

An Application of a Shape Function Based Spatiotemporal Interpolation Method to Ozone and Population-Based Environmental Exposure in the Contiguous U.S.

L. Li^{1*}, X. Zhang², and R. Piltner³

¹Department of Computer Sciences, Georgia Southern University, Statesboro, GA 30460-7997, USA

²The Robert Graham Center for Policy Studies in Family Medicine and Primary Care, Washington, D.C. 20036, USA

³Department of Mathematical Sciences, Georgia Southern University Statesboro, GA 30460-8093, USA

Received 18 May 2007; revised 15 March 2008; accepted 9 April 2008; published online 1 December 2008

ABSTRACT. It is important to conduct research on the connection between air pollution and human health using population-based spatiotemporal environmental exposure assessment on a large scale with respect to area and population. In our paper, using a set of spatiotemporal data with annual ozone concentration measurements in the contiguous U.S. during 1994 and 1999, we address the following challenging issues in conducting such research: spatiotemporal interpolation, comparison of spatiotemporal interpolation methods, visualization, and analysis of population exposure to ozone. A 3D shape function based spatiotemporal interpolation method has been used in this paper to estimate the ozone concentrations at any unmeasured location and time. Using the leave-one-out cross-validation, we compute error statistics to compare the shape function and IDW (Inverse Distance Weighting) methods. For the considered case studies it is observed that the shape function method is better than IDW in terms of MAPE (Mean Absolute Percentage Error) and algorithm complexity. For generating maps of annual ozone concentrations, we propose a new approach to select locations to interpolate and visualize: picking U.S. census block centroids as sample locations. The advantage of this approach is to generate more sample points in areas with more intensive human activities. In our experiment, there were about 8,000,000 sample points selected per year. Traditional GIS techniques are insufficient in handling such kind of spatiotemporal data. The visualization results of ozone concentration distribution at the census tract level in the contiguous U.S. from 1994 to 1999 are illustrated. We also analyze the population exposure to ozone in the year 1999 according to different ozone concentration levels following the recommendations given by the U.S. EPA on air quality. Our finding is that in the year of 1999, 9.8% total population in the contiguous U.S. has been exposed to a high risk ozone level, 78.7% to a moderate risk, and only 11.5% to a low risk.

Keywords: shape functions, spatiotemporal interpolation, ozone, population exposure, leave-one-out cross-validation, geographic information systems (GIS)

1. Introduction

It is important to investigate the association between air pollution and human health. For example, despite its complex etiology and significant, well-documented sociodemographic disparities (Akinbami et al., 2002), the increase in childhood asthma may be better explained by environmental changes than by genetic or evolutionary changes. Evidence regarding the contribution of air pollution is accumulating (Moore and Bates, 2001). Increased childhood asthma hospital visits are linked to the proximity of local traffic and other polluting sites (Oyana and Rivers, 2005; Pandya et al., 2002). Environmental exposure assessment is a critical analytical tool to conduct such research. Population-based environmental exposure modeling and analysis have made significant progress in the

last two decades (Nieuwenhuijsen et al., 2006), especially with the advances in modern geospatial technologies — Geographic Information System (GIS) and Remote Sensing and Global Positioning System (GPS), and probability techniques — Markov Chain Monte Carlo (MCMC) simulation and Bayesian statistical modeling (Briggs, 2005; Christakos and Serre, 2000; Elgethun et al., 2003; Jarup, 2004; Nuckols et al., 2004).

Appropriate interpolation is critical to the environmental exposure assessment. Common spatial interpolation methods, such as *kriging* (Deutsch and Journel, 1998; Krige, 1966), *inverse distance weighting* (IDW) (Shepard, 1968) and *splines* (Goodman and Rourke, 1997; Schoenberg, 1946), are available in major GIS software packages. A review and comparison of spatial interpolation methods are given in (Lam, 1983). These interpolation methods have been used to assess spatial variations in environmental risk factors in most environmental epidemiological studies.

Despite the need for traditional spatial interpolation, more and more environmental exposure studies require the interpolation of temporal data. The temporal dimension of datasets is often ignored, underemphasized, or isolated from the spatial

* Corresponding author. Tel.: +1 912 4867646; fax: +1 912 4867672.

E-mail address: lli@georgiasouthern.edu (L. Li).

domain mainly due to the lack of efficient spatiotemporal interpolation methods. With the additional *time* attribute, the traditional spatial interpolation methods are insufficient, and new spatiotemporal interpolation methods must be developed. For example, suppose that we have recorded data on carbon monoxide at different monitoring sites at different times. Some spatiotemporal interpolation method is needed to estimate the carbon monoxide values at unsampled locations and times to determine the range of potential threat. Currently most researchers assume that time can be incorporated into spatiotemporal interpolation by conducting a *sequence* of spatial interpolations and treat space and time separately. For example, a GIS-based geostatistical approach used kriging to do spatial interpolation for environmental exposure analysis on a sequence of daily ambient particulate matter (PM) concentrations (Liao et al., 2006).

A few exceptions in the literature do investigate spatiotemporal interpolation methods that integrate space and time simultaneously. *Shape functions*, which can be viewed as a spatial interpolation method, are popular in engineering applications; for example, in Finite Element algorithms (Zienkiewicz and Taylor, 2000). Considering piecewise linear approximations, linear shape functions for triangles can be used in a 2D mesh, or 3D linear shape functions for tetrahedrons can be used in a 3D mesh. Shape function-based spatiotemporal interpolation methods have been developed since 2003 and are used in GIS applications (Li and Revesz, 2004; Li et al., 2004; Li, 2003). Several methods have been compared (Li and Revesz, 2002; Li and Revesz, 2004) and applied. For example, shape function interpolation methods were applied to a set of ozone data in the conterminous United States (Li et al., 2006; Li and Zhang, 2007; Li et al., 2007) and a real estate dataset with house prices (Li and Revesz, 2004). Another was used to represent West Nile virus data in constraint databases (Revesz, 2002) and implemented a particular epidemiological system called WeNiVIS that enables the visual tracking of, and reasoning about, the spread of the epidemic in Pennsylvania (Revesz and Wu, 2006).

Furthermore, most of the current research on the relationship between air pollution and asthma has focused on a specific area and study population, such as a small group of asthmatics who reside in downtown areas (Kim et al., 2007), hikers in the Great Smoky Mountains National Park (Girardot et al., 2006), children spending time at swimming pools (Lagerkvist et al., 2004), and infants delivered at certain hospitals in southwestern Virginia (Triche et al., 2006). The results of these studies vary. One indicates that exposure to relatively low concentrations of ozone influences the symptoms of moderate-to-severe asthmatics, regardless of changes in pulmonary function or medication use (Kim et al., 2007), while another suggests no significant association of acute changes in pulmonary function with ozone (O_3) and fine particulate matter ($PM_{2.5}$) (Girardot et al., 2006). The investigation of the relationship between air pollution and asthma can be improved by conducting environmental exposure analysis in a much larger spatial extent.

In our paper, we investigate human population exposure to

ozone in the large spatial scale of the contiguous United States during 1994 and 1999 using a shape function-based spatiotemporal interpolation method. The rest of this paper is organized as follows. Section 2 describes the spatiotemporal ozone dataset. Section 3 introduces general approaches of spatiotemporal interpolation methods; summarizes the advantages of a particular type of method called shape function-based extension methods; and describes the 3D shape function-based spatiotemporal interpolation method introduced by Li and Revesz (2004). In Section 4, the shape function-based method is compared to the IDW-based spatiotemporal interpolation method using the leave-one-out cross-validation by applying both methods on the ozone dataset to obtain the interpolation results. Several types of error statistics are computed to compare the shape function-based and IDW-based methods. In the comparison the shape function-based method provided better results than IDW in terms of MAPE (Mean Absolute Percentage Error) and algorithm complexity. In Section 5, we illustrate the continuous distribution of annual ozone concentrations by six static maps for the six years during 1994 and 1999. Section 6 gives the analysis of population exposure to ozone at the census block level based on the shape function-based spatiotemporal interpolation result. Finally, in Section 7 we give conclusions.

2. Spatiotemporal Ozone Dataset

The spatiotemporal ozone data we are interested in this paper is a set of AIRS (Aerometric Information Retrieval System) data with annual ozone concentration measurements in the contiguous United States (website <http://www.epa.gov/airmarkets/cmap/data/category1.html>). AIRS is a computer based repository of information about airborne pollution in the United States and various World Health Organization (WHO) member countries. The system is administered by the U.S. Environmental Protection Agency (EPA).

We obtained several datasets from the U.S. EPA website (<http://cfpub.epa.gov/gdm>) and reorganized them into a dataset with schema (x, y, t, w) , where x and y attributes are the longitude and latitude coordinates of monitoring sites where the ozone concentrations were collected; t is the year when the ozone concentrations were taken; w records the annual ozone concentration measurements $O34MAX$ (4th max of 1-hour values for O_3).

The reorganized dataset initially has no ozone concentrations at 2904 monitoring sites, which means no measurements was available at those sites. After filtering out the entries at those sites from the dataset, there are 1209 sites left with measurements. Among the 1209 monitoring sites with measurements, some sites have complete measurements of yearly ozone values from 1994 to 1999, while the other sites have only partial records. For example, some sites only have measurements of ozone values from 1998 to 1999. In total, there are 6135 ozone concentration measurements recorded. Each measurement corresponds to the ozone value at a spatiotemporal point (x, y, t) , where (x, y) is the location of one of the 1209 monitoring sites, and t is a year between 1994 and 1999. Table 1 shows instances of ozone measurements at two sites. Figure 1

shows the locations of the 1209 monitoring sites with measurements.

Table 1. Ozone (x, y, t, w)

x	y	t	w
-80.6833	25.3883	1994	0.068
-80.6833	25.3883	1995	0.068
-80.6833	25.3883	1996	0.067
-80.6833	25.3883	1997	0.083
-80.6833	25.3883	1998	0.086
-80.6833	25.3883	1999	0.08
...
-82.5461	27.632799	1998	0.085
-82.5461	27.632799	1999	0.097
...



Figure 1. 1209 AIRS monitoring sites with measurements in the contiguous United States.

3. Shape Function-Based Spatiotemporal Interpolation

3.1. Spatiotemporal Interpolation

There are two fundamentally different ways for spatiotemporal interpolation: reduction and extension (Li and Revesz, 2002).

- **Reduction.** This approach reduces the spatiotemporal interpolation problem to a regular spatial interpolation case using two steps. First, interpolate (using any 1-D interpolation in time) the measured value over time at each sample point. Second, get spatiotemporal interpolation results by substituting the desired time instant into some regular spatial interpolation functions.
- **Extension.** This approach deals with time as another dimension in space and extends the spatiotemporal interpolation problem into a one-higher dimensional spatial interpolation problem.

In general, any spatial interpolation method, such as IDW, kriging, and shape function-based methods, can be developed to a spatiotemporal interpolation method either using the reduction or the extension approach.

Table 2. Difference Coordinate Scales

x-dimension	y-dimension	time-dimension
Meter	Meter	minute
Foot	Foot	hour
Meter	Foot	second

3.2. Advantages of 3D Shape Function-Based Extension Methods

The ozone dataset described in Section 2 is in the domain of 2D space (x, y) and 1D time (t). In this paper, we choose to use the 3D shape function-based spatiotemporal interpolation method introduced in (Li and Revesz, 2004) based on the extension approach. This is because the 3D shape function-based extension interpolation method has the following advantages:

- **Invariance to coordinate scales.** The 3D shape function-based spatiotemporal extension method is invariant to coordinate scales, which means the results will remain the same even if the scale of a dimension (or dimensions) changes. We will not have to worry about what space and time units should be used. For example, using the following three sets of scales on a dataset in 2D space and 1D time as shown in Table 2 will produce the same interpolation results if a shape function-based method is used. However, all the reduction approach based methods and other extension approach based methods, for example using IDW and kriging, are not invariant to coordinate scales. The proof of the invariance of shape function-based interpolation methods and the noninvariance of other methods can be found in (Li and Revesz, 2004).
- **Accuracy.** The 3D shape function-based spatiotemporal interpolation method was compared with other extension approach based methods such as IDW and kriging using a set of real estate data. The comparison results show that the shape function-based method is the most accurate in terms of MAE (Mean Absolute Error) and RMSE (Root Mean Square Error) and the overall best spatiotemporal interpolation method for the discussed examples (Li and Revesz, 2004).
- **Fast performance.** Since the 3D shape function-based extension method is only linear (Li and Revesz, 2004), the simulations are very efficient. For example, when the 3D shape function-based spatiotemporal interpolation method was used to interpolate the annual ozone concentration values at more than 8,000,000 census block centroids per year in the contiguous United States (Li and Zhang, 2007), it took a regular desktop computer with 2.26 GHz CPU and 760 MB RAM no more than four minutes to execute a Matlab program with the shape function algorithm. On the other hand, other extension approach based methods, for example using IDW and kriging, are non-linear. Spatiotemporal interpolation methods based on the reduction approach are at least quadratic (Li and Revesz, 2004). Therefore they are not as efficient as the 3D shape function-based extension method.
- **Easy storage in constraint databases.** The interpolation

results of the 3D shape function-based extension method can be accurately and efficiently stored in constraint databases (Revesz, 2002), which can yield powerful solutions to useful queries (Li and Revesz, 2002; Li and Revesz, 2004; Li et al., 2006; Li, 2003).

3.3. 3D Shape Function-based Spatiotemporal Interpolation Method

For datasets in the domain of 2D space and 1D time such as the ozone dataset in Section 2, we can apply the 3D shape function based extension method to interpolate. Interpolation by the 3D shape functions requires dividing the (x, y, t) domain into a finite number of tetrahedra with the sampled data points as corner vertices. Figure 2 illustrates one tetrahedron in the tetrahedral mesh. Assume the unmeasured point (x, y, t) is located inside the tetrahedron and we would like to interpolate the value at this point based on the measured values w_1, w_2, w_3, w_4 at the corner vertices. The spatiotemporal interpolation equation is (Li and Revesz, 2004):

$$w(x, y, t) = N_1(x, y, t)w_1 + N_2(x, y, t)w_2 + N_3(x, y, t)w_3 + N_4(x, y, t)w_4 \quad (1)$$

where N_1, N_2, N_3 and N_4 are the following 3D shape functions:

$$N_i(x, y, t) = \frac{V_i}{V}, \quad i = 1, 2, 3, 4 \quad (2)$$

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

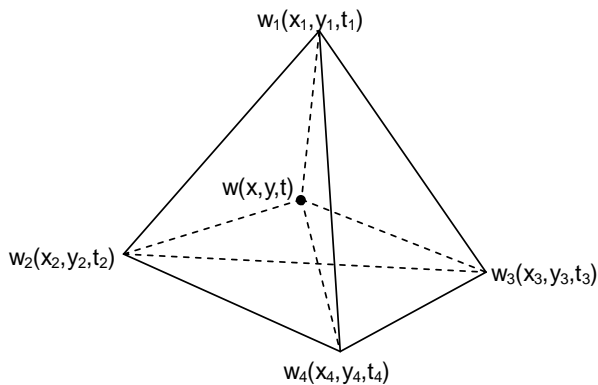


Figure 2. Spatiotemporal interpolation using 3D shape functions.

4. Comparison with IDW Interpolation

Li and Revesz (2004) compare different spatiotemporal interpolation methods such as shape functions, IDW (Inverse

Distance Weighting), and kriging methods based on a set of actual real estate data set with house prices. The comparison criteria include interpolation accuracy, error-proneness to time aggregation, invariance to scaling on the coordinate axes, and the type of constraints used in the representation of the interpolated data. The experimental results indicate that the extension method based on shape functions is the most accurate and the overall best spatiotemporal interpolation method.

In this section, we discuss the IDW based spatiotemporal interpolation method and compare the shape function method and the IDW method based on the ozone dataset introduced in Section 2.

4.1. Spatiotemporal Interpolation Based on IDW

As other spatial interpolation methods, IDW (Inverse Distance Weighting) interpolation is based on the assumption that things that are close to one another are more alike than those that are farther apart. (Revesz and Li, 2002) uses IDW to visualize spatial interpolation data. In IDW, the measured values (known values) closer to an estimation location will have more influence on the estimated value (unknown value) than those farther away. More specifically, IDW assumes that each measured point has a local influence that diminishes with distance. Thus, points in the near neighborhood are given high weights, whereas points at a far distance are given small weights.

The general formula of IDW interpolation is the following:

$$w(x, y) = \sum_{i=1}^N \lambda_i w_i, \quad \lambda_i = \frac{(1/d_i)^p}{\sum_{k=1}^N (1/d_k)^p} \quad (3)$$

where $w(x, y)$ is the estimated value at location (x, y) , N is the number of nearest known points surrounding (x, y) , λ_i s are the weights assigned to each measured value w_i at location (x_i, y_i) , d_i s are the Euclidean distances between each (x_i, y_i) and (x, y) , and p the exponent, which influences the weighting of w_i on w .

Although IDW is originally a spatial interpolation method, we can apply the reduction and extension approaches to IDW to obtain spatiotemporal interpolation methods based on IDW (Li and Revesz, 2004). In our experiment, we apply extension method based on IDW to interpolate our ozone dataset. The interpolation formula is of the form of Formula (3) with $d_i =$

$$\sqrt{(x_i - x)^2 + (y_i - y)^2 + (t_i - t)^2} .$$

4.2. Performance Comparison between Shape Function and IDW

For our experimental data, the observation points are the spatiotemporal points (x, y, t) , where (x, y) is the location of a monitoring site and t is the year when the ozone concentration was taken. In the end of the process of the leave-one-out cross-validation, each of the observation points not only has its original ozone concentration, but also has an interpolated value. The original and interpolated concentrations at each

Table 3. Major Statistics of the Shape-Function-Based Spatiotemporal Interpolation Errors for U.S. Ozone Data (1994-1999)

Year	N	MIN	MAX	MAE	MBE	RMSE	MAPE
1994	941	-0.081	0.089	0.012	0.000	0.017	12.359
1995	990	-0.099	0.068	0.009	0.002	0.013	9.361
1996	1000	-0.085	0.064	0.008	-0.003	0.011	9.194
1997	1041	-0.065	0.039	0.008	-0.001	0.011	9.117
1998	1063	-0.118	0.080	0.009	0.002	0.013	9.470
1999	1064	-0.078	0.064	0.008	-0.001	0.012	9.393
1994 - 1999	6099	-0.118	0.089	0.009	0.000	0.013	9.779

Table 4. Major Statistics of the IDW (with 5 Neighbors and Exponent 2) Spatiotemporal Interpolation Errors for U.S. Ozone Data (1994-1999)

Year	N	MIN	MAX	MAE	MBE	RMSE	MAPE
1994	956	0.045	0.223	0.010	0.000	0.014	10.702
1995	993	0.036	0.209	0.010	0.000	0.014	10.389
1996	1004	0.039	0.205	0.009	0.000	0.012	9.187
1997	1041	0.043	0.172	0.009	0.001	0.012	9.550
1998	1063	0.032	0.211	0.009	0.000	0.014	10.015
1999	1078	0.01	0.171	0.008	0.000	0.011	9.134
1994 - 1999	6135	0.01	0.223	0.009	0.000	0.013	9.813

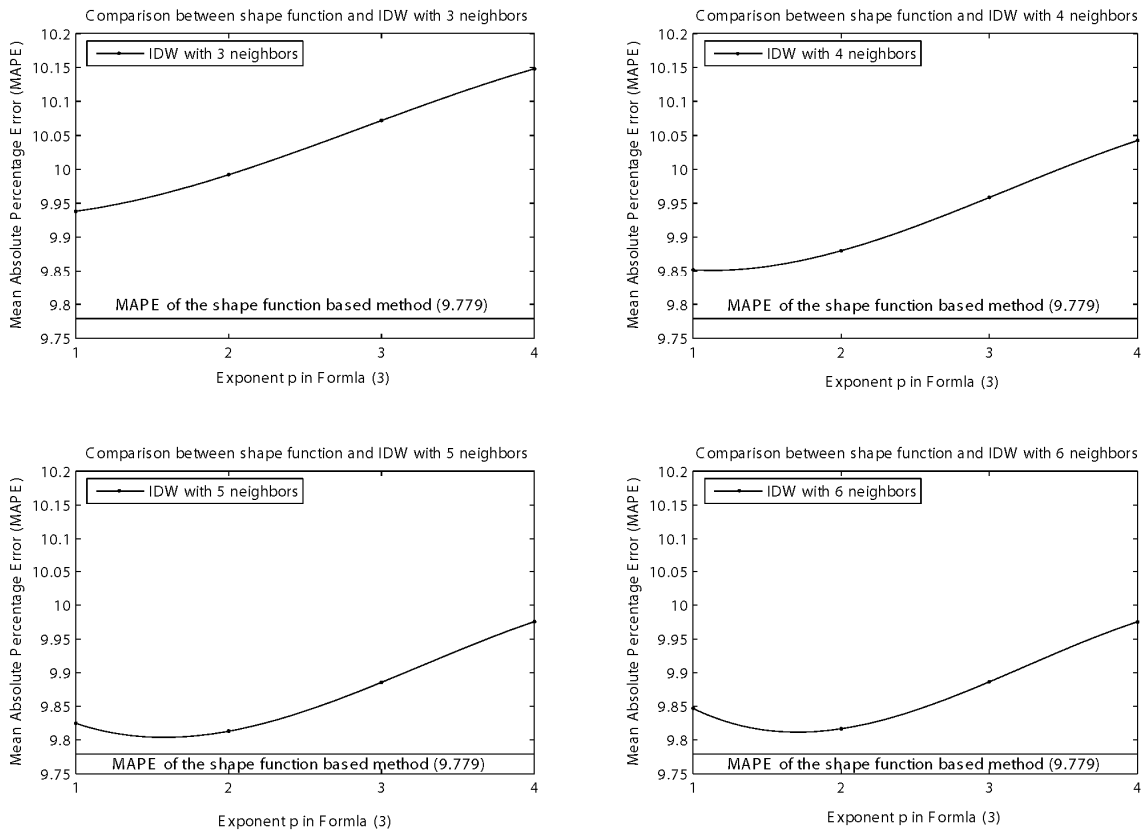


Figure 3. Accuracy comparison of the shape function method and the IDW method.

observation point can be compared for the purpose of an error analysis.

Based on the results of the leave-one-out cross-validation,

we use the following four types of error statistics to compare the performance of the shape function and the IDW interpolation methods: MAE (Mean Absolute Error), MBE (Mean

Biased Error), RMSE (Root Mean Squared Error) and MAPE (Mean Absolute Percentage Error). The less the error, the better the performance. Each of the error statistics is calculated for each year and the whole period. The definitions of MAE, MBE, RMSE and MAPE are as follows:

$$MAE = \frac{\sum_{i=1}^N |I_i - O_i|}{N}, MBE = \frac{\sum_{i=1}^N (I_i - O_i)}{N},$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (I_i - O_i)^2}{N}}, MAPE = \frac{\sum_{i=1}^N \frac{|I_i - O_i|}{O_i}}{N}. \quad (4)$$

where N is the number of observations, I_i is the interpolated value, and O_i is the original value. Since the ozone concentration values are quite small, the error statistics such as MAE, MBE and RMSE have small values. We think the MAPE is a better statistic error measurement for the ozone dataset since it measures percentage errors.

4.2.2. Error Analysis of the Shape Function Interpolation Method

The shape function interpolation method can interpolate values inside the convex hull after generating the tetrahedral mesh. It cannot interpolate for the points that are outside the convex hull. This can be considered as a limitation of the shape function-based method. There are 6135 original spatio-temporal observation points in the ozone dataset (see Section 2). The leave-one-out cross-validation method requires generating 6135 different tetrahedral meshes with 6134 vertices for each mesh. It is possible that for a particular tetrahedral mesh, the point to be estimated is outside the convex hull. In this case, the value at this selected point cannot be interpolated. In our experiment, the number of such points that cannot be interpolated is 36. Therefore, the rest 6099 points with interpolation results are used for our error analysis.

Applying the error analysis approach described in Section 4.2.1 to the shape function based extension method, Table 3 shows that the 1994 spatiotemporal interpolation has the largest MAPE as expected, since 1994 is the year with the most missing data. There is no significant difference during 1995 ~ 1999. The overall MAPE during 1994 ~ 1999 is 9.779%.

4.2.3. Error Analysis of the IDW Interpolation Method

The IDW interpolation method is dependent on the number of nearest neighbors and the exponent [see Formula (3)]. We applied the following 16 IDW-based extension methods to the ozone dataset and performed the leave-one-out cross-validation:

- IDW with 3 neighbors and exponent 1, 2, 3 and 4.
- IDW with 4 neighbors and exponent 1, 2, 3 and 4.
- IDW with 5 neighbors and exponent 1, 2, 3 and 4.
- IDW with 6 neighbors and exponent 1, 2, 3 and 4.

Figure 3 illustrates the overall MAPE (Mean Absolute Percentage Error) results of the above 16 IDW methods, compared with the MAPE of the shape function based interpolation method. As shown in Figure 3 that the IDW method with 5 neighbors and exponent 2 has the best performance in terms of the overall MAPE, which is 9.813%. Table 4 shows the detailed error analysis results of this IDW approach.

4.2.4. Comparison of Shape Function and IDW

First, it is clear from Figure 3 that the shape function based method is the best in terms of the overall MAPE. The overall MAPE (Mean Absolute Percentage Error) of the 3D shape function-based method is 9.779% during 1994 and 1999, compared to the 9.813% overall MAPE result of the best IDW method with 5 neighbors and exponent 2. Second, it is hard to find the optimal configuration of n (the number of neighbors) and p (exponent) for the IDW interpolation method. After the leave-one-out cross-validation, we find out that for our ozone dataset, the IDW method with $n = 5$ and $p = 2$ is the best IDW method. But even this IDW method is still not as good as the shape function based interpolation method. Third, the shape function based interpolation method is more efficient than the IDW method in terms of the algorithm complexity. The shape function based extension method is only linear, while the IDW interpolation method is polynomial because of the Euclidean distance calculation and the exponent p in Formula (3).

5. Visualization

First, we need to select numerous locations in the contiguous U.S. where the ozone values should be interpolated based on the set of measured ozone data. The possible existing methods of taking sample points are to take random points and to take evenly spaced points in the area extent of interest. However, these existing methods do not consider the human population activities factor on the ground. In our work, we propose a new method to pick U.S. census block centroids as sample locations to be interpolated, which generates more sample points in the areas with more intensive human activities. In our experiment, there were about 8,000,000 sample points selected per year. Traditional GIS techniques are insufficient in handling such kind of spatiotemporal data. Second, we applied the shape function based spatiotemporal interpolation method to interpolate the ozone concentrations at the selected 8,000,000 sample locations per year from 1994 and 1999. Third, we generated continuously changing maps for each year to visualize the interpolation results by ArcGIS. Note that in our visualization procedure, the basic interpolation is at the census block level. The spatial patterns of population exposure to ozone could be displayed at multiple scales, such as census tracts, counties and states, and linked with the health outcomes and other risk factors at different scales. Figures 4 to 9 show the visualization results of ozone concentration distribution at the census tract level in the contiguous U.S. from 1994 to 1999.

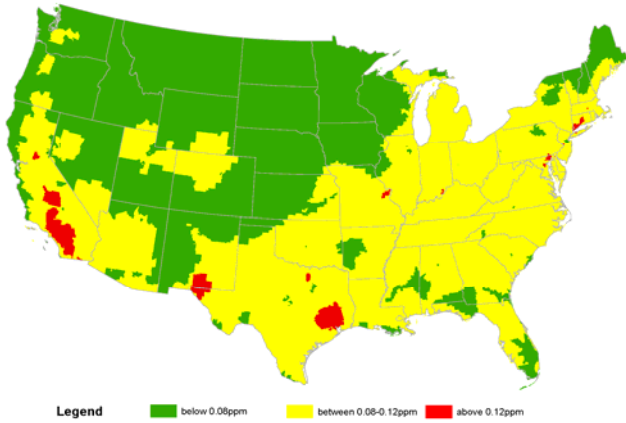


Figure 4. Ozone concentration visualization for 1994 at census tract level.

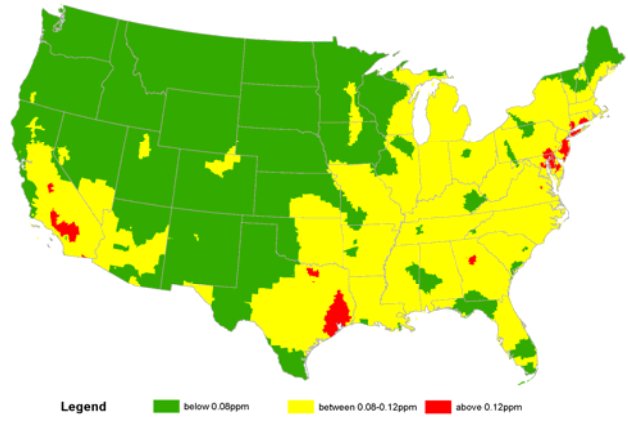


Figure 7. Ozone concentration visualization for 1997 at census tract level.

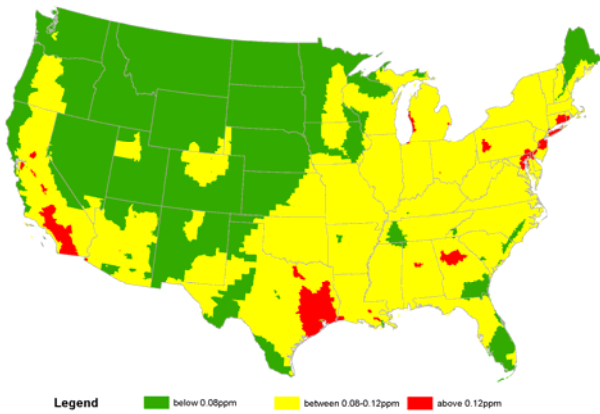


Figure 5. Ozone concentration visualization for 1995 at census tract level.

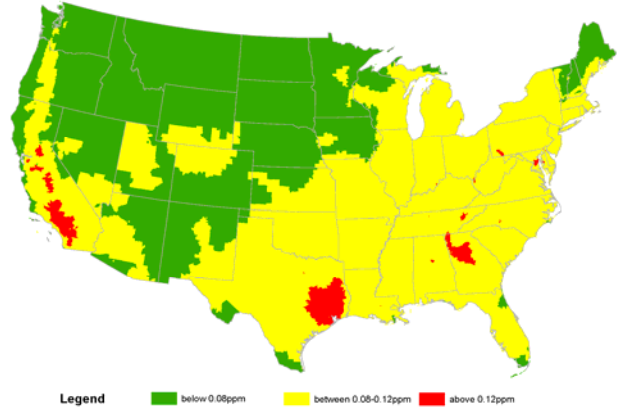


Figure 8. Ozone Concentration Visualization for 1998 at Census Tract Level.

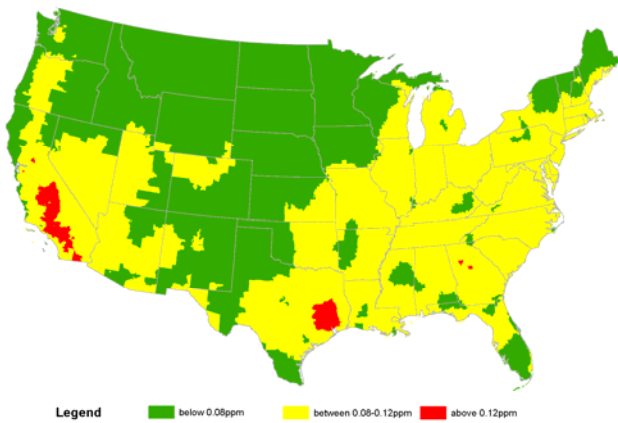


Figure 6. Ozone concentration visualization for 1996 at census tract level.

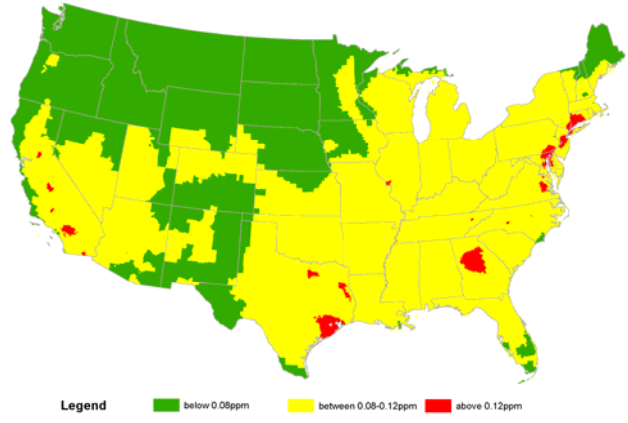


Figure 9. Ozone concentration visualization for 1999 at census tract level.

6. Analysis of Population Exposure to Ozone

High ozone concentration level has a series of health effects on human population. Ozone is one of air pollutants significantly associated with asthma and other respiratory diseases (Donnay, 2004; Girardot et al., 2006; Lagerkvist et al., 2004; Thurston and Bates, 2004). Some studies show even exposure to relatively low concentrations of ozone influences the symptoms of moderate to severe asthmatics (Kim et al., 2007; Triche et al., 2006). According to the recommendations given by the U.S. EPA on air quality (website <http://www.epa.gov/03healthtraining/aqi.html>), we used the block centroid's ozone concentration level for the whole block and estimated the population exposed to the following three risks according to different ozone concentration levels:

- lower risk: ozone concentration level is below 0.08;
- moderate risk: ozone concentration level is between 0.08 and 0.12 ppm;
- high risk: ozone concentration level is above 0.12 ppm.

We obtained the 1999 census block ozone levels by the shape function interpolation. As shown in Table 5, in the year of 1999, 9.8% total population in the contiguous United States has been exposed to a high risk ozone level, 78.7% to a moderate risk, and only 11.5% to a low risk.

Table 5. The Population Exposed to Different Ozone Levels in the Contiguous U.S.

Risk Level	1999 Population	Percentage
Low	32,137,501	11.49%
Moderate	219,933,840	78.66%
High	27,513,906	9.84%

7. Conclusions

Population-based spatiotemporal environmental exposure assessment on a large scale with respect to area and population has been conducted. Spatiotemporal interpolation is a crucial tool in this kind of research. The 3D shape function-based method has been selected as the spatiotemporal interpolation method in this paper. This method has been compared with the IDW-based method based on the leave-one-out cross-validation results. It is shown that the shape function method is better than IDW in terms of MAPE (Mean Absolute Percentage Error) and algorithm complexity.

Continuous maps of annual ozone concentration distributions have been generated for the contiguous United States during 1994 and 1999. A new approach to select locations to interpolate and visualize has been proposed and implemented: picking U.S. census block centroids as sample locations. The advantage of this approach is to generate more sample points in the areas with more intensive human activities. In our experiment, there were about 8,000,000 sample points selected per year. Traditional GIS techniques are insufficient in handling such kind of spatiotemporal data.

The population exposure to ozone in the year of 1999 has

been analyzed according to different ozone concentration levels following the recommendations given by the U.S. EPA on air quality. Our finding is that in the year of 1999, 9.8% total population in the contiguous U.S. has been exposed to a high risk ozone level, 78.7% to a moderate risk, and only 11.5% to a low risk.

In the future, we plan to continue to compare the performances of the shape function-based interpolation method with other extension approach-based methods such as using kriging and Lagrange, using the same set of ozone concentration data and using the leave-one-out cross-validation.

Based on the interpolation results, we will also further explore the socio-spatial disparities in population ozone exposure to exam the hypothesis that low socioeconomic and minority populations are more likely to be exposed to risky ozone concentration levels than other population groups.

References

- Akinbami, L.J., LaFleur, B.J. and Schoendorf, K.C. (2002). Racial and income disparities in childhood asthma in the United States, *Ambulatory Pediatrics*, 2(5), 382-387, doi:10.1367/1539-4409(2002)002<0382:RAIDIC>2.0.CO;2.
- Briggs, D. (2005). The role of GIS: coping with space (and time) in air pollution exposure assessment, *J. Toxicol. Environ. Health A*, 68(13-14), 1243-1261, doi:10.1080/15287390590936094.
- Christakos, G. and Serre, M.L. (2000). Spatiotemporal analysis of environmental exposure-health effect associations, *J Expo Anal Environ Epidemiol.*, 10(2), 168-87.
- Deutsch, C.V. and Journel, A.G. (1998). *GSLIB: Geostatistical Software Library and User's Guide*, Oxford University Press, New York.
- Donnay, A. (2004). Ozone and asthma, *J. Am. Med. Assoc.*, 423.
- Elgethun, K., Fenske, R.A., Yost, M.G. and Palcisko, G.J. (2003). Time-location analysis for exposure assessment studies of children using a novel global positioning system instrument, *Environ. Health Perspect.*, 111(1), 115-22, doi:10.1289/ehp.5350.
- Girardot, S.P., Ryan, P.B., Smith, S.M., Davis, W.T., Hamilton, C.B., Obenour, R.A., Renfro, J.R., Tromatore, K.A. and Reed, G.D. (2006). Ozone and PM_{2.5} Exposure and Acute Pulmonary Health Effects: A Study of Hikers in the Great Smoky Mountains National Park, *Environ. Health Perspect.*, 114(7), 1044-1052, doi:10.1289/ehp.8637.
- Goodman, J.E., and Rourke, J.O. (1997). *Handbook of Discrete and Computational Geometry*, Boca Raton, New York.
- Hjorth, U. (1994). *Computer Intensive Statistical Methods, Validation, Model Selection, and Bootstrap*, Chapman and Hall.
- Jarup, L. (2004). Health and environment information systems for exposure and disease mapping, and risk assessment, *Environ. Health Perspect.*, 112(9), 995-7, doi:10.1289/ehp.6736.
- Kim, D.H., Kim, Y.S., Park, J.S., Kwon, H.J., Lee, K.Y., Lee, S.R. and Jee, Y.K. (2007). The Effects of On-site Measured Ozone Concentration on Pulmonary Function and Symptoms of Asthmatics, *J. Korean Med. Sci.*, 22(1), 30-36.
- Krige, D. (1966). Two dimensional weighted moving average trend surfaces for ore evaluation, *Journal of the Society of African Institutes of Mining and Metallurgy*, 66, 13-38.
- Lagerkvist, B.J., Bernard, A., Blomberg, A., Bergstrom, E., Forsberg, B., Holmstrom, K., Karp, K., Lundstrom, N.G., Segerstedt, B., Svensson, M., and Nordberg, G. (2004). Pulmonary Epithelial Integrity in Children: Relationship to Ambient Ozone Exposure and

- Swimming Pool Attendance, *Environ. Health Perspect.*, 112(17), 1768-1771, doi:10.1289/ehp.7027.
- Lam, N.S. (1983). Spatial Interpolation Methods: A Review, *The American Cartographer*, 10(2), 129-149, doi:10.1559/152304083783914958.
- Li, L. and Revesz, P. (2002). A Comparison of Spatio-Temporal Interpolation Methods, *Proc. of the Second International Conference on GIScience*, 2002, 145-160.
- Li, L. and Revesz, P. (2004). Interpolation methods for spatio-temporal geographic data, *Computers, Environment and Urban Systems*, 28(3), 201-227, doi:10.1016/S0198-9715(03)00018-8.
- Li, L., Zhang, X. and Piltner, R. (2006). A Spatiotemporal Database for Ozone in the Conterminous U.S., *Proc. of the Thirteenth International Symposium on Temporal Representation and Reasoning*, IEEE, Budapest, Hungary, 168-176.
- Li, L., Li, Y. and Piltner, R. (2004). A New Shape Function Based Spatiotemporal Interpolation Method, *Proc. of the First International Symposium on Constraint Databases 2004*, 25-39.
- Li, L. (2003). *Spatiotemporal Interpolation Methods in GIS*, Ph.D. Thesis, University of Nebraska-Lincoln, Lincoln, Nebraska.
- Li, L. and Zhang, X. (2007). Estimation of Population Exposure to Ozone Using Shape Function Interpolation Methods, *Proc. of the Second National Conference on USGS Health-Related Research*, USGS, Reston, VA, 43.
- Li, L., Zhang, X. and Piltner, R. (2007). Spatiotemporal Interpolation and Constraint Databases for a GIS Application: Ozone in the Contiguous U.S., *Proc. of the Dagstuhl-Seminar on Constraint Databases*, Geometric Elimination and GIS, Dagstuhl, Germany.
- Liao, D., Peuquet, D.J., Duan, Y., Whitsel, E.A., Dou, J., Smith, R.L., Lin, H.M., Chen, J.C. and Heiss, G. (2006). GIS Approaches for the Estimation of Residential-Level Ambient PM Concentrations, *Environ. Health Perspect.*, 114(9), 1374-1380, doi:10.1289/ehp.9169.
- Moore, C. and Bates, D. (2001). *Air Pollution Causes Asthma: A Review of Recent Studies*, Health and Clean Air Newsletter.
- Nieuwenhuijsen, M., Paustenbach, D. and Duarte-Davidson, R. (2006). New developments in exposure assessment: the impact on the practice of health risk assessment and epidemiological studies, *Environ Int.*, 32(8), 996-1009.
- Nuckols, J.R., Ward, M.H. and Jarup, L. (2004). Using geographic information systems for exposure assessment in environmental epidemiology studies, *Environ. Health Perspect.*, 112(9), 1007-15, doi:10.1289/ehp.6738.
- Oyana, T.J. and Rivers, P.A. (2005). Geographic variations of childhood asthma hospitalization and outpatient visits and proximity to ambient pollution sources at a US-Canada, border crossing, *Int. J. Health Geogr.*, 4(14), 11, doi:10.1186/1476-072X-4-14.
- Pandya, R.J., Solomon, G., Kinner, A. and Balmes, J.R. (2002). Diesel exhaust and asthma: hypotheses and molecular mechanisms of action, *Environ. Health Perspect.*, 110, 103-112.
- Revesz, P. (2002). *Introduction to Constraint Databases*, Springer, New York.
- Revesz, P. and Li, L. (2002). Constraint-Based Visualization of Spatial Interpolation Data, *Proc. of the Sixth International Conference on Information Visualization*, IEEE, London, England, 563-569.
- Revesz, P. and Wu, S. (2006). Spatiotemporal reasoning about epidemiological data, *Artificial Intelligence in Medicine*, 38(2), 157-170, doi:10.1016/j.artmed.2006.05.001.
- Schoenberg, I.J. (1946). Contributions to the problem of approximation of equidistant data by analytic functions, *Quarterly of Applied Mathematics*, 4, 45-99, 112-141.
- Shepard, D. (1968). *A two-dimensional interpolation function for irregularly spaced data*, 23rd National Conference ACM, 517-524.
- Thurston, G.D., and Bates, D.V. (2004). Ozone and asthma-Reply, *J. Am. Med. Assoc.*, 423-424.
- Triche, E.W., Gent, J.F., Holford, T.R., Belanger, K., Bracken, M.B., Beckett, W.S., Naeher, L., McSharry, J.E. and Leaderer, B.P. (2006). Low-level Ozone Exposure and Respiratory Symptoms in Infants, *Environ. Health Perspect.*, 114(6), 911-916, doi:10.1289/ehp.8559.
- Zienkiewicz, O.C. and Taylor, R.L. (2000). *Finite Element Method*, Vol. 1, The Basis, Butterworth Heinemann, London.