Estimating Population Exposure to Fine Particulate Matter in the Conterminous U.S. using Shape Function-based Spatiotemporal Interpolation Method

A County Level Analysis

Lixin Li

Department of Computer Sciences
Georgia Southern University
Statesboro, GA, U.S.A.
lli@georgiasouthern.edu

Xingyou Zhang, James B. Holt
Centers for Disease Control and Prevention
Atlanta, GA, U.S.A.
gyx8@cdc.gov, jgh4@cdc.gov

Jie Tian

Department of Geology and Geography
Georgia Southern University
Statesboro, GA, U.S.A.
jtian@georgiasouthern.edu

Reinhard Piltner

Department of Mathematical Sciences
Georgia Southern University
Statesboro, GA, U.S.A.
rpiltner@georgiasouthern.edu

Abstract — **This paper investigates spatiotemporal interpolation methods for the application of air pollution assessment. The air pollutant of interest in this paper is fine particulate matter $PM_{2.5}$. The choice of the time scale is investigated when applying the shape function-based method. It is found that the measurement scale of the time dimension has an impact on the quality of interpolation results. Based upon the result of 10-fold cross validation, the most effective time scale out of four experimental ones was selected for the $PM_{2.5}$ interpolation. The paper also estimates the population exposure to the ambient air pollution of $PM_{2.5}$ at the county-level in the contiguous U.S. in 2009. The interpolated county-level $PM_{2.5}$ has been linked to 2009 population data and the population with a risky $PM_{2.5}$ exposure has been estimated. The risky $PM_{2.5}$ exposure means the $PM_{2.5}$ concentration exceeding the National Ambient Air Quality Standards. The geographic distribution of the counties with a risky $PM_{2.5}$ exposure is visualized. This work is essential to understanding the associations between ambient air pollution exposure and population health outcomes.**

Keywords- spatiotemporal interpolation, fine particulate matter, air pollution exposure, time scale

I. INTRODUCTION

Spatial interpolation methods have been well developed to estimate values at unknown locations based upon values that are spatially sampled in GIS (Geographic Information Systems). These methods assume a stronger correlation among points that are closer than those farther apart. They are characterized as either deterministic or stochastic depending on whether statistical properties are utilized. Deterministic interpolation methods determine an unknown value using mathematical functions with predefined parameters such as distances in Inverse Distance Weighting (IDW) [1], areas and volumes in shape function (SF) based methods [2-5], while stochastic interpolation methods such as Kriging [6, 7], investigate the spatial autocorrelation.

Although spatial interpolation methods have been adopted in various applications, many critical problems remain unsolved. One of them is that traditional GIS researchers tend to treat space and time separately when interpolation needs to be conducted in the continuous space-time domain. The primarily strategy identified from the literature is to reduce spatiotemporal interpolation problems to a sequence of snapshots of spatial interpolations [8]. In order to interpolate for an unmeasured time instance, temporal interpolation can then be conducted based on the spatial interpolation results at each location [9].

A. Motivation

Integrating space and time simultaneously is anticipated to yield better interpolation results than treating them separately for certain typical GIS applications. For example, the following set of ozone data collected 8 annual ozone concentration measurements in 1994, 4 in 1995, 3 in 1996, 6 in 1997, and 8 in 1998, as shown by the solid circular dots in Fig. 1. Suppose we need to estimate the ozone concentration at the location illustrated by the empty dot W in 1996. The traditional strategy can only utilize the existing measurements at the locations A, B and C in 1996 to interpolate the ozone concentration at the location W, although A, B and C are rather far away from W. It would be inappropriate to use the ozone concentrations around the San Francisco area to interpolate the concentrations for New York, even if they were from the same time instance.

Therefore, a more integrative research framework that allows us to incorporate measurements from other years in the interpolation for the space-time point W needs to be developed. Depending on the characteristics of various spatiotemporal interpolation methods, different groups of measurements may be chosen to interpolate for W. One possibility would be choosing F and G in 1997 and D and E in 1995 since these four points are relatively closer to W in space and time. In general,

spatiotemporal interpolation is based on the assumption: things that are closer in the space-time domain are more alike than those that are farther apart.
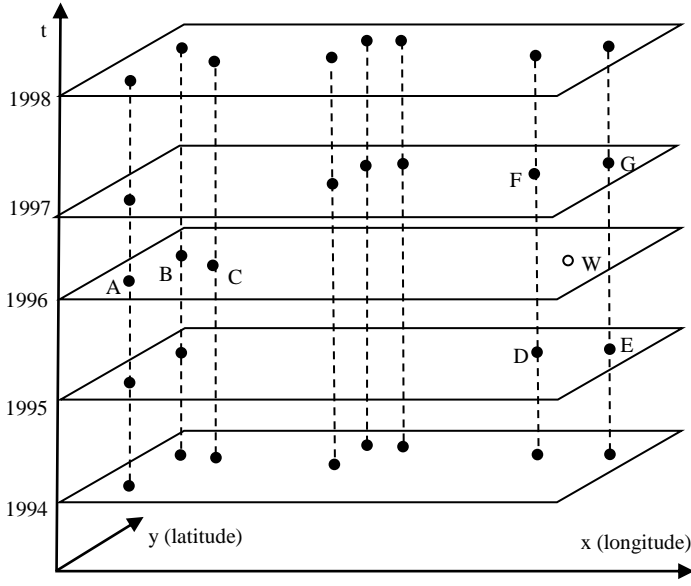

Figure 1. Annual ozone concentration sample points from 1994 to 1998. The solid circular dots (•) are the points with known measurements, while the empty dot (o) is unmeasured and needs to be interpolated for 1996.

B. Existing Approach and Its Limitation

In order to integrate space and time simultaneously, a so called "extension approach" has been proposed in [4, 10]. This approach treats time as another dimension in space and therefore extending the spatiotemporal interpolation problem into a higher-dimensional spatial interpolation problem. Some applications using the extension approach can be found in [3, 11-13].

Although the idea of the extension approach is intriguing, extending spatial interpolation to incorporate a dimension of time poses many challenges. One of the biggest challenges is that space and time have different and incomparable scales. The physically sound distance in a space-time domain is therefore difficult to define or compute. For example, how would you define the spatiotemporal distance between two points P_1 (x_1, y_1, t_1) and P_2 (x_2, y_2, t_2) when the extension approach of IDW is applied? This challenge has been rarely investigated and impedes to a very large degree the development of a more generalized and logically sound approach of spatiotemporal interpolation. The existing practice used in references [3, 11, 12] is to choose some convenient scales coming from the raw GIS data.

C. Contributions

There are two main contributions of our paper. The first contribution is to examine whether the choice of time scale affects the quality of SF (Shape Function) based spatiotemporal interpolation results. SF-based 2-D triangular interpolation methods are proven to be invariant to coordinate scales [3]. However, in reference [3], the authors did not realize that the invariance is only true after the mesh is constructed. Mesh generation is the foundation of SF-based interpolation methods.

It divides the total domain into a finite number of simple sub-domains that can be triangles in the case of 2-D problems and tetrahedra in the case of 3-D problems. However, will a different choice of time scale (e.g. month or day) influence the construction of a mesh and thereafter the performance of a spatiotemporal interpolation? According to our knowledge, this is a topic that has been rarely investigated.

The second contribution is to set an important initial step to investigate the associations between ambient air pollution exposure and population health outcomes. Since $PM_{2.5}$ concentrations are only measured at certain monitoring sites and time instances, the $PM_{2.5}$ concentrations at unsampled locations and times need to be estimated using an effective method. In this paper, an efficient 3-D SF-based spatiotemporal interpolation method is applied to the $PM_{2.5}$ data set. The interpolated county-level $PM_{2.5}$ has been linked with 2009 population data and the population with a risky $PM_{2.5}$ exposure has been estimated. The geographic distribution of the counties exceeding the $PM_{2.5}$ air quality standards is displayed.

II. SF-BASED SPATIOTEMPORAL INTERPOLATION AND ITS INVARIANCE AND VARIANCE TO COORDINATE SCALES

Shape functions, which can be viewed as a spatial interpolation method, are popular in engineering applications such as finite element algorithms [14, 15]. There are various types of 2-D and 3-D shape functions. We are interested in 2-D shape functions for triangles and 3-D shape functions for tetrahedra, both of which are linear approximation methods. Shape functions have been adopted and further developed as spatiotemporal interpolation methods for GIS applications [3, 16, 17]. Reference [3] also describes IDW and Kriging based spatiotemporal interpolation methods and compares them to SF-based methods by using an actual real estate data set with house prices. SF-based spatiotemporal interpolation methods have been applied to various applications including air quality mapping [11, 12] and many others [18-20].

A. SF-based Spatiotemporal Interpolation: A Review

At the beginning stage of applying SF-based interpolation methods, a mesh that divides the total domain into a finite number of simple sub-domains or elements should be generated. For example,

- for a 2-D spatial problem, a mesh composed of triangular elements should be generated if one wants to use shape functions for triangles to interpolate unknown values in the (x, y) coordinate system;

- for a 3-D spatial problem, a mesh composed of tetrahedral elements should be generated if one wants to use shape functions for tetrahedra to interpolate unknown values in the (x, y, z) coordinate system.

Quite successful algorithms have been developed to generate triangular or tetrahedral meshes, including the popular method of Delaunay triangulation meshing [21, 22]. Delaunay triangulation is related to the construction of the so called "Voronoi diagram", which is related to "Convex Hull" problems. The techniques to compute 2-D Delaunay triangulations can be extended to create 3-D tetrahedral meshes. Fig. 2 shows a tetrahedron.
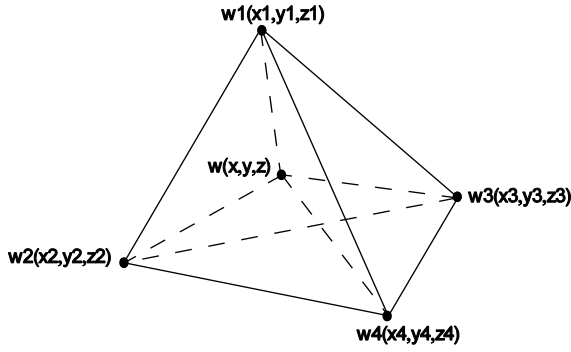
Figure 2. Computing 3-D shape functions by tetrahedral volume divisions. w_1, w_2, w_3 and w_4 are measured values, while the value w at location (x, y, z) is unmeasured and needs to be interpolated.

Considering the tetrahedron in Fig. 2, the SF-based interpolation result at a point (x, y, z) located inside the tetrahedron can be obtained by using the measurement values w_1, w_2, w_3, and w_4 at the four corners as below [3]:

$$w(x, y, z) = N_1(x, y, z)w_1 + N_2(x, y, z)w_2 + \\ N_3(x, y, z)w_3 + N_4(x, y, z)w_4 \tag{1}$$

where N_1, N_2, N_3 and N_4 are the following shape functions

$$N_1(x, y, z) = \frac{V_1}{V}, \; N_2(x, y, z) = \frac{V_2}{V}, \\ N_3(x, y, z) = \frac{V_3}{V}, \; N_4(x, y, z) = \frac{V_4}{V} \tag{2}$$

V_1, V_2, V_3, and V_4 are the volumes of the four sub-tetrahedra $ww_2w_3w_4$, $w_1ww_3w_4$, $w_1w_2ww_4$, and $w_1w_2w_3w$, respectively; and V is the volume of the outside tetrahedron $w_1w_2w_3w_4$ as shown in Fig. 2.

Our paper focuses on spatiotemporal interpolation problems in the domain of 2-D space and 1-D time. Using the extension approach of SF-based interpolation methods, we treat time as the imaginary third dimension z in space. Therefore, substituting the z variable by the time variable t in (1) and (2) gives a SF-based spatiotemporal interpolation method for 2-D space and 1-D time problems:

$$w(x, y, t) = N_1(x, y, t)w_1 + N_2(x, y, t)w_2 + \\ N_3(x, y, t)w_3 + N_4(x, y, t)w_4 \tag{3}$$

where N_1, N_2, N_3 and N_4 are the following shape functions

$$N_1(x, y, t) = \frac{V_1}{V}, \; N_2(x, y, t) = \frac{V_2}{V}, \; N_3(x, y, t) = \frac{V_3}{V}, \; N_4(x, y, t) = \frac{V_4}{V} \tag{4}$$

Extending spatial interpolation to incorporate a dimension of time using the SF-based interpolation method has shown promising results compared with the extension methods based on IDW and Kriging [3, 12]. The challenge of mixing incomparable scales of space and time has been investigated in [3]. It has been proven that the SF-based 2-D triangular interpolation method is invariant to coordinate scales. However, the invariance only holds after the mesh is generated. From the application point of view, one often needs to decide coordinate scales before meshing, especially the time

scale for spatiotemporal interpolation. There is a lack of literature investigating the metrical integration of space and time prior to the mesh construction. For the rest of this section, we first show the proof of the invariance to coordinate scales of 3-D tetrahedral shape functions after meshing. Then we investigate its variance to coordinate scales before meshing.

B. Invariance to Coordinate Scaling After Meshing

Lemma 1. *3-D tetrahedral shape functions are invariant to coordinate scaling after the mesh construction.*

Proof. Similar as the proof for 2-D triangular shape functions in [3], we obtain a proof for the invariance of 3-D tetrahedral shape functions to coordinate scaling after meshing as below.

Since it is assumed that coordinate scaling happens after the mesh is constructed, each tetrahedron in the mesh has the same set of corner vertices before and after coordinate scaling. For a given tetrahedron, comparison can be made for the value of $N_1(x,y,t)$ in (3) and (4) before and after coordinate scaling. We first calculate the results of V_1, V_2, V_3, and V_4 in (4) using the coordinates of the corner vertices. Then we have

$$N_1(x, y, t) =$$
$$[x_2y_3t_4 - x_2y_4t_3 - x_3y_2t_4 + x_3y_4t_3 + x_4y_2t_3 - x_4y_3t_3 + (-y_3t_4 + y_4t_3 \\ + y_2t_4 - y_4t_2 - y_2t_3 + y_3t_2)x + (-x_2t_4 + x_2t_3 + x_3t_4 - x_3t_2 - x_4t_3 \\ + x_4t_2)y + (-x_2y_3 + x_2y_4 + x_3y_2 - x_3y_4 - x_4y_2 + x_4y_3)t] \; \div \\ [(x_2y_3t_4 - x_2t_3y_4 - x_3y_2t_4 + x_3t_2y_4 + x_4y_2t_3 - x_4t_2y_3 - x_1y_3t_4 + x_1t_3y_4 \\ + x_3y_1t_4 - x_3t_1y_4 - x_4y_1t_3 + x_4t_1y_3 + x_1y_2t_4 - x_1t_2y_4 - x_2y_1t_4 + x_2t_1y_4 \\ + x_4y_1t_2 - x_4t_1y_2 - x_1y_2t_3 + x_1t_2y_3 + x_2y_1t_3 - x_2t_1y_3 - x_3y_1t_2 + x_3t_1y_2)].$$

Assume that the t dimension enlarges to n times of the original scale. Then N_1 will be calculated as b after the scaling

$$N_1'(x, y, t) =$$
$$[nx_2y_3t_4 - nx_2y_4t_3 - nx_3y_2t_4 + nx_3y_4t_3 + nx_4y_2t_3 - nx_4y_3t_3 + (-ny_3t_4 \\ + ny_4t_3 + ny_2t_4 - ny_4t_2 - ny_2t_3 + ny_3t_2)x + (-nx_2t_4 + nx_2t_3 + nx_3t_4 - nx_3t_2 \\ - nx_4t_3 + nx_4t_2)y + (-x_2y_3 + x_2y_4 + x_3y_2 - x_3y_4 - x_4y_2 + x_4y_3)nt] \; \div \\ [(nx_2y_3t_4 - nx_2t_3y_4 - nx_3y_2t_4 + nx_3t_2y_4 + nx_4y_2t_3 - nx_4t_2y_3 - nx_1y_3t_4 \\ + nx_1t_3y_4 + nx_3y_1t_4 - nx_3t_1y_4 - nx_4y_1t_3 + nx_4t_1y_3 + nx_1y_2t_4 - nx_1t_2y_4 \\ - nx_2y_1t_4 + nx_2t_1y_4 + nx_4y_1t_2 - nx_4t_1y_2 - nx_1y_2t_3 + nx_1t_2y_3 + nx_2y_1t_3 \\ - nx_2t_1y_3 - nx_3y_1t_2 + nx_3t_1y_2)].$$

It is obvious that $N_1'(x,y,t)$ is the same as $N_1(x,y,t)$. Therefore N_1 is invariant to t coordinate scaling after meshing. Invariance to x or y coordinate scaling is straightforward too. Similarly, we can prove that N_2, N_3 and N_4 in (3) and (4) are also invariant to coordinate scaling after meshing. □

C. Variance to Coordinate Scaling Before Meshing

When the SF-based spatiotemporal interpolation using the extension approach is applied, an important question has been neglected in the literature: will different time scales lead to different meshes using the Delaunay triangulation method?

In order to answer this question, a simple 2-D example can demonstrate whether coordinate scales affect the mesh result:

- We first used the 76 house locations of the real estate data in [3] and generated the 2-D Delaunay triangular mesh as shown in Fig. 3.
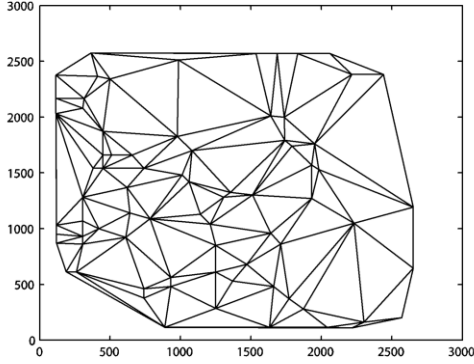

Figure 3. Delaunay triangulation result with the original coordinate scales.

- Then we double the y-coordinate and generated the new mesh as shown in Fig. 4.
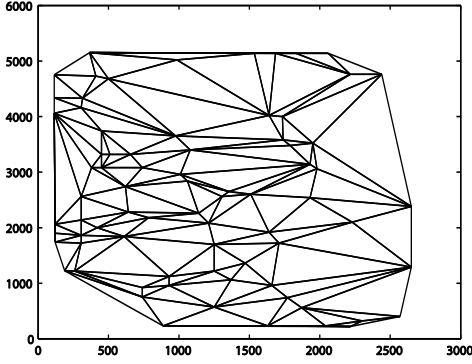

Figure 4. Delaunay triangulation result with the *y* coordinate values doubled.

It is not hard to find that the two triangulations in Figures 3 and 4 are not the same, which gives the proof to Lemma 2. In other words, the construction of a mesh is sensitive to the scales chosen for the dimensions.

Lemma 2. *3-D tetrahedral shape functions are variant to coordinate scaling before the mesh construction.*

Different meshes will eventually lead to different interpolation results, because each point to interpolate may be located in different elements. Therefore, a reliable mesh constructed with an appropriate time scale is fundamental to the success of a SF-based spatiotemporal interpolation using the extension approach.

III. APPLICATION

Particle pollution (also known as "particulate matter") in the air includes a mixture of solids and liquid droplets. Such particles are either emitted directly or form in the atmosphere when other pollutants react. Particles come in a wide range of sizes. The EPA (Environmental Protection Agency) is concerned about particles that are 10 micrometers in diameter or smaller because those are the particles that generally pass through the throat and nose and enter the lungs. Ten

micrometers are smaller than the width of a single human hair. Once inhaled, these particles can affect the heart and lungs and cause serious health effects. EPA groups particle pollution into two categories (http://www.epa.gov/air/particlepollution):

- Inhalable coarse particles (PM_{10}), such as those found near roadways and dusty industries, are larger than 2.5 micrometers and smaller than 10 micrometers in diameter.

- Fine particles ($PM_{2.5}$), such as those found in smoke and haze, are 2.5 micrometers in diameter or smaller. These particles can be directly emitted from sources such as forest fires, or they can form when gases emitted from power plants, industries and automobiles.

The data used in this study is daily $PM_{2.5}$ concentration measured in 2009 by monitoring sites over the contiguous U.S.

A. Experimental Data

The $PM_{2.5}$ data set with measurements. The data coverage contains point locations of the monitoring sites, the daily concentration level measurements of $PM_{2.5}$, and the days of the measurements. We obtained a number of data sets from the U.S. EPA (http://www.epa.gov/ttn/airs/airsaqs/detaildata) and reorganized them into a data set with schema *(id, x, y, [time], w)*, where x and y are the longitude and latitude coordinates of the monitoring sites, [time] is (year, month, day) when a $PM_{2.5}$ measurement is taken, and w is the measured $PM_{2.5}$ values. The reorganized data set has some entries with $PM_{2.5}$ values as zero, which means no measurements available at a particular site and on a particular day. After all the zero entries are filtered out, there are 146,125 daily measurements at 955 monitoring sites, which are shown as stars (*) in Fig. 5.

Figure 5. Monitoring sites with $PM_{2.5}$ measurements in 2009.

The county data set with points to interpolate. The data set with locations to interpolate are the centroids of 3,109 counties in the contiguous United States. This data set has the format of (id, x, y). The estimated $PM_{2.5}$ values at each county location and on each day in 2009 need to be computed. Therefore, there are $3,109 \times 365 = 1,134,785$ spatiotemporal $PM_{2.5}$ values to interpolate.

B. Applying SF-based Spatiotemporal Interpolation using an Appropriate Time Scale

When implementing SF-based spatiotemporal interpolation, it is important to decide the time scale, as discussed in the previous section. The four time scales shown in Table 1 have been tested.

time	ScaleA	ScaleB (ScaleA/10)	ScaleC (ScaleA/5)	ScaleD (ScaleA/15)
01/01/2009	1	0.1	0.2	0.067
01/02/2009	2	0.2	0.4	0.133
…	…	…	…	…
12/31/2009	365	36.5	73	24.333

Accuracy Assessment	ScaleA	ScaleB (ScaleA/10)	ScaleC (ScaleA/5)	ScaleD (ScaleA/15)
$\overline{MAE}$	3.1538	3.5621	3.2526	3.7344
$\overline{MSE}$	77.2331	74.5937	74.7896	73.0842
$\overline{RMSE}$	8.6521	8.4539	8.4200	8.3536
$\overline{MARE}$	3.2384	0.4286	0.3866	0.4486
$\overline{MSRE}$	5462.2800	35.4845	36.6819	35.1620
$\overline{RMSRE}$	73.3605	3.1898	3.4048	3.2607

In order to decide which is the best time scale to use for interpolation, *10-fold cross validation* [23] was implemented using the following steps.

1. The PM$_{2.5}$ data set with measurements was randomly split into ten nearly equally sized folds.

2. For each of the four time scales in Table 1, ten iterations of training and validation were performed such that within each iteration a different fold of the data was held-out for validation while the remaining nine folds were used for learning. Within each iteration, the following two actions have been taken:

    a. The spatiotemporal points in one fold (validation fold) were interpolated using the remaining nine folds (learning folds). Each point in the validation fold had both the original PM$_{2.5}$ measurement and an estimated value, after the interpolation.

    b. Six accuracy assessments were made to compare the original and estimated PM$_{2.5}$ values in the validation fold: MAE (Mean Absolute Error), MSE (Mean Squared Error), RMSE (Root Mean Squared Error), MARE (Mean Absolute Relative Error), MSRE (Mean Squared Relative Error) and RMSRE (Root Mean Squared Relative Error). They are defined as follows:

$$MAE = \frac{\sum_{i=1}^{N} |I_i - O_i|}{N} \qquad MSE = \frac{\sum_{i=1}^{N} (I_i - O_i)^2}{N}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (I_i - O_i)^2}{N}} \qquad MARE = \frac{\sum_{i=1}^{N} \frac{|I_i - O_i|}{O_i}}{N} \qquad (5)$$

$$MSRE = \frac{\sum_{i=1}^{N} \frac{(I_i - O_i)^2}{O_i}}{N} \qquad RMSRE = \sqrt{\frac{\sum_{i=1}^{N} \frac{(I_i - O_i)^2}{O_i}}{N}}$$

where *N* is the number of observations, I_i's are the interpolated values, and O_i's are the original values.

3. Since there are ten iterations and a different validation fold is chosen within each iteration, for each accuracy assessment, the average of ten accuracy results has been calculated. Table 2 shows the average results of $\overline{MAE}$, $\overline{MSE}$, $\overline{RMSE}$, $\overline{MARE}$, $\overline{MSRE}$ and $\overline{RMSRE}$. In Table 2, *ScaleA* gives different values for $\overline{MSRE}$ and $\overline{RMSRE}$ compared to the other scale results. *ScaleB*, *ScaleC* and *ScaleD* yield similar results. Since *ScaleC* shows the best result for $\overline{MARE}$, we decided to use *ScaleC* for our PM$_{2.5}$ data interpolation.

C. SF-based Interpolation Results

The SF-based interpolation for the county-level PM$_{2.5}$ data was implemented in Matlab using *ScaleC*. A total of 1,134,785 (3,109 × 365) PM$_{2.5}$ values were computed for 3,109 county centroids in the contiguous U.S. on each day in 2009.

IV.   POPULATION EXPOSURE ANALYSIS

The interpolated county-level PM$_{2.5}$ was linked to 2009 county population data. The population with a risky PM$_{2.5}$ exposure was estimated. The revised EPA National Ambient Air Quality Standards for PM$_{2.5}$ in 2006 was adopted here (http://www.epa.gov/pm/standards.html):

- 35 micrograms per cubic meter (35 µg/m^3) for 24-hours,

- 15 micrograms per cubic meter (15 µg/m^3) for the annual mean.

The results from our spatiotemporal interpolation suggest:

- there is a population of 33,147,335 (33.1million) residing in counties with an annual PM$_{2.5}$ exceeding the national standard of $15 µg/m^3 and

- more than one third of the U.S. population (111,752,669) residing in counties where PM$_{2.5}$ exceeded 35 µg/m^3 for at least one day in 2009.

Fig. 6 shows the geographic distribution of such counties with the annual and/or 24-hour PM$_{2.5}$ exceeding EPA National Ambient Air Quality Standards.
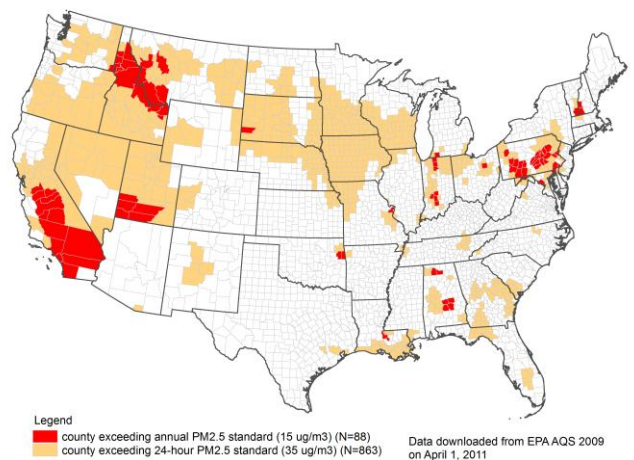

Figure 6.   Geographic distribution of counties in the contiguous United States that exceeded the PM$_{2.5}$ air quality standards in 2009.

V. Future work

First, we tested only four possible time measurement scales and chose the best one for the $PM_{2.5}$ air pollution data. A more systematic and effective method should however be developed to help decide the most appropriate time scale that should be chosen in a particular application.

Second, the use of county centroids in this study could have caused biases in county population level exposure to $PM_{2.5}$. In future studies, a finer geographic resolution, such as census block groups and tracts, may provide a more solid base for evaluating population exposure to air pollutants.

Third, it would be of great value to link air quality with population health outcomes, such as asthma and other respiratory diseases. Such future study will hopefully further support the utility of our method presented.

DISCLAIMER

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention (CDC), USA.

REFERENCE

[1] D. Shepard, "A two-dimensional Interpolation Function for Irregularly Spaced Data," in *23nd National Conference ACM*, 1968, pp. 517-524.

[2] P. Revesz and L. Li, "Representation and Querying of Interpolation Data in Constraint Databases," in *the Third National Conference on Digital Government Research*, Los Angeles, California, 2002, pp. 225-228.

[3] L. Li and P. Revesz, "Interpolation Methods for Spatio-temporal Geographic Data," *Journal of Computers, Environment and Urban Systems,* vol. 28, pp. 201-227, 2004.

[4] L. Li, *Spatiotemporal Interpolation Methods in GIS: Exploring Data for Decision Making*: VDM, 2009.

[5] P. Revesz, *Introduction to Databases: From Biological to Spatio-Temporal*. New York: Springer, 2010.

[6] D. G. Krige, "A Statistical Approach to Some Mine Valuations and Allied Problems at the Witwatersrand," Master Thesis, University of Witwatersrand, 1951.

[7] M. A. Oliver and R. Webster, "Kriging: A method of Interpolation for Geographical Information Systems," *International Journal of Geographical Information Systems,* vol. 4, pp. 313-332, 1990.

[8] D. Liao, D. J. Peuquet, Y. Duan, E. A. Whitsel, J. Dou, R. L. Smith, H.-M. Lin, J.-C. Chen, and G. Heiss, "GIS Approaches for the Estimation of Residential-Level Ambient PM Concentrations," *Environmental Health Perspectives,* vol. 114, pp. 1374-1380, 2006.

[9] J. S. Borak and M. F. Jasinski, "Effective Interpolation of Incomplete Satellite-derived Leaf-area Index Time Series for the Continental United States," *Agricultural and Forest Meteorology,* vol. 149, pp. 320-332, 2009.

[10] L. Li, "Spatiotemporal Interpolation Methods in GIS," Ph.D. Thesis, Department of Computer Science and Engineering, University of Nebraska - Lincoln, 2003.

[11] L. Li, X. Zhang, and R. Piltner, "A Spatiotemporal Database for Ozone in the Conterminous U.S.," in *the Thirteenth International Symposium on Temporal Representation and Reasoning*, Budapest, Hungary, 2006, pp. 169-176.

[12] L. Li, X. Zhang, and R. Piltner, "An Application of the Shape Function Based Spatiotemporal Interpolation Method on Ozone and Population Exposure in the Contiguous U.S.," *Journal of Environment Informatics,* vol. 12, pp. 120-128, 2008.

[13] L. Li, "Constraint Databases and Data Interpolation," in *Encyclopedia of Geographic Information System*, S. Shekhar and H. Xiong, Eds., ed: Springer, 2008, pp. 144-153.

[14] G. R. Buchanan, *Finite Element Analysis*. New York: McGraw-Hill, 1995.

[15] O. C. Zienkiewics and R. L. Taylor, *Finite Element Method, Vol. 1, The Basis*. London: Butterworth Heinemann, 2000.

[16] L. Li and P. Revesz, "A Comparison of Spatio-Temporal Interpolation Methods," in *the Second International Conference on GIScience*, 2002, pp. 145-160.

[17] L. Li, Y. Li, and R. Piltner, "A New Shape Function Based Spatiotemporal Interpolation Method," in *the First International Symposium on Constraint Databases 2004*, Paris, France, 2004, pp. 25-39.

[18] J. Gao and P. Revesz, "Voting Prediction using New Spatiotemporal Interpolation Methods," in *the Seventh Annual International Conference on Digital Government Research*, San Diego, 2006, pp. 293-300.

[19] J. Li, R. Narayanan, and P. Revesz, "A Shape-based Approach to Change Detection and Information Mining in Remote Sensing," in *Frontiers of Remote Sensing Information Processing*, C. H. Chen, Ed., ed: WSP, 2003, pp. 63-86.

[20] P. Revesz and S. Wu, "Spatiotemporal Reasoning about Epidemiological Data," *Artificial Intelligence in Medicine,* vol. 38, pp. 157-170, 2006.

[21] F. P. Preparata and M. I. Shamos, *Computational Geometry: An Introduction*: Springer-Verlag, 1985.

[22] J. E. Goodman and J. O'Rourke, *Handbook of Discrete and Computational Geometry*. Boca Raton, New York: CRC Press, 1997.

[23] P. Refaeilzadeh, L. Tang, and H. Liu, "Cross-Validation," in *Encyclopedia of Database Systems*, L. Liu and M. T. Özsu, Eds., ed: Springer, 2009, pp. 532-538.

AUTHOR BIOGRAPHIES

Dr. Lixin Li is an Associate Professor in Computer Science at Georgia Southern University. She received her B.S. and M.S. degrees in Computer Science from the Southwest Jiaotong University, Chengdu, China in 1997 and 1999, respectively. She received her Ph.D. in Computer Science in 2003 from the University of Nebraska-Lincoln. Her research interests include Geographic Information Systems (GIS) with a focus on spatiotemporal interpolation, database systems with a focus on constraint databases.

Dr. Xingyou Zhang is a geographer and statistician at Division of Population Health (proposed), Centers for Disease Control and Prevention (CDC). He holds a Ph.D. in Geography (2004) with an emphasis on GIS and spatial analysis and a M.S. in Statistics (2003) from University of Cincinnati. His current research focuses on geographic contextual effects (social and physical environments) on chronic disease outcomes, such as obesity, physical activity, Chronic Obstructive Pulmonary Disease (COPD), and asthma. His methodological interests include Bayesian disease mapping, multilevel spatial modeling and analysis and small area estimation.

Dr. James Holt is the Team Leader for Analytic Methods, in the CDC National Center for Chronic Disease Prevention and Health Promotion. Jim has been with the U.S. Centers for Disease Control and Prevention (CDC) in Atlanta since 1992. One of a small group of Ph.D. trained geographers at the CDC, his research focuses on the application of geospatial analysis methods and cartographic visualization techniques for chronic disease prevention and health promotion.

Dr. Jie Tian is an Assistant Professor in the Department of Geology and Geography at Georgia Southern University. He received his Ph.D. (2009) and Master's (2004) both in Geography (GIS and Remote Sensing) from Queen's University, Kingston, Ontario and the University of Western Ontario, London, Ontario, respectively. His undergraduate training was received in Earth Science (major) and General Science (minor) from Peking University, China. His research interests broadly focus on Geospatial Science & Technology and their application in Environmental Monitoring and Modeling, Landscape Ecology, Public Health, Air Quality, Renewable Energy, etc.

Dr. Reinhard Piltner received his doctoral degree from the Ruhr-University in Bochum, Germany while working at the Institute for Mechanics in the Computational Mechanics group. During postdoctoral research at the University of California at Berkeley in the Structural Engineering, Mechanics and Materials group, he concentrated on finite element and boundary element methods while keeping his interest on theoretical and applied mechanics. He taught a variety of courses at the Ruhr-University Bochum, University of Nebraska and Georgia Southern University. His current interests include Biomedical Engineering and CT image based 3D simulations.