizontal direction (E–W):

$$EV_{3L}(0, 2) \to \tfrac{1}{729}[134\gamma(0, 1) + 19\gamma(0, 2) + 72\gamma(1, 0) - 40\gamma(1, 1) - 40\gamma(-1, 1) + 4\gamma(1, 2) + 4\gamma(-1, 2)],$$

$$EV_{3L}(0, 3) \to \tfrac{1}{729}[308\gamma(0, 1) - 50\gamma(0, 2) + 96\gamma(1, 0) - 46\gamma(1, 1) - 46\gamma(-1, 1) - 2\gamma(1, 2) - 2\gamma(-1, 2)],$$

$$EV_{3L}(0, 4) \to \tfrac{1}{729}[134\gamma(0, 1) + 55\gamma(0, 2) + 108\gamma(1, 0) - 40\gamma(1, 1) - 40\gamma(-1, 1) - 14\gamma(1, 2) - 14\gamma(-1, 2)].$$

In the calculation, we only need to get the exact formula for 3×3 grid since $E[D_{3L}(j/3, k/3) - D_{3L}\{(j + s_1)/3, (k + s_2)/3\}]^2$ is the same as $E[D_{3L}(j/3 + h_1, k/3 + h_2) - D_{3L}\{(j + s_1)/3 + h_1, (k + s_2)/3 + h_2\}]^2$ for any integers h_1, h_2, j, k, s_1, s_2. Further we note that the three limiting values coincide when the low resolution process has theoretical variogram $\gamma(k_1, k_2) \propto |k_1| + |k_2|$.

References

Berman, S., Ku, J.Y., Rao, S.T., 1999. Spatial and temporal variation in the mixing depth over the northeastern United States during the summer of 1995. J. Appl. Meteorol. 38, 1661–1673.

Byun, D.W., Ching, J.K.S., 1999. Science algorithms of the EPA Models-3 community multiscale air quality (CMAQ) modeling system. Available at: ⟨http://www.epa.gov/asmdnerl/models3/doc/science/science.html⟩.

Ching, J.K.S., Herwehe, J., Swall, J., 2005. On joint deterministic grid modelling and sub-grid variability conceptual framework for model evaluation. Atmospheric Environment 40, 4935–4945.

Cressie, N., 1993. Statistics for Spatial Data. Wiley, New York.

Dirks, K.N., Hay, J.E., Stow, C.D., Harris, D., 1998. High-resolution studies of rainfall on Norfolk Island part II: interpolation of rainfall data. J. Hydrol. 208, 187–193.

Gupta, R.P., 2002. Remote Sensing Geology. second ed. Springer, New York.

Huang, H.C., Cressie, N., 2001. Multiscale graphical modeling in space: applications to command and control. In: Moore, M. (Ed.), Spatial Statistics: Methodological Aspects and Applications. Springer Lecture Notes in Statistics, vol. 159. Springer, NY, pp. 83–113.

Huang, H.C., Cressie, N., Gabrosek, J., 2002. Fast resolution-consistent spatial prediction of global processes from satellite data. J. Comput. Graph. Statist. 11, 63–88.

Irwin, J.S., 1997. Draft standard practice for statistical evaluation of atmospheric dispersion models. American Society for Testing and Materials, Z6849Z.

Johannesson, G., Cressie, N., 2004. Variance–covariance modeling and estimation for multi-resolution spatial models. In: Sanchez-Vila, X., Carrera, J., Gomez-Hernandez, J. (Eds.), geoENV IV-Geostatistics for Environmental Applications. Kluwer, Dordrecht, pp. 319–330.

Jun, M., Stein, M.L., 2004. Statistical comparison of observed and CMAQ modeled daily sulfate levels. Atmos. Environ. 38, 4427–4436.

Lacser, A., Otte, T.L., 2002. Implementation of an urban canopy parameterization in MM5. Preprints, Fourth Symposium on Urban Environment, Norfolk, VA, American Meteorological Society, 153–154.

Legates, D.R., Deliberty, T.L., 1993. Precipitation measurement biases in the United States. Water Resource Bull. 29, 855–861.

Matsushita, B., Xu, M., Chen, J., Kameyama, S., Tamura, M., 2004. Estimation of regional net primary productivity (NPP) using a process-based ecosystem model: how important is the accuracy of climate data? Ecol. Model. 178, 371–388.

Otte, T.L., Lacser, A., Dupont, S., Ching, J.K.S., 2004. Implementation of an urban canopy parameterization in a mesoscale meteorological model. J. Appl. Meteorol. 43, 1648–1665.

Phillips, D.L., Dolph, J., Marks, D., 1992. A comparison of geostatistical procedures for spatial analysis of precipitations in mountainous terrain. Agr. Forest. Meteorol. 58, 119–141.

Pratt, W.K., 1991. Digital Image Processing. second ed. Wiley-Interscience, New York.

Satherley, B.L., Jones, R.D., Bones, P.J., 1996. EEG spectral topography in neurology: II a new system and a theoretical comparison of interpolation techniques. Australasian Phys. Eng. Sci. Medicine 19, 183–193.

Shao, X., Stein, M.L., 2006. Statistical conditional simulation of a multiresolution numerical air quality model. J. Geophys. Res.111, D15211, doi: 10.1029/2005JD007037.

Soufflet, L., Toussaint, M., Luthringer, R., Gresser, J., Minot, R., Macher, J.P., 1991. A statistical evaluation of the main interpolation methods applied to 3-dimensional EEG mapping. Electroen. Clin. Neuro. 79, 393–402.

Stein, M.L., 2005. Space–time covariance functions. J. Amer. Statist. Assoc. 100, 310–321.

Tabios, G.Q., Salas, J.D., 1985. A comparative analysis of techniques for spatial interpolation of precipitation. Water Resource Bull. 21, 365–380.

Toth, D., 2004. Adding interior points to an existing Brownian sheet lattice. Statist. Probab. Lett. 66, 221–227.

Wartenberg, D., Uchrin, C., Coogan, P., 1991. Estimating exposure using kriging—a simulation study. Environ. Health Persp. 94, 75–82.

Zhao, Y., Wall, M.M., 2004. Investigating the use of the variogram for lattice data. J. Comput. Graph. Statist. 13, 719–738.