



Application of geostatistical approaches to predict the spatio-temporal distribution of summer ozone in Houston, Texas

Ryan Michael¹ · Cassandra R. O'Lenick¹ · Andrew Monaghan¹ · Olga Wilhelmi¹ · Christine Wiedinmyer² · Mary Hayden¹ · Mark Estes³

Received: 12 February 2018 / Revised: 9 September 2018 / Accepted: 17 September 2018 / Published online: 19 November 2018
© Springer Nature America, Inc. 2018

Abstract

Mitigation of adverse effects of air pollution requires understanding underlying exposures, such as ambient ozone concentrations. Geostatistical approaches were employed to analyze temporal trends and estimate spatial patterns of summertime ozone concentrations for Houston, Texas, based on hourly ozone observations obtained from the Texas Commission on Environmental Quality. We systematically assess the accuracy of several spatial interpolation methods, comparing inverse distance weighting, simple kriging, ordinary kriging, and universal kriging methods utilizing the hourly ozone observations and meteorological measurements from monitoring sites. Model uncertainty was assessed by leave-one-out cross-validation. Kriging methods performed better, showing greater consistency in the generated surfaces, fewer interpolation errors, and lower biases. Universal kriging did not significantly improve the interpolation results compared to ordinary kriging, and thus ordinary kriging was determined to be the optimal method, striking a balance between accuracy and simplicity. The resulting spatial patterns indicate that the more industrialized areas east and northeast of Houston exhibit the highest summertime ozone concentrations. Estimated daily maximum 8 h ozone concentration fields generated will be used to inform research on population health risks from exposure to surface ozone in Houston.

Keywords: Kriging · ozone · spatial interpolation · urban pollution

Introduction

Seasonally higher temperatures and increased sunlight in summer result in increases in emissions of biogenic and anthropogenic hydrocarbons, and facilitate reactions of those hydrocarbons with nitrogen oxides to form surface ozone [1]. Background levels of summertime surface ozone have increased over the last century due to increasing levels

of anthropogenically emitted nitrogen oxides [2]. Warming in the forthcoming century due to climate change may contribute to increases [3] in the intensity, frequency, and duration of daily maximum surface ozone concentrations, especially during the summer months [4, 5].

Elevated surface ozone concentrations are a concern because of the harmful effects on human health [1]. Short- and long-term exposures to elevated levels of ambient ozone have been associated with a variety of adverse health outcomes including respiratory [6], cardiovascular [7], and neurological conditions [8]. Furthermore, sensitivity to extreme ozone events varies within urban populations, with elderly and socioeconomically disadvantaged sub-populations being disproportionately affected [9]. Numerous studies report associations between ambient ozone levels and respiratory hospital admissions among the elderly [4].

An important input for modeling ozone-related health risks is accurate, spatially continuous surface ozone concentration data over the region of interest. However, such data are not readily available since ozone observational data are most often collected from monitoring stations with large and irregular spatial gradients. Spatial interpolation methods

Electronic supplementary material The online version of this article (<https://doi.org/10.1038/s41370-018-0091-4>) contains supplementary material, which is available to authorized users.

✉ Ryan Michael
ryanm@ucar.edu

¹ Research Applications Laboratory, National Center for Atmospheric Research, Boulder, CO, USA

² Atmospheric Chemistry Observations and Modeling Laboratory, National Center for Atmospheric Research, Boulder, CO, USA

³ Air Modeling and Data Analysis Section, Texas Commission on Environmental Quality, Austin, TX, USA

provide a means of generating spatially continuous data from these point observations [5, 10, 11]. Ozone is amenable to spatial interpolation methods due to its spatial distribution, correlation, and constant variance across well-defined geographic regions [10, 12].

The application of geostatistical methods to estimate spatial-temporal trends in ozone and other air pollutants is well supported in the literature, including spatial averaging [13, 14], nearest neighbor [15–19], inverse distance weighing [14, 18], and kriging [10–12, 21]. The variety of interpolation methods available have led to questions about relative accuracy and appropriate application for different scenarios. Previous studies have compared spatial interpolation methods, with emphasis on understanding the factors that affect model performance, such as sample density [21], data variation [22], sampling design [23], sources of errors in data, and factors affecting reliability [24, 25]. Results have shown that the performance of spatial interpolation methods depends on features of the method itself, as well as the data variation and sample density. However, there still exist uncertainties in selecting an appropriate method when large variabilities in sample frequencies and network densities exist.

The primary objective of this study is to systematically compare the performance of several spatial interpolation methods, and to identify an optimum method for the generation of ozone surfaces for metropolitan Houston, Texas, during summer. A review of the literature indicates that this systematic evaluation of geostatistical methods for generating ozone surfaces is the first of its kind for Houston, and is a key motivating factor for this work. These generated ozone concentration surfaces will be used as inputs to a health risk model in a follow-on study that more broadly examines impacts of ozone and extreme heat on elderly populations indoors and outdoors [26].

After presenting the methods, the results and discussion are divided into two major sections. First, we describe the observational data and discuss trends in ambient ozone concentrations. Second, we assess the spatio-temporal estimates of ozone concentrations for the Houston area generated by several geostatistical approaches. Specifically, we compare inverse distance weighing (IDW), simple kriging (SK), ordinary kriging (OK), and universal kriging utilizing varying combinations of temperature, relative humidity, and wind speed as covariates.

Methods

Study area

Our geographic domain is the Houston–Galveston–Brazoria (HGB) metropolitan area, with emphasis on the city of

Houston, Texas. Specifically, we defined an approximately 20,000 km² domain centered on the city of Houston (Fig. 1). Houston is the largest city in Texas, the fourth largest city in the United States, and the most ethnically diverse metropolitan area in the United States [27]. Along with its growth and diversity, come challenges such as an aging population, educational and income disparities, and poor air quality. A concern about ambient levels of air pollution in Houston has existed for decades and Harris County is known to be a “severe ozone non-attainment area” for the 1 h standard of the Clean Air Act [28, 29]. An extensive transportation network accounts for high emissions of nitrogen oxides (NO_x) and volatile organic compounds (VOCs) from mobile sources in the region [30]. Additionally, the presence of large amounts of vegetative and forested areas in the northeast of Houston allows for substantial contributions of biogenic VOCs [31]. Furthermore, the Houston Ship Channel is home to one of the largest concentrations of petrochemical industries in the United States, and represents a substantial source of NO_x and reactive VOCs in the region [32, 33]. This combination of emissions from anthropogenic NO_x sources and biogenic VOC emissions under favorable meteorological conditions, especially during summer months, can contribute to the formation of high O₃ concentrations in the study area [31, 34].

Description of observation data

Hourly observations of ozone during the summer season (1 June to 30 September) were obtained from the Texas Commission on Environmental Quality (TCEQ) monitoring network in the HGB metropolitan area for 1990–2016. TCEQ maintains an extensive network of Continuous Ambient Monitoring Stations (CAMS) that measure ambient ozone concentrations located on the perimeter, as well as the urban core of the Houston area. Figure 1 shows the geographic domain and the distribution of the monitoring sites. A total of 86 sites reported data. Of these, 61 sites were designated as regulatory sites by TCEQ, identified as meeting the requirements for assessing the federal ozone standards. The remainder were classified as “lite” or “non-regulatory” sites. Ozone monitors at the “lite” sites were not calibrated as often or as thoroughly as those at regulatory sites. Monitors at “non-regulatory” sites were well calibrated, but located on the tops of buildings or towers instead of at ground level. We used data from all monitoring sites to create ozone concentration interpolated surfaces and analyze spatio-temporal ozone trends.

Valid sample days were defined as those having more than 18 h of data. We calculated daily maximum 8 h ozone concentrations (MD8) by applying an 8 h moving window to the hourly time series and selecting the 8 h time window with the highest ozone concentration value during each 24 h

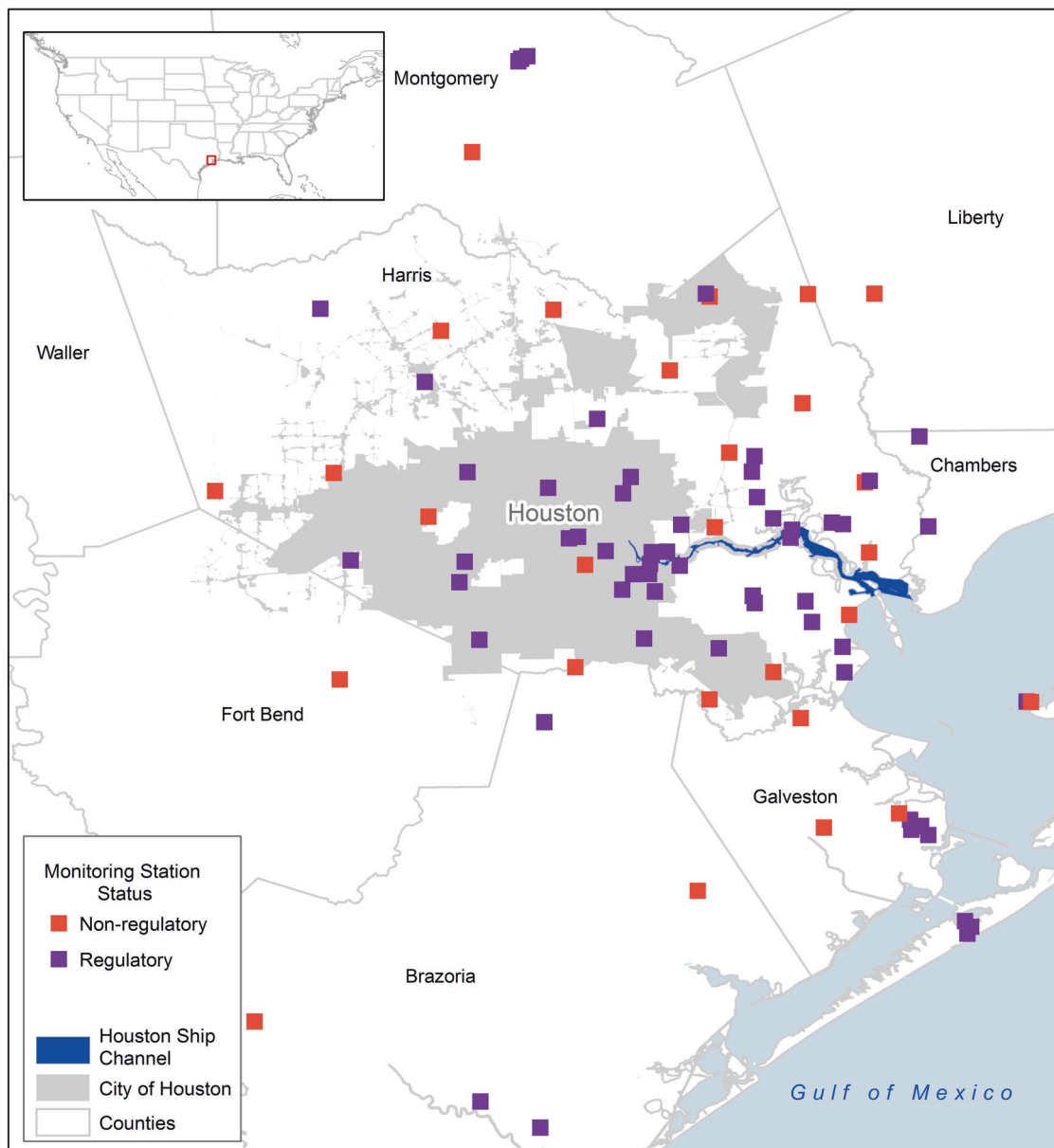


Fig. 1 Map of Houston–Galveston–Brazoria metropolitan area, with the City of Houston superimposed, showing the locations of ozone monitoring sites and the Houston Shipping Channel

period starting at local midnight. The 8 h windows were determined as missing if ≥ 3 missing hours occurred in the window. We applied MD8 as our summary statistic for assessing temporal trends and modeling the spatial distribution of summertime ozone. We utilized the entire temporal extent of the data (1990–2016) to elucidate ozone temporal trends. Temporal trends were computed by fitting linear regression lines through the annual (June–September) values of the 95th, 75th, 50th, 25th, and 5th percentiles of MD8. The trend was considered statistically significant if $p < 0.05$ according to Student's t -test. For the analysis of the spatial distribution of MD8, we restricted the years to 2000–

2016 which corresponds to the period for which we have health surveillance data for the follow-on study that will utilize these estimates.

Interpolation methods

There are several well-developed interpolation techniques for modeling spatial data. These include deterministic methods such as triangulation, local polynomial interpolations, trend surface analysis, splines, IDW, and geostatistical methods such as kriging and its many iterations [35–37]. Triangulations produce a continuously

differentiable surface but give no measure of prediction accuracy, while local polynomial interpolations and trend surfaces do not model account for fine-scale variations, and thus are not applicable when local prediction accuracy is important [10, 38]. Kriging is a best linear unbiased predictor of a spatial variable that produces a set of predictions that minimizes the error variance. It accounts for clustering, is an exact estimator, and produces error estimates [11, 39]. It must be highlighted that the variability in kriging estimates will be less than the variability in the true spatial process due to the “smoothing” nature of the method, and its results depends entirely on the representative sampling data for the region of interest. A non-uniform or sparse network may limit the accuracy of the resulting interpolated surface due to insufficient sampling of the extreme sub-regions of concentrations in the spatial domain. Consequently, kriging may not be able to resolve small scale spatial trends, such as titration of ozone near NOx sources.

IDW produces estimates that are simply weighted averages of the nearby data points, where the averaging is based on some criteria. Previous studies indicate that with careful consideration to the choice of parameter values, IDW can provide estimates with nearly the same prediction accuracy as kriging [10, 39, 40].

Here, we investigate IDW and kriging. We choose to evaluate kriging because it provides a solution to the problem of estimation of a surface by taking spatial correlation into account. The deterministic IDW was chosen for comparison due to the simplicity of its formulation and the fact that it combines the idea of estimation based on proximity, and the gradual change of a trend surface. Both of these methods are weighted average methods with the same basic mathematical formulation. Essentially, we seek to compute ozone concentration, z , at an unsampled location, x_0 , given a set of neighboring values sampled at locations denoted by x_i . The interpolating relationship is given by: [36, 41]

$$z(x_0) = \sum_{i=1}^n \lambda_i \cdot z(x_i) \dots \text{where } \sum_{i=1}^n \lambda_i = 1,$$

where λ_i represents the weights assigned to each of the neighboring values, and the sum of the weights is one. Interpolation involves defining the search area around the point to be predicted, locating the observed data points within the neighborhood, and assigning appropriate weights to each observed data point.

In IDW, interpolation weights are computed as a function of the distance between the observation locations and the predicted/unknown locations. An observed value closer to the unknown location of interest is assigned a heavier weight. IDW assumes that each measured point has a local influence that diminishes with distance, and is characterized

by the following formulation: [39]

$$z(x_0) = \frac{\sum_{i=1}^n w(d_i) \cdot z(x_i)}{\sum_{i=1}^n w(d_i)},$$

where $z(x_0)$ and $z(x_i)$ represent the predicted and observed values respectively, n is the number of measured sample points used in the prediction, $w(d_i)$ is the weighting function, and d_i is the distance from x_0 to x_i . Here, the weight is assigned as the inverse of the distance raised to a mathematical power. This power parameter facilitates the control of the significance of known points on the interpolated values based on their distance from the output point. A higher power value places more emphasis on the nearest points. Thus, nearby data will have the most influence, and the surface will have more detail (be less smooth). Specifying a lower value for power has been shown to result in undue influence being assigned to surrounding points that are farther away, resulting in a smoother surface. Since the IDW formula is not linked to any real physical process, there is no way to determine that a particular power value is too large. A default value of 2 is typically used, however, and we conducted sensitivity testing on power values ranging from 0.5 to 3, and considered the value with the minimum mean absolute error as optimal.

Kriging is a stochastic technique similar to IDW, in that it uses a linear combination of weights at known points to estimate the value at an unknown point; however, in contrast to the deterministic IDW, kriging takes into account the spatial correlation between measurement points in providing a solution. The spatial correlation between the measurement points is quantified by means of a variogram function: [39, 42]

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i + h)]^2,$$

where γ_h is the estimated semivariance at a separation distance, h , and $z(x_i)$ and $z(x_i + h)$ are the observed values at x_i and $x_i + h$ separated by h . $N(h)$ is the number of pairs of measurement points with distance h apart. The variogram is used to compute weights, λ_i , which minimize the variance in the estimated value. The semivariance can be a function of both distance and direction, and most often increases as h increases, indicating that points close together tend to be more similar than those far apart. A parametric function is used to model the semivariance for different values of h . Although the spherical model is most widely used, we also explored Gaussian, exponential, and Matern models. Once the model variogram is fit to the empirical data, it is used to compute the weights, λ_i , such that the estimation variance is less than the variance for any other linear combination of the observed values [41, 43].

We explored simple kriging, ordinary kriging, and universal kriging, utilizing observed meteorological variables (temperature, relative humidity, and wind speed) from monitors co-located at the ozone monitoring sites to improve estimates. When spatial correlation between a covariate and the variable of interest is high, universal kriging has been shown to give better results for the estimates than ordinary kriging [39]. Additionally, high ozone pollution episodes have been shown to be correlated with high temperatures, low wind speeds, clear skies, and stagnant weather [44–46]. Simple kriging assumes that the mean value is known, while ordinary kriging assumes that the mean is unknown, focuses on the spatial component, and only uses samples in the local neighborhood for the estimate. Universal kriging explores non-stationary variation by assuming a trend in average values across the domain [39, 43]. We applied each interpolation method to generate daily MD8 ozone concentration surfaces at 1 km × 1 km spatial resolution for the 20,000 km² (100 km × 200 km) domain.

Assessment of interpolation methods

We assessed the spatial interpolation methods in two ways. First, we plotted the spatial MD8 patterns generated by each method for a randomly sampled summer case day in order to provide a visual depiction of the patterns and differences among methods, and to assess predictions of MD8 quantiles. Second, we computed numerous model fit statistics over a 5-year period in order to robustly assess and compare the methods with a large set of independent MD8 observations.

We randomly selected summer 4 August 2010, as our case day, and used the MD8 ozone concentration as our test statistic to evaluate initial model parameters. We estimated an empirical variogram by comparing both the classical and Cressie robust estimators for binned and un-binned distances, and settled on a binned variogram with a maximum distance restricted to 100 km [47, 48]. Next, we estimated the parameters of several candidate parametric variograms, comparing among exponential, Matern, and Gaussian covariance models, and between ordinary least squares and weighted least square estimation procedures for each method. The parameters from the fitted variogram model were then used to implement and assess the kriging methods.

We selected 2012–2016 (June–September) for our 5-year model fit assessment, using leave-out-one cross-validation to evaluate the performance of each interpolation method. We used the period 2012–2016 since this period coincided with the period of highest monitor density in the observation network. This was achieved by taking each observation in turn out of the sample dataset and estimating it from the remaining observations. A total of 6501 ozone

concentration surfaces were generated from 591 days and 8 interpolation techniques. This process allowed us to estimate mean error (ME) and the root mean squared error (RMSE) test statistics for each interpolation. The ME was used to detect bias, and should ideally be zero if the predictions are centered on the measurement values. The RMSE was used to compare the ability of the interpolation methods to predict the measured values. A smaller RMSE suggests better model performance. We also calculated the 95% prediction interval coverage probability (Cov95) and the mean prediction standard deviation (AveSE) as metrics for evaluating model performance. The validated model was applied to produce spatial estimates of MD8 ozone concentrations for the Houston area; these estimates will inform our efforts to understand population health risks from extreme ozone episodes. Spatial interpolation methods were performed using the geoR (Version 1.7.5.2) [52] package on the R (version 3.4.1) [50].

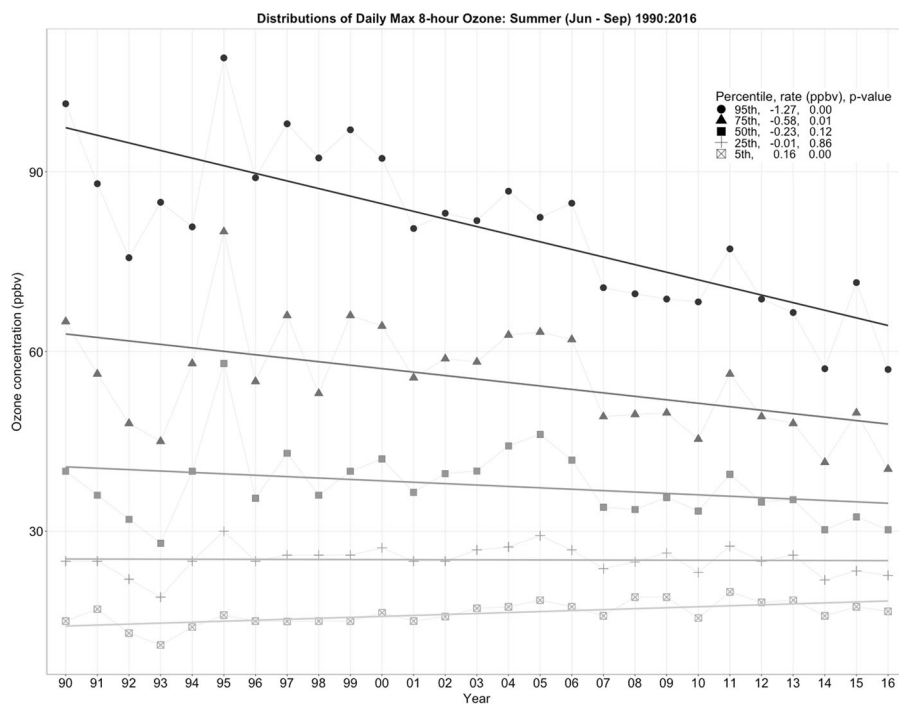
Results and discussion

Trends in ozone observations

We examined the trend in ozone observed by 62 active monitoring sites for the summer months from 1990 to 2016 (Fig. 1S in the Supplementary Materials). In the first decade of the interval, an average of 13 sites were active per year. This number increased to 35 sites in the second decade, and 45 in the final 6 years. Considering the trend across the entire interval, reporting from active sites was generally less than 50% prior to 2004 and increased substantially thereafter (greater than 60%). The fraction of valid station days observed was consistently high across all years, averaging greater than 85% over the period.

We also examined the observed MD8 ozone concentrations for June–September 1990–2016, emphasizing station days when the MD8 ozone exceeded the regulatory standard of 70 ppb (National Ambient Air Quality Standards (2015 NAAQS) as defined in the US Code of Federal Regulations (80 FR 65292)) (Fig. 2S, Supplementary Material). A greater proportion of station days exceed the 8 h ozone standard in the earlier years of the period, and also a greater number of exceptionally high ozone station days with values exceeding 120 ppbv, classified as severe non-attainment for the 8 h ozone standard. One-fifth of summer station days in 1999 and 2000 exceeded the 8 h ozone standard. This trend decreased over the summers of subsequent years, ranging from an average of 14% of summer station days exceeding the standard during the 2001–2006 interval, to an average of 4% of days for the remaining years (2007–2016). The number of occurrences of exceptionally high ozone station days also displayed a decreasing trend,

Fig. 2 Temporal trends over the period 1990–2016 for the 95th, 75th, 50th, 25th, and 5th percentiles of summer MD8 ozone concentrations, respectively. The solid line associated with each percentile gives the trend derived by linear regression, and the legend shows the trend rate and p value



especially in the last 8 years of the period, with relatively few station days exceeding the standard compared with the preceding interval. The observed trends in MD8 ozone concentrations exceedances are largely attributed to changes in the ozone standard over the period. This includes the change in 1997 from a 1 h, 120 ppbv ozone US NAAQS to an 8 h, 80 ppbv ozone standard (NAAQS). This standard was further revised in 2015 (NAAQS) from 80 ppbv to 70 ppbv by the Environmental Protection Agency (EPA) [51, 52].

Figure 2 gives the trend in the 95th, 75th, 50th, 25th, and 5th percentile distributions of MD8 ozone, respectively, over the interval. The 95th and 75th percentile distributions demonstrated a decreasing trend that was significant at the 0.05 level, at a rate of -1.3 and -0.6 ppbv/year, respectively. The median ozone rate, while not significant, also demonstrated a decrease (-0.2 ppbv/yea) over the period. At the lower extreme, the 5th percentile distribution showed an increasing trend over the interval that was significant at 0.2 ppbv/year. Overall, MD8 ozone concentrations in the study area demonstrated a decreasing trend.

The temporal characteristics in MD8 summer ozone concentrations are presented in Fig. 3. Summer ozone concentrations, both the extreme (Fig. 3a) and average (Fig. 3b), displayed an increasing trend in the first decade of the period, peaked in 1995, and then gradually decreased over the remainder of the interval. This trend was consistent across the majority of the monitoring sites with some spatial variation. We do note two inflection points in the later part of the interval at 2011 and 2015.

The monthly averages of MD8 ozone concentrations are shown in Fig. 3c exhibit interesting temporal variability within the summer season. We observe a decrease in the mean MD8 ozone concentrations from June to July, before increasing for the remainder of the summer months. This mid-summer decrease is attributed to meteorological phenomenon called the Bermuda High, a quasi-permanent high pressure system that influences summertime weather over the eastern and southern United States [53, 54]. The system extends further west in mid-summer than during other times of the year and brings clean maritime air over the eastern half of Texas, usually carried by relatively brisk winds. The result of this influx of clean air and associated winds is a decrease in ozone concentrations along the path and the resultant mid-summer inflection point in July, as demonstrated here.

Figure 3d shows the mean summer diurnal cycle in ozone concentrations observed at all stations over the period. Daily summer ozone across Houston area demonstrated the typical mono-modal pattern indicative of tropospheric ozone chemistry. Ozone concentrations were lowest between 04:00 and 06:00 h local time, and increased through the day to peak between 12:00 and 17:00 h. There is substantial spatial variation in hour of daily max ozone, with some stations peaking as much as 4 h later than others. To investigate this spatial variability, we placed a 40 km resolution grid (representing different “zones”) on the domain, centered on Houston (Fig. 4a). We then plotted the diurnal cycles of all monitors within each 40 km zone, color coded according to the zone (Fig. 4b). The results indicate

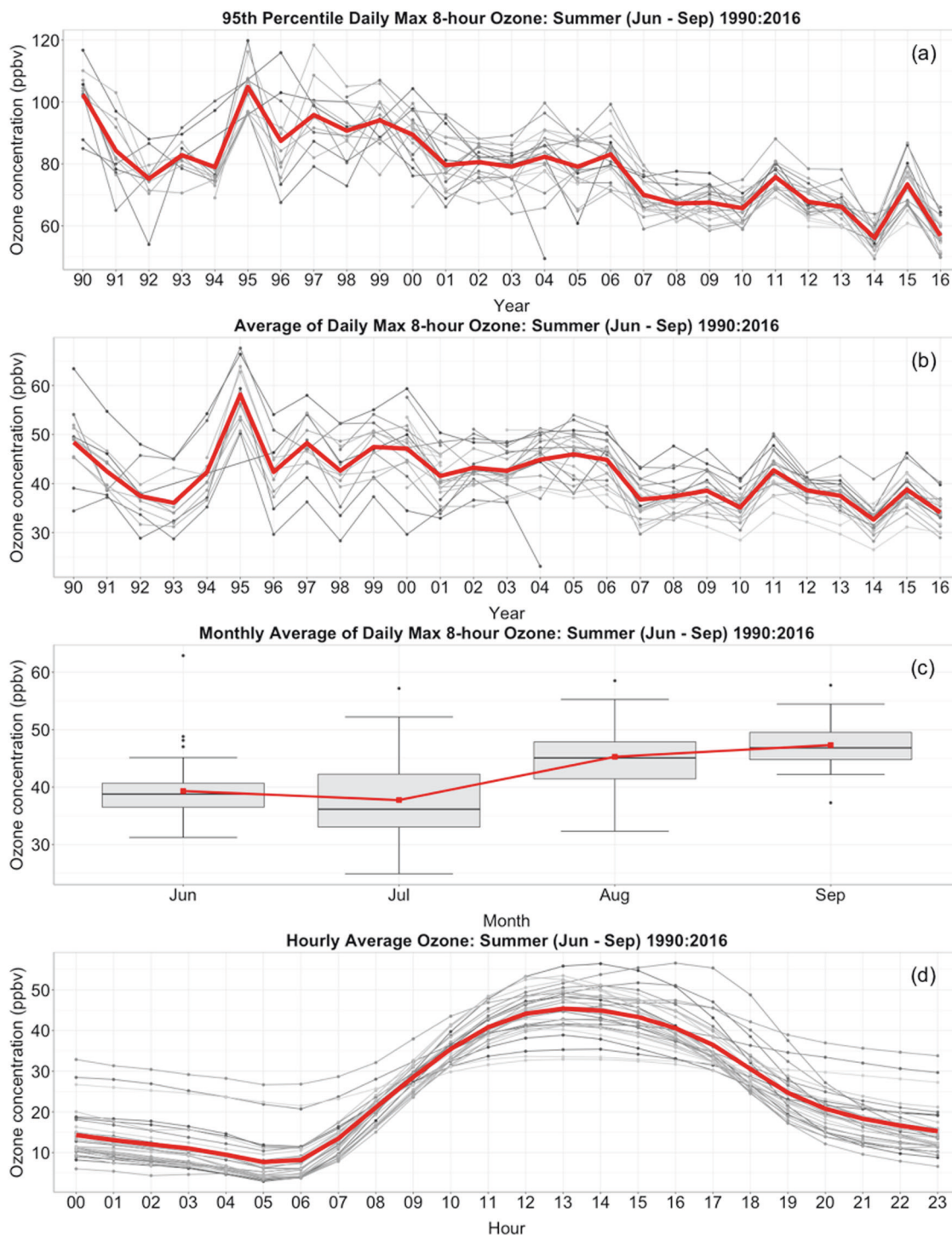


Fig. 3 Temporal variation in max daily 8 h (MD8) ozone concentration for summers (June–September) over the period 1990–2016 from all reporting monitors. Each dot (line plots) represents the summary statistic calculated for each monitor for the representative year. The weighted line represents the moving average across all monitors. Subplot (a) displays the 95th percentile of MD8 ozone concentrations,

and subplot (b) gives the average of the MD8 ozone. The boxplot (subplot c) gives the variation in monthly averaged ozone across monitors in the domain and the red dots/lines show the average across stations by month. Subplot (d) gives the average diurnal cycle of ozone over the interval

that ozone peaks in the southeast of the domain (near the Houston Ship Channel) earlier in the cycle, and at lower concentrations, then migrates across the domain in a SE to

NW direction, peaking further inland at locations increasingly distant from the industrial area with each successive hour. This observed spatio-temporal trend highlights the

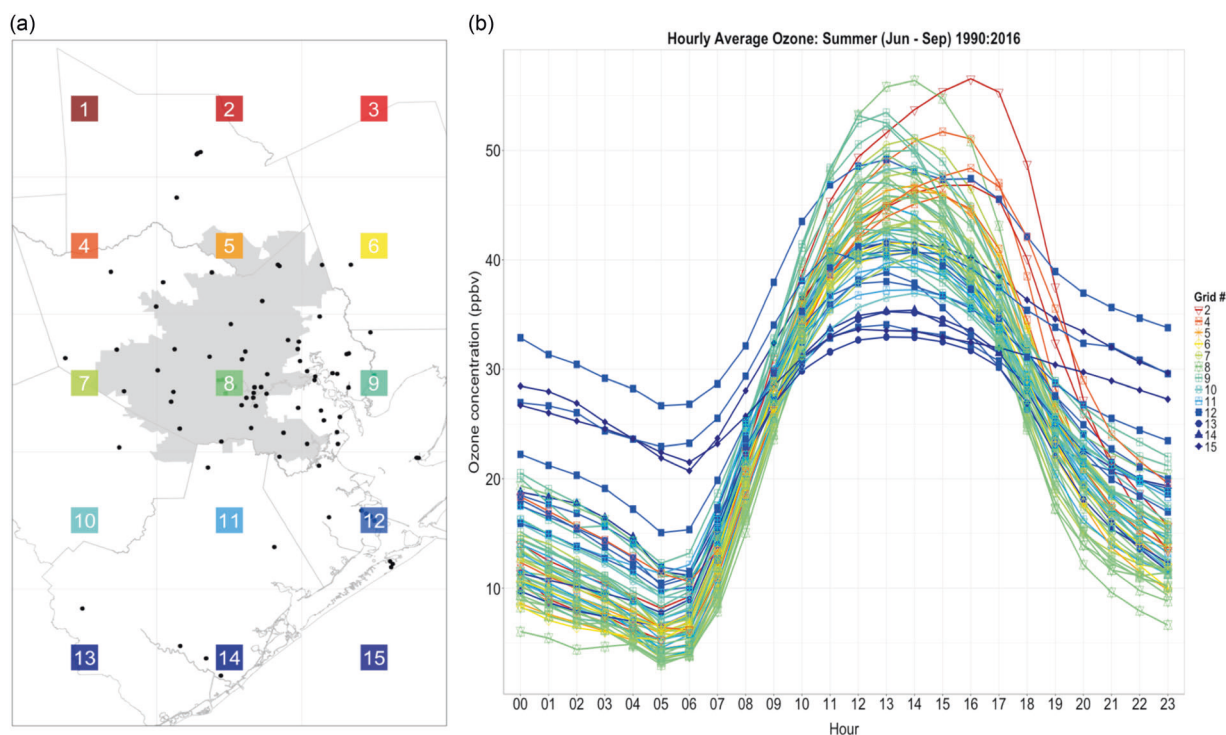


Fig. 4 **a** The 40 km resolution grid placed on the domain. The numbers/colors identify the grid each monitor occupies. The black dots identify the monitors. **b** Diurnal cycles of hourly averaged ozone. Each

line identifies a monitor. The colors/shapes depict the grid each monitor occupies

role of industrial emissions as the primary cause of the highest ozone, and is consistent with studies done in the Houston area [34, 51]. For example, TCEQ identified the highest ozone (>125 ppbv) concentrations in the HGB area as resulting from rapid and efficient ozone formation plumes, originating from highly reactive volatile organic compounds and nitrogen oxides co-emitted from petrochemical facilities, and identified the Houston Ship Channel (HSC) as the origin of the plumes with the highest ozone concentrations [32]. Dispersion of ozone plumes is aided by a prominent sea breeze driven by land–sea contrasts along the coasts of the Gulf of Mexico and Galveston Bay which cause air to be drawn during the day from Galveston Bay northward into Houston. The resultant effect is the transport of ozone and ozone precursors away from the heavily industrialized area of the HSC into more populated areas of Houston, and the presence of transient high ozone events at the observation sites [29, 55–59].

Comparison of interpolation methods

Figure 5 shows the spatial variability in MD8 ozone concentrations observed for the randomly sampled case day, 4 August 2010. Forty-one sites in the 20,000 km² domain reported observations for this day. MD8 ozone concentration varied from 23.0 to 77.1 ppbv across monitoring stations, with a mean of 45 ppbv and a median value of 41

ppbv. MD8 ozone observations were substantially higher in the north-eastern part of the domain, a predominantly industrial region of the Houston area; the highest concentrations occur northeast of the HSC.

After comparing several candidate variogram models, we applied a Gaussian model, with a weighted least squares estimation procedure, and Cressie inverse-variance weights. We selected the Gaussian covariance function because it outperformed the other methods when comparing weighted sum-of-square error. There were no significant performance gains when comparing between ordinary least squares and weighted least squares estimation procedures, and the nugget estimates were consistent across covariance functions. Estimated parameters of the final semi-variogram included a nugget of 25.5, a marginal variance of 262, and leveled off to the sill at 64 km.

Table 1 compares the summary statistics of MD8 ozone concentrations observed at each monitor location on 4 August 2010, with the values predicted by the interpolation method assessed here. Unsurprisingly, IDW reproduced the distribution of the data well due to its deterministic nature. Simple kriging underestimated the MD8 ozone concentrations at both the minimum and the maximum, but reproduced the median, 25th and 75th percentiles well. Ordinary kriging overestimated the minimum and 25th percentile MD8 ozone concentration but underestimated the maximum. Universal kriging performed similarly to ordinary

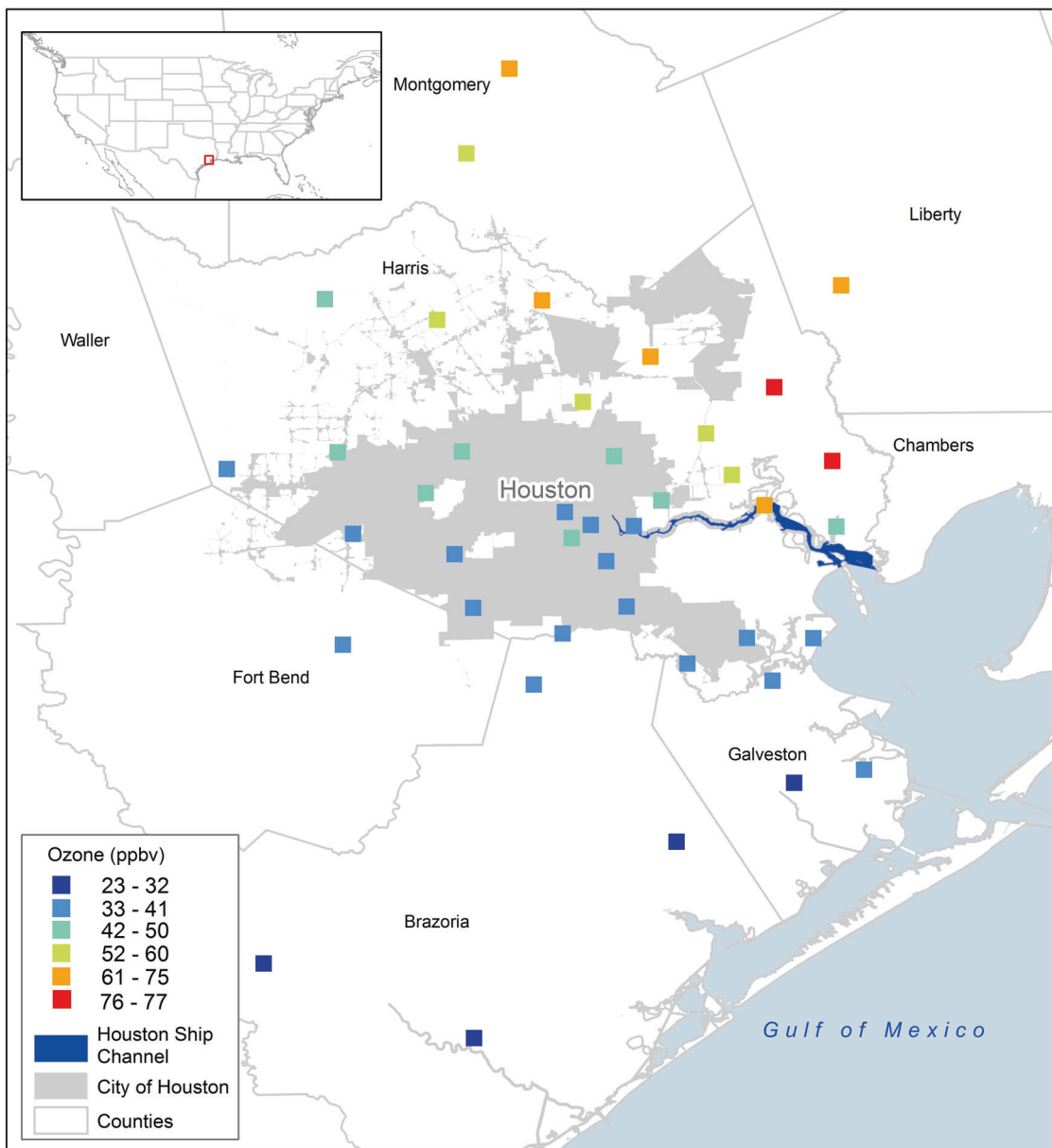


Fig. 5 Spatial variability in MD8 ozone concentration for 4 August 2010 from all reporting monitors. Each dot represents the value of the summary statistic calculated at each monitor

Table 1 Comparison of summary statistics of the measured MD8 ozone concentrations for 4 August 2010 to the interpolated MD8 ozone concentrations at each monitor in ppbv/year, utilizing inverse distance weighing (IDW), simple kriging (SK), ordinary kriging (OK), and universal kriging with daily maximum relative humidity (maxRH), maximum temperature (maxT), and mean wind speed (meanWS) as covariates, respectively

	MD8 Ozone	IDW	SK	OK	UKmaxRH	UKmaxT	UKmeanWS
Min	23.00	23.00	22.36	24.11	23.65	23.59	24.04
25th%	34.38	34.61	34.83	36.13	36.34	36.41	36.31
Median	40.62	40.62	40.10	40.00	39.93	39.87	40.00
Mean	45.22	45.22	44.70	45.21	45.57	45.57	45.53
75th%	56.50	56.49	56.48	56.32	57.31	57.19	57.40
Max	77.12	77.01	74.25	73.55	74.96	74.33	74.68

kriging, overestimating the lower extremes, reproducing the median and 75th quartile, and underestimating the maximum. We also observe a slight but consistent increase in the range of estimates for the universal kriging methods (decrease in the minimum and increase in the maximum) as compared to the ordinary kriging estimates.

Table 2 compares the summary statistics of the spatial prediction standard errors across the predicted surface for the kriging models for MD8 ozone concentrations. We observe similar distributions of standard errors for surfaces estimated with the simple and ordinary kriging methods. Larger differences in the distribution of standard errors were

Table 2 Comparison of summary statistics of spatial prediction standard errors across the predicted surface for 4 August 2010, utilizing simple kriging (SK), ordinary kriging (OK), and universal kriging with daily maximum relative humidity (maxRH), maximum temperature (maxT), and mean wind speed (meanWS), respectively

	SK	OK	UKmaxRH	UKmaxT	UKmeanWS
Min	1.77	1.77	1.83	1.85	1.83
25th%	3.01	3.01	3.25	3.32	3.18
Median	4.63	4.64	4.87	4.92	4.81
Mean	5.30	5.34	5.88	5.86	5.75
75th%	7.06	7.10	8.03	8.06	7.84
Max	14.51	15.20	19.60	16.13	16.63

observed for the universal methods when compared with the simple and ordinary kriging methods, with increases observed in all categories of the summaries. This suggests that while the quality of the fits provided by the two models comparable, there is no significant value gained but the inclusion of additional covariates, and thus the simple model, ordinary kriging can be used to interpolate the spatial region with adequate results.

Figure 6 gives the predicted MD8 ozone concentration surfaces for 4 August 2010 for each interpolation method on a regular grid of 1 km by 1 km resolution across the domain. Based on visual inspection, IDW appears to have the poorest performance of the interpolation methods (which is confirmed in the statistical validation presented below). It is evident that the weight assigned to points was influenced by neighboring points when they were more clustered. Additionally, isolated points were allowed to exert undue influence in all directions, thus resulting in the characteristic bull's eye pattern seen in surfaces generated using this method. Since IDW is an exact interpolator, it reproduced the minimum and maximum values in the observations, but high variability in the observations resulted in a rougher surface produced.

The surfaces generated by kriging appear to provide a more realistic representation of the spatial variation in ozone concentrations in the domain, based on previous studies

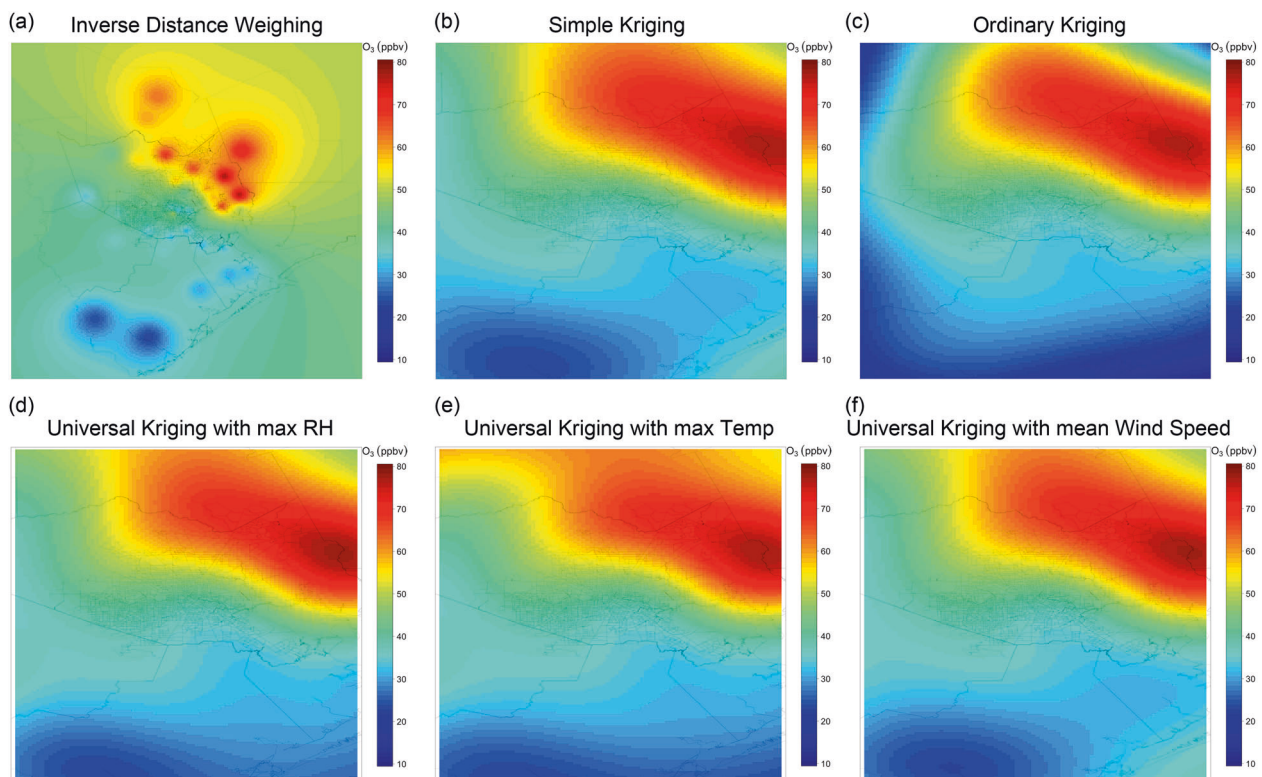


Fig. 6 Gridded predictions of MD8 ozone concentrations for 4 August 2010 using inverse distance weighing (IDW, **a**), simple kriging (SK,

b), ordinary kriging (OK, **c**), and universal kriging with max relative humidity (**d**), max temperature (**e**), and mean wind speed (**f**)

indicating “smooth“ ozone spatial variability [37, 41, 59]. We observe an ozone concentration plume in the north-eastern quadrant of the domain that is reflective of the high observation values recorded at monitors located there. Compared to the surface generated by simple kriging, the ordinary kriging exhibited lower prediction error overall. The differences in prediction error were higher in areas of the domain where the monitoring network was sparse, as well as in domain areas with large variations between nearby observations. Simple kriging did a poor job of reproducing the values at the lower extreme of the observed concentration range, while ordinary kriging was able to generate values representative of both extremes. Universal kriging with all covariates did not exhibit any substantial

improvements in the interpolated surfaces over those gained by ordinary kriging, but performed better than simple kriging, showing similar trends in the predicted surfaces, and good reproduction of both the maximum and minimum observations.

Figures 7a–d further examines the contrast between ordinary kriging and universal and simple kriging spatial predictions. The figures were derived by subtracting the universal and simple kriging estimate from the ordinary kriging estimate at each predicted location. Contours representing quantiles of the differences between predicted model estimates were used to understand spatial agreement between model estimates. The range of differences between simple and ordinary kriging estimates are relatively large;

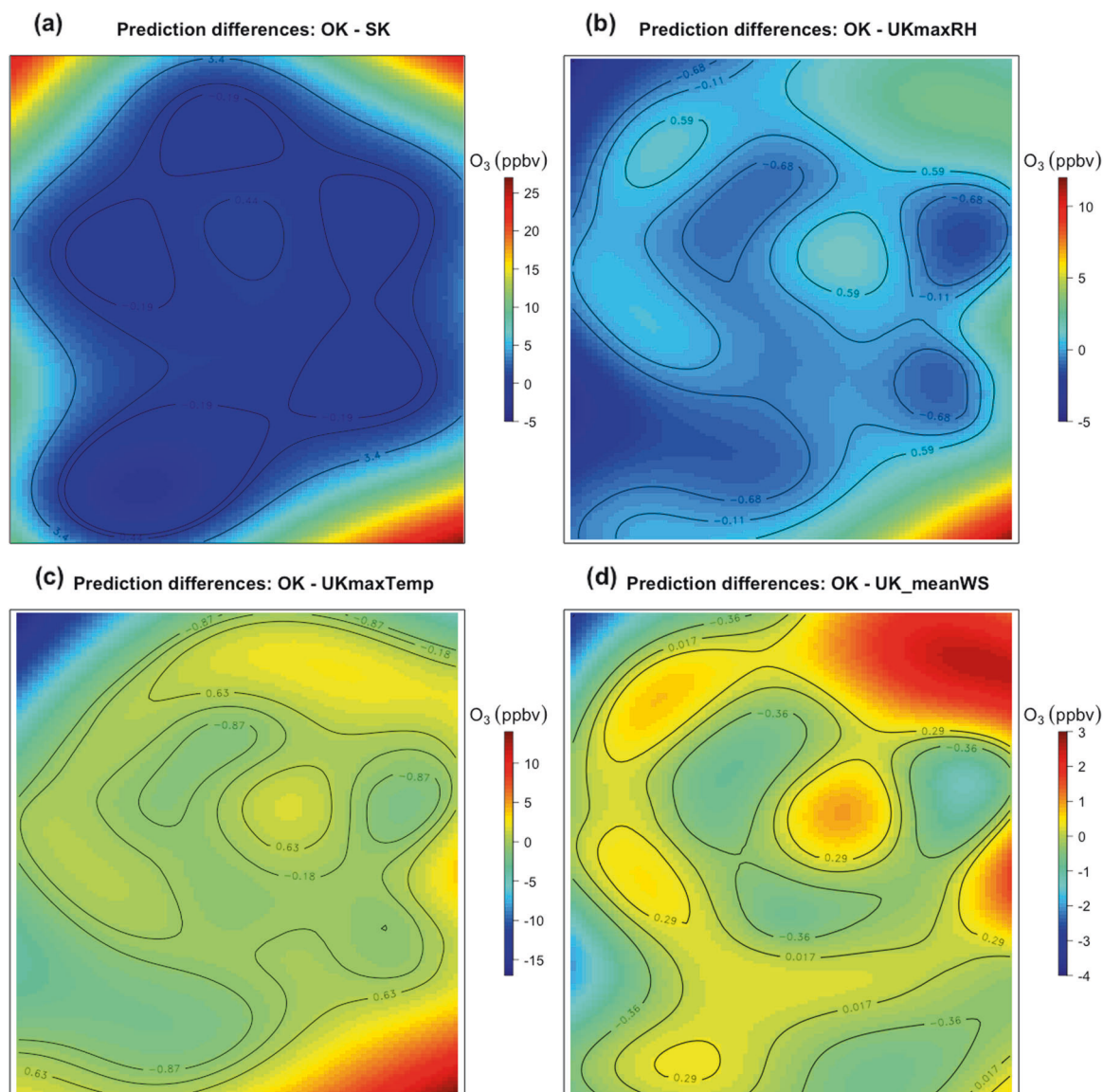


Fig. 7 Gridded predictions differences of MD8 ozone concentrations for 4 August 2010 between ordinary kriging and simple kriging (SK, **a**), universal kriging with max relative humidity (**b**), max temperature

(**c**), and mean wind speed (**d**). The contour lines give the quantiles of differences of the kriging estimate from the ordinary kriging

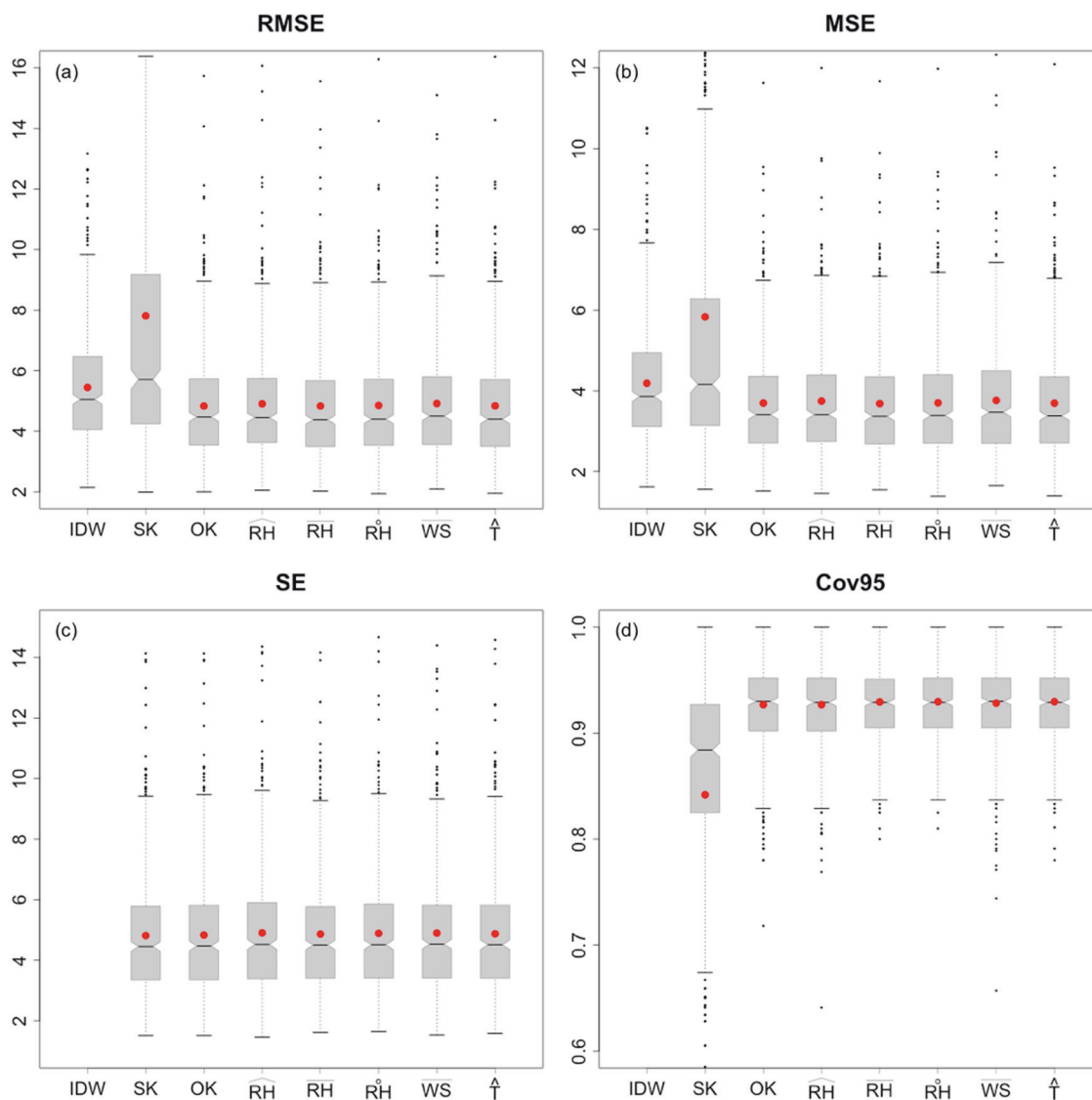


Fig. 8 (a) Root mean squared error (RMSE, ppbv), (b) mean absolute error (MSE, ppbv), (c) mean prediction error (SE, ppbv), and (d) 95% coverage interval (Cov95) calculated from leave-out-one cross-validation of MD8 ozone concentration interpolation methods performed on five years (2012 - 2016) of summer ozone observations in

the domain. Methods assessed here are inverse distance weighted (IDW), simple kriging (SK), ordinary kriging (OK), and universal kriging with daily maximum relative humidity (\overline{RH}), mean relative humidity (\overline{RH}), minimum relative humidity (RH), mean wind speed (\overline{WS}), and maximum temperature (\overline{T}) as covariates, respectively

however, greater than 75% of the predicted surface display good agreement. In comparison, the universal method estimates demonstrated better agreement with the ordinary kriging estimate as evidenced by the narrower range of prediction differences and greater coverage (>85% across all universal kriging methods). While the low and high regions tend to be clustered, the midrange of the differences was evenly distributed, suggesting that the universal kriging estimates did not detect any important trend features missed by the ordinary kriging model.

Finally, Fig. 8a–d give statistical metrics calculated from the leave-out-one cross-validation of the interpolation

methods. Since kriging explicitly accounts for spatial variance, in contrast to IDW, it tends to give lower RMSE and ME values, as is evident in the results observed here. Simple kriging was consistently the poorest interpolation method, displaying high interpolation errors and greater bias. Overall, ordinary kriging and universal kriging were the better performing methods, displaying lower RMSE and MSE, indicating that the methods were substantially unbiased. There was little difference in the statistical metrics between ordinary kriging and the universal kriging methods, indicating that no obvious increases in performance are achieved by including additional covariates via universal kriging.

Previous model inter-comparison studies have assessed the ability of spatial interpolation methods to estimate ozone concentrations at subject exposure points in Houston, Texas, with emphasis on IDW, kriging in space, and kriging in space and time [60], and ordinary kriging [61, 62]. Gorai et al. [63] explored the influence of local climatic factors on the spatial distribution of ground level ozone concentrations, investigating the role of temperature, wind speed, wind direction, and NO₂ level ozone concentrations over Eastern Texas. Higher concentrations of NO₂ were associated with higher concentrations of ozone, and while the distribution patterns of ozone were influenced by wind speed and direction, no significant correlation was found with the temperature profile of the domain. Studies have shown that the scale of the domain may affect the contributions of climate variable to affect the spatial model [64].

Conclusion

We analyzed 27 years (1990–2016) of summer ozone observations from the TCEQ monitoring network in the HGB metropolitan area to understand spatial and temporal trends. We also explored spatial interpolation methods for generating representative concentration surfaces, and provided a systematic comparison between different interpolation methods to identify the optimal method for the generation of ozone surfaces for metropolitan Houston, Texas. This approach is generalizable and provides information on methodological uncertainty by evaluating multiple methods utilizing networks with varied spatial coverage and sampling frequencies. This approach can be extended by incorporating advanced methods into the comparison scheme, such as emission-based air quality modeling, and regression methods, and the inclusion of multiple pollutants.

The temporal trend in summer ozone concentrations in the study area indicated greater concentrations in the first decade of observation in both the extreme and the mean, before decreasing over the remainder of the period. The 95th and 75th percentile distributions of MD8 ozone demonstrated a statistically significant decreasing trend that was significant over the period. Summer ozone also exhibited a spatio-temporal trend of lower peaks earlier in the diurnal cycle in the southeastern region of the domain, and greater concentration peaks later in the cycle predominantly in the north-north western region. This pattern is facilitated by the emissions of ozone precursors from the heavily industrialized zone of the HSC, and the presence of a prominent sea breeze pushing ozone plumes north.

Evaluation of the spatial interpolation methods indicated that when compared with the deterministic IDW in this study, kriging methods performed better, showing greater

consistency in the generated surfaces, and lower errors and bias. Ordinary kriging was determined to be the optimal kriging method, striking a good balance between accuracy and simplicity. The inclusion of additional covariates did not significantly improve the interpolation results. The surfaces generated here contributed to better understanding of spatial and temporal variability of ozone over a large urban area. Estimated daily maximum 8 h ozone concentration fields from the ordinary kriging model will inform our research on population health risks associated with extreme ozone episodes, and will be applied to assess exposures for empirical and predictive health risk models.

Acknowledgements This research was supported in part by Assistance Agreement No. 83575401 awarded by the US Environmental Protection Agency. It has not been formally reviewed by the EPA. The views expressed in this document are solely those of the authors and do not necessarily reflect those of the Agency.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Knowlton K, Rosenthal JE, Hogrefe C, B. L, Gaffin S, Goldberg R, et al. Assessing ozone-related health impacts under a changing climate. *Environ Health Perspect.* 2004;112:1557–63.
2. Logan JA. Tropospheric ozone: seasonal behavior, trends, and anthropogenic influence. *J Geophys Res.* 1985;90(D6): 10463–82.
3. Intergovernmental Panel on Climate Change (IPCC). Managing the risks of extreme events and disasters to advance climate change adaptation. Cambridge: Cambridge University Press; 2012. p. 582.
4. Stafoggia M, Forastiere F, Faustini A, Biggeri A, Bisanti L, Cadum E, et al. Susceptibility factors to ozone-related mortality: a population-based case-crossover analysis. *Am J Respir Crit Care Med.* 2010;182:376–84.
5. Adam-Poupart A, Brand A, Fournier M, Jerrett M, Smargiassi A. Spatiotemporal modeling of ozone levels in Quebec (Canada): a comparison of kriging, land-use regression (LUR), and combined bayesian maximum entropy-LUR approaches. *Environ Health Perspect.* 2014;122:970–6.
6. Bell ML, Dominici F, Samet JM. A meta-analysis of time-series studies of ozone and mortality with comparison to the National Morbidity, Mortality, and Air Pollution Study. *Epidemiology.* 2005;16:436–45.
7. Ruidavets J-B. Ozone air pollution is associated with acute myocardial infarction. *Circulation.* 2005;111:563–9.
8. Genc S, Zadeoglulari Z, Fuss SH, Genc K. The adverse effects of air pollution on the nervous system. *J Toxicol.* 2012;2012:1–23.
9. Conlon Kathryn, Monaghan Andrew, Hayden Mary, Wilhelmi O. Correction: Potential impacts of future warming and land use changes on intra-urban heat exposure in Houston, Texas. *PLoS ONE.* 2016;11:e0151226–4..
10. Liu LJS, Rossini AJ. Use of kriging models to predict 12-hour mean ozone concentrations in metropolitan Toronto - a pilot study. *Environ Int.* 1996;22:677–92.

11. Knotters M, Brus DJ, Oude Voshaar JH. A comparison of kriging, co-kriging and kriging combined with regression for spatial interpolation of horizon depth with censored observations. *Geoderma*. 1995;67:227–46.
12. Lefohn AS, Simpson J, Knudsen HP, Bhumralkar C, Logan JA. An evaluation of the kriging method to predict 7-h seasonal mean ozone concentrations for estimating crop losses. *J Air Pollut Control Assoc*. 1987;37:595–602.
13. Chestnut LG, Schwartz J, Savitz DA, Bruchfiel CM. Pulmonary function and ambient particulate matter: epidemiological evidence from NHANES I. *Arch Environ Heal*. 1991;46:135–44.
14. Kinney PL, Aggarwal M, Nikiforov SV, Nadas A. Methods development for epidemiologic investigations of the health effects of prolonged ozone exposure. Part III. An approach to retrospective estimation of lifetime ozone exposure using a questionnaire and ambient monitoring data (U.S. sites). *Res Rep Health Eff Inst*. 1998;81:79–121.
15. Tashkin DP, Clark VA, Simmons M, Reems C, Coulson AH, Bourque LB, et al. The UCLA population studies of chronic obstructive respiratory disease. VII. Relationship between parental smoking and children's lung function. *Am Rev Respir Dis*. 1984;129:891.
16. Schwartz J, Zeger S. Passive smoking, air pollution, and acute respiratory symptoms in a diary study of student nurses. *Am Rev Respir Dis*. 1990;141:62–7.
17. Stern BR, Raizenne ME, Burnett RT, Jones L, Kearney J, Franklin CA. Air pollution and childhood respiratory health: Exposure to sulfate and ozone in 10 Canadian rural communities. *Environ Res*. 1994;66:125–42.
18. Künzli N, Lurmann F, Segal M, Ngo L, Balmes J, Tager IB. Association between lifetime ambient ozone exposure and pulmonary function in college freshmen - results of a pilot study. *Environ Res*. 1997;72:8–23.
19. Vedal S, Petkau J, White R, Blair J. Acute effects of ambient inhalable particles in asthmatic and nonasthmatic children. *Am J Respir Crit Care Med*. 1998;157(4 PART I):1034–43.
20. Lefohn AS, Knudsen HP, McEvoy LR. The use of kriging to estimate monthly ozone exposure parameters for the Southeastern United States. *Environ Pollut*. 1988;53:27–42.
21. Stahl K, Moore RD, Floyer JA, Asplin MG, McKendry IG. Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density. *Agric Meteorol*. 2006;139:224–36.
22. Schloeder CA, Zimmerman NE, Jacobs MJ. Comparison of methods for interpolating soil properties using limited data. *Soil Sci Soc Am J*. 2001;65:470.
23. Zimmerman D, Pavlik C, Ruggles A, Armstrong MP. An experimental comparison of ordinary and universal kriging and inverse distance weighting. *Math Geol*. 1999;31:375–90.
24. Li J, Heap AD. A review of comparative studies of spatial interpolation methods in environmental sciences: performance and impact factors. *Ecol Inf*. 2011;6:228–41.
25. Li J, Heap A. A review of spatial interpolation methods for environmental scientists. *Geoscience Australia, Record 2008/23*, 137.
26. Sailor D. Determinants of indoor and outdoor exposure to ozone and extreme heat in a warming climate and the health risks for an aging population. In: Science to Achieve Results (STAR) Indoor Air & Climate Change Progress Review Meeting and Webinar. Science to Achieve Results (STAR) Indoor Air & Climate Change Progress Review Meeting and Webinar, Office of Research and Development (ORD) (United States Environmental Protection Agency, Washington, DC 2016).
27. Emerson MO, Bratter J, Howell J, Jeanty PW, Cline, M. Houston region grows more racially/ethnically diverse, with small declines in segregation. A Joint Report Analyzing Census Data From 1990, 2000 and 2010. Kinder Institute for Urban Research & The Hobby Center for the Study of Texas; 2012.
28. Bethel HL, Sexton K, Linder S, Delclos G, Stock T, Abramson S, et al. A closer look at air pollution in Houston: identifying priority health risks. Houston: Mayor's Task Force on the Health Effects of Air Pollution; 2006.
29. Couzo E, Jeffries HE, Vizuete W. Houston's rapid ozone increases: preconditions and geographic origins. *Environ Chem*. 2013;10: 260–68.
30. Li G, Zhang R, Fan J, Tie X. Impacts of biogenic emissions on photochemical ozone production in Houston, Texas. *J Geophys Res Atmos*. 2007;112:D10309.
31. Wiedinmyer C, Guenther A, Estes M, Strange IW, Yarwood G, Allen DT. A land use database and examples of biogenic isoprene emission estimates for the state of Texas, USA. *Atmos Environ*. 2001;35:6465–77.
32. Cowling E, Furness C, Dimitriades B, Parrish D. Final rapid science synthesis report: findings from the Second Texas Air Quality Study (TexAQ5 II)– Final report to the Texas Commission on Environmental Quality, TCEQ Contract Number 582–4-65614. Southern Oxidants Study Office, North Carolina State University; 2007.
33. Daum PH. A comparative study of O₃ formation in the Houston urban and industrial plumes during the 2000 Texas Air Quality Study. *J Geophys Res*. 2003;108:4715
34. Vizuete W, Kim BU, Jeffries H, Kimura Y, Allen DT, Kioumourtzoglou MA, et al. Modeling ozone formation from industrial emission events in Houston, Texas. *Atmos Environ*. 2008;42: 7641–50.
35. Collins FC. A comparison of spatial interpolation techniques in temperature estimation. Blacksburg: Virginia Polytechnic Institute and State University; 1995.
36. Lam NS-N. Spatial interpolation methods: a review. *Cartogr Geogr Inf Sci*. 1983;10:129–50.
37. Nikiforov SV, Aggarwal M, Nadas A, Kinney PL. Methods for spatial interpolation of long-term ozone concentrations. *J Expo Anal Environ Epidemiol*. 1998;8:465–82.
38. Switzer, P., Sailor, D., Lam, N. S.-N., Battisti, D. S., Naylor, R. L., G., N., Le Sueur, D. Geostatistics, rare disease and the environment. *Env Heal Perspect*. 2004;22:67–85.
39. Toggweiler J, Key R. Ocean circulation: thermohaline circulation. *Encycl Atmos Sci*. 2001;4:1549–55.
40. Baafi EY, Kim YC. Comparison of different ore reserve estimation methods using conditional simulation. *Min Eng*. 1983;35:12.
41. Wong DW, Yuan L, Perlin SA. Comparison of spatial interpolation methods for the estimation of air quality data. *J Expo Anal Environ Epidemiol*. 2004;14:404–15.
42. Oliver MA. Geostatistics, rare disease and the environment. *Spatial Analytical Perspectives on GIS*, (Taylor and Francis: London, 1996). p. 67–85.
43. Shen L, Mickley LJ, Tai APK. Influence of synoptic patterns on surface ozone variability over the eastern United States from 1980 to 2012. *Atmos Chem Phys*. 2015;15:10925–38.
44. Camalier L, Cox W, Dolwick P. The effects of meteorology on ozone in urban areas and their use in assessing ozone trends. *Atmos Environ*. 2007;41:7127–37.
45. Bloomer BJ, Vinnikov KY, Dickerson RR. Changes in seasonal and diurnal cycles of ozone and temperature in the eastern U.S. *Atmos Environ*. 2010;44:2543–51.
46. Gotway CA. Fitting semivariogram models by weighted least squares. *Comput Geosci*. 1991;17:171–2.
47. Cressie N, Hawkins DM. Robust estimation of the Variogram .1. *J Int Assoc Math Geol*. 1980;12:115–25.
48. Cressie N. Statistics for spatial data. Vol. 4. Terra Nova (John Wiley & Sons, New York: 1992)
49. Ribeiro jr. PJ, Diggle PJ. Summary of contents of this issue. *Chem Fibers Int*. 2005;55:89.

50. R Development Core Team. R Internals. Vol. 1. Vienna: R Development Core Team; 2015. p. 63.
51. Kembell-Cook S, Parrish D, Ryerson T, Nopmongcol U, Johnson J, Tai E, et al. Contributions of regional transport and local sources to ozone exceedances in Houston and Dallas: Comparison of results from a photochemical grid model to aircraft and surface measurements. *J Geophys Res Atmos.* 2009;114:D00F02.
52. NARSTO Synthesis Team. An Assessment of Tropospheric Ozone Pollution—A North American Perspective. Pasco: North American Research Strategy for Tropospheric Ozone; 2000.
53. Wang Y, Jia B, Wang SC, Estes M, Shen L, Xie Y. Influence of the Bermuda High on interannual variability of summertime ozone in the Houston-Galveston-Brazoria region. *Atmos Chem Phys.* 2016;16:15265–76.
54. Zhu J, Liang XZ. Impacts of the bermuda high on regional climate and ozone over the United states. *J Clim.* 2013;26:1018–32.
55. Nielsen-gammon JW. The Houston heat pump: modulation of a land-sea breeze by an urban heat island. College Station: Department of Atmospheric Sciences; 2000.
56. Banta RM, Senff CJ, Nielsen-Gammon J, Darby LS, Ryerson TB, Alvarez RJ, et al. A bad air day in Houston. *Bull Am Meteorol Soc.* 2005;86:657–69.
57. Vizuete W, Jeffries HE, Tesche TW, Olaguer EP, Couzo E. Issues with ozone attainment methodology for Houston, TX. *J Air Waste Manag Assoc.* 2011;61:238–53.
58. Couzo E, Olatosi A, Jeffries HE, Vizuete W. Assessment of a regulatory model's performance relative to large spatial heterogeneity in observed ozone in Houston, Texas. *J Air Waste Manag Assoc.* 2012;62:696–706.
59. Mulholland JA, Butler AJ, Wilkinson JG, Russell AG, Tolbert PE. Temporal and spatial distributions of ozone in Atlanta: Regulatory and epidemiologic implications. *J Air Waste Manag Assoc.* 1998;48:418–26.
60. Hopkins LP, Ensor KB, Rifai HS. Empirical evaluation of ambient ozone interpolation procedures to support exposure models. *J Air Waste Manag Assoc.* 1999;49:839–46.
61. Gorai AK, Jain KG, Shaw N, Tuluri F, Tchounwou PP. Kriging analysis for spatio-temporal variations of ground level ozone concentration. *Asian J Atmos Environ.* 2015;9:247–58.
62. Kethireddy SR, Tchounwou PB, Ahmad HA, Yerramilli A, Young JH. Geospatial interpolation and mapping of tropospheric ozone pollution using geostatistics. *Int J Environ Res Public Health.* 2014;11:983–1000.
63. Gorai AK, Tuluri F, Tchounwou PB, Ambinakudige S. Influence of local meteorology and NO₂ conditions on ground-level ozone concentrations in the eastern part of Texas, USA. *Air Qual Atmos Heal.* 2015;8:81–96.
64. Yao X, Fu B, Lü Y, Sun F, Wang S, Liu M. Comparison of four spatial interpolation methods for estimating soil moisture in a complex terrain catchment. *PLoS ONE.* 2013;8:e54660