Relationship between particulate matter and childhood asthma – basis of a future warning system for Central Phoenix

R. Dimitrova¹, N. Lurponglukana², H. J. S. Fernando¹, G. C. Runger², P. Hyde³, B. C. Hedquist³, J. Anderson³, W. Bannister⁴, and W. Johnson⁴

¹Department of Civil Engineering & Geological Sciences, University of Notre Dame, Notre Dame, IN, USA
²Center for Health Information and Research, Arizona State University, Tempe, AZ, USA
³School for Engineering of Matter, Transport and Energy, Arizona State University, Tempe, AZ, USA
⁴Industrial Engineering, Arizona State University, Tempe, AZ, USA

Received: 16 June 2011 – Accepted: 10 October 2011 – Published: 25 October 2011

Correspondence to: R. Dimitrova (reneta.dimitrova@nd.edu)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

Statistically significant correlations between increase of asthma attacks in children and elevated concentrations of particulate matter of diameter 10 microns and less (PM$_{10}$) were determined for metropolitan Phoenix, Arizona. Interpolated concentrations from a five-site network provided spatial distribution of PM$_{10}$ that was mapped onto census tracts with population health records. The case-crossover statistical method was applied to determine the relationship between PM$_{10}$ concentration and asthma attacks. For children ages 5–17, a significant relationship was discovered between the two, while children ages 0–4 exhibited virtually no relationship. The risk of adverse health effects was expressed as a function of the change from the 25th to 75th percentiles of mean level PM$_{10}$ (36 µg m$^{-3}$). This increase in concentration was associated with a 12.6% (95% CI: 5.8%, 19.4%) increase in the log odds of asthma attacks among children ages 5–17. Neither gender nor other demographic variables were significant. The results are being used to develop an asthma early warning system for the study area.

1 Introduction

The prevalence of asthma, the single-most common chronic childhood disease in developed nations (Janson et al., 1997; Committee TISoAaAiCIS, 1998; WHO, 2003, 2007), has been reported as increasing in many countries (Magnus and Jaakkola, 1997; Akinbami et al., 2002; Burney, 2002; Lin et al., 2005), including in the United States (Brown et al., 1997; Vollmer et al., 1998; Holgate, 1999; CDC, 2011). The prevalence of asthma in the US has increased 75% from 1980 until 1994, according the Centers for Disease Control and Prevention (CDC, 1998). Although asthma has become a major public health problem for Americans of all ages, children are particularly affected. According to surveys, the most substantial increase has occurred amongst children of age 0–4 yr (160%, from 22.2 per 1000 to 57.8 per 1000) and persons of age
5–14 yr (74 %, from 42.8 per 1000 to 74.4 per 1000). Over 10 million US children under age 18 were diagnosed with asthma in 2008 (CDC, 2008). During the past decade, the statewide prevalence of asthma in Arizona has continued to increase, now being higher than the national average (CDC, 2003). In Phoenix, Arizona, asthma is a major public health issue amongst children of age 0–19, which represent 38 % of all asthma cases in Arizona (ADHS, 2009a). In 2009, 21 % of youth and 15 % of adults in Arizona have been diagnosed with asthma (ADHS, 2009b). Increased asthma is a menace for many growing cities in arid regions, such as Phoenix, and hence the United States Environmental Protection Agency (US EPA) is paying particularly close attention to its causes.

Many studies have reported that measurable excesses in pulmonary function decrements, respiratory symptoms, hospital and emergency department admissions, and mortality in human population are associated with ambient levels of various indicators of PM exposure, most notably PM$_{10}$ and PM$_{2.5}$ (US EPA, 2004). Environmental exposures may increase the risk of developing asthma, and they may also increase the frequency and severity of asthma incidents. A number of studies have investigated how short-term exposure to PM is related to mortality or to lung functions and respiratory symptoms. An increase of PM$_{10}$ results in a greater use of asthma medication, frequent visits to medical practitioners and increase of asthma-related hospital admissions (Schwartz et al., 1993; Lipsett et al., 1997; Peters et al., 1997; Atkinson et al., 1999; Norris et al., 1999; Delfino et al., 2002; Maciejczyk et al., 2004; Monteil, 2008; Yeatts et al., 2009; Scapellato et al., 2009). In these studies the association between asthma attacks and different triggers as such as PM$_{10}$, ozone (O$_3$), nitrogen dioxide (NO$_2$), sulfur dioxide (SO$_2$), and carbon monoxide (CO) have been investigated. Some reported significant correlations between asthma-related hospital admissions and concentrations of PM$_{10}$ and NO$_2$ (Lipsett et al., 1997), PM$_{10}$ and SO$_2$ and/or SO$_4$ (Peters et al., 1997; Atkinson et al., 1999) or PM$_{2.5}$ or PM$_{10}$ and CO (Norris et al., 1999; Shepard et al., 1999) whereas others, using multi-pollutant models, contend that asthma is related to PM concentrations only (Nauenberg and Basu, 1999; Tolbert et al., 2000). All
studies that used single-pollutant models have reported significant association of particles and asthma admissions (Schwartz et al., 1992; Maciejczyk et al., 2004; Monteil, 2008; Scapellato et al., 2009). In spite of ozone being considered as an asthma trigger itself some studies indicate that both PM$_{10}$ and O$_3$ are associated with increases in asthma (Karin et al., 2009), whereas others disagree with this notion (Lipsett et al., 1997; Atkinson et al., 1999; Nauenberg and Basu, 1999). Note that high pollution episodes are largely governed by local conditions, high O$_3$ levels are typical of the summer (generated by photochemical reactions between oxides of nitrogen and volatile organic compounds VOC) and elevated PM$_{10}$ levels may occur during the summer or winter (as in Arizona and California for example). Many of the epidemiologic studies on PM$_{10}$ morbidity have been conducted under relatively high PM$_{10}$ levels, but there are studies that dealt with PM$_{10}$ concentrations as low as 50 µg m$^{-3}$ or below (Jacobs et al., 1997; Nauenberg and Basu, 1999; Sheppard et al., 1999; Tolbert et al., 2000), which are much lower than the USEPA’s 24-h PM$_{10}$ National Ambient Air Quality Standards (NAAQS) of 150 µg m$^{-3}$. Although the causes of childhood asthma are not well understood, air pollution appears to have a clear role in triggering attacks.

A review of historical data of “criteria” pollutants in metropolitan Phoenix shows that PM$_{10}$ and 8-h ozone concentrations exceed NAAQS. Ozone concentrations are high in the summer and exceed NAQQS on some days; in contrast, winter concentrations of ozone are low. The occurrence of high concentrations of ozone and PM$_{10}$ over two disparate periods provided an opportunity to investigate the influence of PM$_{10}$ and ozone on asthma attacks. The indications are that in the summer the asthma incidences in Phoenix are reduced (so as PM$_{10}$), despite an increase in ozone, suggesting that the major trigger is particulate matter. Therefore, it is instructive to investigate how asthma is related to PM$_{10}$, irrespective of other pollutants. To this end, the US EPA awarded a challenge grant to Arizona Department of Environmental Quality (ADEQ) to study how PM$_{10}$ is correlated with asthma occurrences. PM$_{10}$ is comprised of both fine particles less than 2.5 µm (PM$_{2.5}$) and coarse particles (smaller than 10 and larger than 2.5 µm). Note that PM$_{2.5}$ also can be a trigger, but insufficient data is available from the
Many studies consider only fine particles, but coarse particles, which are of mainly geological origin, contribute significantly to the total mass of PM$_{10}$ in arid environments. In metropolitan Phoenix in 2005 and 2006, for example, the four air monitoring sites with both PM$_{10}$ and PM$_{2.5}$ monitors had an average ratio of PM$_{2.5}$ to PM$_{10}$ of 0.28, considerably higher than the PM$_{2.5}$/PM$_{10}$ mass ratios of 0.1 to 0.17 reported in most other studies (Countess and Richard, 2003; Pace, 2005; Ono, 2005; MRI, 2005, 2006). While we acknowledge that a complex mix of gaseous, semi-volatile, and particulate pollutants and their physicochemical properties contribute to health outcomes, in the present study we concentrated on PM$_{10}$, as this is the principal pollutant that exceeds NAAQS during late autumn and winter. It was hoped to develop an association between PM$_{10}$ and asthma-related hospital admittances so that an early warning system could be developed for Phoenix based on predicted PM$_{10}$ concentrations. The results of the present study are also applicable to climatologically similar urban areas, for example, many Southern California cities (Cal/EPA, 2006), Salt Lake City, Utah, Denver, Colorado, El Paso, Texas, and so on.

The interpolation of PM$_{10}$ measurements from a five-site network of continuous monitors enabled the use of case-crossover statistical method to be applied to census-tract-based health records to determine the relationship between PM$_{10}$ and asthma incidences. Spatial concentration fields of PM$_{10}$ for 2005 and 2006 were constructed and then linked with the public health data. Inverse Distance Weighting (IDW) was used in this study, but it was evaluated and verified by the more sophisticated Ordinary Kriging (OK) with additional data from an expanded network in December 2007 through February 2008.

The goal of this study is to clarify the association of asthma incidents (primarily emergency department visits and hospital admissions with a diagnosis of asthma) with elevated concentrations of particulate matter 10 µm and smaller (PM$_{10}$). It breaks new ground in two ways: it demonstrates how adequate concentration fields can be derived by interpolation from a sparse (five site) network of continuous PM$_{10}$ monitors; and, it specifically focuses on associations of childhood asthma with PM$_{10}$, a focus that could...
lead to the development of a warning system for the asthmatic children of Phoenix. This work complements numerous previous studies on health outcomes of air pollution for particular groups, e.g. pregnant women and infants (Viana et al., 2008) and by specific causations e.g. congenital defects or prematurity (Dolk and Vrijheid 2003; Hansen et al., 2006).

2 Data and methods

2.1 Air quality data

Spatially distributed PM$_{10}$ concentrations were estimated by interpolating the measured concentrations from a permanent network of five continuous monitors (hourly data) in Central Phoenix operated by two public agencies: Maricopa County Department of Air Quality (MCDAQ) and Air Quality Division of the Arizona Department of Environmental Quality (ADEQ). As the only NAAQS of interest here was for the 24-h average of PM$_{10}$, and the aim was to examine the relationship between PM$_{10}$ concentrations and daily asthma incidents, daily averages of PM$_{10}$ were calculated from reported hourly averages measured by the permanent monitors. Each of these permanent sites has its unique combination of local emission sources and susceptibility to urban transport that together, along with the desert background, result in the measured concentrations. Each has its own degree of representativeness as a footprint: i.e. how far away from the monitor do the measured concentrations actually prevail? The PM$_{10}$ variation among these five sites can be seen in their 24-h averages, given as a time series for 2005–2006 (Fig. 1). The names of the monitoring stations are given in the captions of Figs. 1 and 2.

West 43rd Avenue (WF) and Durango complex (DC) sites have PM$_{10}$ concentrations considerably higher than the other three, especially in winter, and both have a relatively high frequency of violations of NAAQS. This pattern can be understood in light of dense localized emissions from the extractive and material handling industries along
the Salt River combined with their low-elevation location that places the area downwind of most nocturnal and early morning urban emissions. They represent the worst PM$_{10}$ air quality in metropolitan Phoenix.

### 2.2 Asthma data

The primary health data were obtained from the Arizona Department of Health Services (ADHS) for the asthma incidents reported between 1 January 2004 and 31 December 2006. Incident locations are based on the census track of the patient's home address. The analysis focuses on school-age children, and it is assumed that school locations are near the home census tract. These data include both hospitalization and emergency room visits throughout the Maricopa County, as well as patient demographics such as gender and age. In spite of standard quality control, additional steps were taken to ensure the accuracy of the data. For example, there were cases where the same patient went to an emergency room and was later admitted to the hospital. This single patient had multiple records on the same day with his status changing from emergency to inpatient. Such redundant records were deleted to leave all but a single asthma incident. Also, those subsequent medical encounters occurring sufficiently long after the initial visit were considered to be a distinct incident. This time period is commonly referred to as a "washout" period or "gap". Several washout periods were explored in this study. The change in the number of incidents is minimal for washout periods of 7 to 28 days (from 10,095 to 9,631). Therefore, a washout period of seven days was selected because it is a common choice for studies such as this. If the same person had multiple records within the seven days, only the initial record was kept.

Preliminary analyses defined the demographic groups, sample size, data cleaning strategies, and the precise study area for the final analyses of the work presented here. The present study focused on asthma incidents within the 168 census tracts of metropolitan Phoenix from 1 January 2005 through 30 September 2006 because of lack of full records of air quality data for 2004. All asthma events within five miles of a continuous PM$_{10}$ monitor were included (Fig. 2). This study area accounted for 75% of the total asthma incidents of Maricopa County.
2.3 Interpolation of PM$_{10}$ concentrations

Practical constraints make it impossible to obtain data at every desired point. Thus, interpolation is imperative for graphing, analyzing and understanding the spatial distribution of air pollution throughout the metropolitan area. Inverse distance weighting (IDW) is one of the simplest interpolation methods. A neighborhood in the vicinity of the interpolated point was identified and a weighted average was taken of the observation values within it. The weights are a decreasing function of distance. The user has control over the mathematical form of the weighting function, the size of the neighborhood (expressed as a radius or a number of points), in addition to other options.

As mentioned, Kriging methods were also tested for interpolation, including ordinary kriging. This method utilizes a variogram model to express the spatial variation of data points, and then minimize their error. The word “kriging” is synonymous with “optimal prediction”. Ordinary kriging (OK) is a geostatistical approach to modeling. Instead of weighting nearby data points by some power of their inverted distance, it relies on the spatial correlation structure of the data to determine the weighting values. This is a more rigorous approach than IDW, as correlation between data points determines the estimated value at an unsampled point. Ordinary kriging makes the assumption of normality among the data points, as well as an unknown constant trend. This differs from simple or universal kriging that assumes a known constant trend and/or model.

In our study, IDW interpolation was used due to lack of enough data points (at least 10) to apply OK. Nevertheless, as discussed later, the adequacy of IDW was established by comparing the two methods over a limited time period; for this, a stop-gap monitoring network was installed to augment continuous monitors (Fig. 1).

Six additional monitors were placed (marked in blue at Fig. 2) – four temporary sites and two permanent ones (GR and SP). The newly established permanent sites began operating in 2007. The four temporary sites were deployed from December 2007 through February 2008. These additional sites enabled the use of OK, and thus enabling evaluation of IDW for the special deployment period and hence for the earlier period with five instead of eleven PM$_{10}$ monitors.
The interpolated 24-h averaged concentrations were linked to 168 census tract areas in metropolitan Phoenix from 1 January 2005 through 30 September 2006, a period for which complete air monitoring records were available. These data were used further for data analyses and linking with the health data.

2.4 IDW method verification

The surfaces of PM$_{10}$ obtained by both methods were compared, with Fig. 3 being only one example of such comparisons of concentration fields for one particular day 21 December 2007. Values from five PM$_{10}$ monitors were used in IDW and eleven in OK. Differences for the 24-h average PM$_{10}$ can be seen in the northwest, where the additional data led to reduced PM$_{10}$ concentrations (Fig. 3).

In a comparison between the two interpolation methods for December 2007 (Fig. 4), it can be seen that better agreement is achieved in the central part of the study area, in close proximity to the five permanent PM$_{10}$ monitors. The overall difference between the two methods in the central part of the study domain is less than 8%. Mean Absolute Error is in range 0–5 µg m$^{-3}$ and the average concentration is 64 µg m$^{-3}$ (in range 56–72 µg m$^{-3}$). Even though the OK interpolation had the benefit of two additional monitors in the central part of the study area (GR and VGC), the differences between the interpolated surfaces was surprisingly low. Therefore, the five monitors used with IDW interpolation are considered representative of the spatial concentration gradients in Central Phoenix. As expected, greater disagreement occurs at the outer edges of the study area, in census tracts close to the temporary PM$_{10}$ monitors.

Some statistical measures (description is given in Appendix A) for IDW interpolation performance, compared with the OK method, are listed in Table 1. Spatial averaged was made over the entire study area (Fig. 4) in addition to time averaged, and the minimum and maximum values within the study domain are also shown. The averaged bias are positive (IDW slightly overestimates the concentrations provided by OK method), but negative bias (underestimation) are also seen for a few census tracts. The average errors are less than 15% (NME, MFR) and NMSE is less than 1%. The index of
agreement between both surfaces is reasonably high (82.3%) with a correlation coefficient 0.68.

In all, the IDW was shown to be sufficiently accurate to produce spatial concentration fields for the health analysis within five miles of each monitoring location, with the resulting interpolation domain for health study shown in Fig. 2.

2.5 Methods of analysis

The case-crossover approach (Maclure, 1991), similar to crossover and matched-pair case-control studies (Hosmer and Lemeshow, 2000), has been developed to study transient, short-term exposure effects on the risk of rare acute events. The "case" is a person with the event of interest at a certain time called the “hazard period”, while the control (called the referent) is the same person at a different time called the “control period”. The key advantage of this design is that each case comes with its own control. The exposure information for each subject during the hazard period is compared with the exposure information during the control or “referent” period. With an appropriate referent period, the case-crossover analysis controls for long-term trends, seasonal effects, and other covariates that change slowly with time.

In the bidirectional approach, Bateson and Schwartz (1999, 2001) examined sampling methods for the referent to reduce the bias by selecting from either side of the event. Also, disjoint referent periods have been recommended for bias concerns (Levy et al., 2001). With a control close in time to the event, these methods avoid confounding due to subject differences and other long-term effects such as seasonality. A balance must be maintained, however, between a control too close in time that generates autocorrelation and a control too far removed in time that confounds long-term effects. Alternatives within a few weeks of the events have been considered. Neas et al. (1999) studied daily mortality in Philadelphia where the case period was the 48 h ending at midnight on the day of death (Schwartz and Dockery, 1992) and the control period was the same day of week 7, 14, or 21 days before and after the case period. Medina-Ramon et al. (2006) used a matching scheme from Bateson and Schwartz (1999,
2001) and a time-stratified partition by Lumley and Levy (2000) that chose control days only in the same month as the hospital admission for chronic obstructive pulmonary disease (http://www.nhlbi.nih.gov/health/dci/Diseases/Copd/Copd_WhatIs.html). Lin et al. (2005) calculated a one to seven day exposure average ending on the admission date as the exposure in the case period with control periods of two weeks before and after the admission date. Peel et al. (2005) studied the hospital admissions associated with ambient air pollution levels and respiratory health effects. The case period of three-day moving averages of pollutant concentration was selected within two weeks of the case period. The average was the average of pollutant concentrations on the same day as the visit, one day before, and two days before.

A case-crossover study was conducted in the present work with the referent period based on a time-stratified scheme to control confounding effects from seasonality and time trends. The referents were selected within a 28-day stratum (four weeks) at 7, 14, and 21 days before or after the case (same day of the week), provided they occurred in the same stratum. Consequently, each case has three controls.

Several potential primary and secondary effects were considered. The first primary effect was the 24-h average PM$_{10}$ concentration from the IDW interpolation of the observations on the day of the asthma event. The second primary effect, called the “lag”, was the average PM$_{10}$ concentration one to five days before the event, here the “lag” being the delay in days between the exposure and the health effect. For example, a one-day lag is the average PM$_{10}$ of the previous day and the event day, abbreviated as “lag1”. Secondary effects were the covariates of age, gender, ethnicity, and place of service. Place of service is a categorical variable with two values: emergency and in-patient. Age is encoded into four categories: 0–4, 5–9, 10–14, and 15–17 yr. Ethnicity followed US Census Bureau definitions. Preliminary analysis detected no differences between places of service; therefore, this variable was not considered further. Even though the case-crossover design controls for the main effects of these patient-level covariates, the interactions of these covariates with the event-day PM$_{10}$ concentration and its various lags were also considered.
The analysis considered subsets of these covariates through conditional logistic regression (Hosmer and Lemeshow, 2000). All-subsets regression was used to investigate models with demographic variables and with and without interaction (cross-product) terms. Short description of terms used in this analysis is listed below (Appendix B).

2.6 Statistical model validation

The model used for the results discussed in the following section was a conditional linear regression model. The traditional assumption for a logistic regression model is linear relationship of the log odds to the predictor variables. However, to assess the linearity assumption a model without any linear or monotonic assumptions was employed for validation. A recommended approach is to use categorical variables to define ranges of PM$_{10}$ (Steenland and Deddens, 2004) and model the log odds. The categories are defined from the quartiles because results are often reported as the effect from a change equal to the interquartile range of PM$_{10}$. The PM$_{10}$ concentrations were linked to asthma events by the indicator variables in the regression models with the first quartile as the baseline. The odds ratio can be readily interpreted as a measure of how likely or unlikely it is for an asthma incident to be associated with a daily mean PM$_{10}$ concentration. Odds ratios greater than 1.0 imply that the asthma event is more likely to be associated with the PM$_{10}$ concentration and vice versa. The natural logarithm of the odds ratio (log odds) of an asthma incident, “log odds” of the y-axis, plotted against the median PM$_{10}$ in the category in Fig. 5, illustrates the monotonic, approximately linear relationship of the log odds versus the quartiles of the daily mean.

3 Results and discussion

Asthma incidents were described by the age group, day of the week, and month, resulting in a dataset containing about 10,000 events within a five-mile radius from the
nearest monitoring station (Figs. 6–8). About two thirds of these events are for ages 0–17.

The event day PM\(_{10}\), also called “Dailymean”, and lag variables were considered in a statistical analysis as shown in Table 2. Only the PM\(_{10}\) on the day of the event (Dailymean) and the lag 1 PM\(_{10}\) are significant at a 1\% significance level. Although there variability is the \(p\)-values if one applied a traditional Bonferroni correction to adjust the significance level for the nine multiple tests, one might require a \(p\)-value less than 0.05/9 = 0.0056 and only the Dailymean is significant. Consequently, further analysis was restricted to the Dailymean as the principal covariate of PM\(_{10}\).

An all-subsets regression considered Dailymean, the demographic variables, and interactions in models for the health effects. The conditional regression analysis indicated that the Dailymean alone was a sufficient covariate for asthma incidents, but that age was a potential effect modifier. The age effect is discussed further below. No effect due to gender or the other demographic variables was detected.

The case-crossover statistics are presented and discussed for three different samples of the childhood asthma population: all events (ages 0–17), preschool events only (ages 0–4) and events excluding the preschool group (ages 5–17). The results from the different samples exhibit certain differences (Table 3).

The coefficients in Table 3 vary considerably indicating somewhat different relationships between the increased probability of an asthma event and the event-day PM\(_{10}\) concentrations among different subpopulations, especially for the case of preschool only (negative coefficient).

The change in 24-h average PM\(_{10}\) from the 25th to 75th percentile, defined as the interquartile range (IQR), is 36 µg m\(^{-3}\). A change in PM\(_{10}\) concentration equal to the IQR is commonly used as the change magnitude to report the adjusted odds ratio. As shown in Table 3 the adjusted odds ratio for all events is estimated to be 1.094, meaning that every additional 36 µg m\(^{-3}\) of Dailymean PM\(_{10}\) is estimated to increase the odds ratio of an asthma incident for a person under 17 yr old by 9.4\% (4.1\%, 14.7\%), where the values in parentheses provide a 95\% confidence interval. In the preschool group
(ages 0–4 yr) only, no effect from Dailymean PM$_{10}$ was found. The odds ratio was estimated at 0.997 and the estimated change was not significant at $-0.3\%$ ($-8.9\%, 8.2\%$) with a 95% confidence interval that contains zero. Further analysis detected a stronger relationship between Dailymean and asthma incidents when the preschool age group was excluded. All-subsets regression again indicated that Dailymean alone was a sufficient covariate for asthma incidents. Every additional 36 $\mu$g m$^{-3}$ of Dailymean PM$_{10}$ increases the odds ratio of an asthma incident for a child in ages 5–17 yr old by 12.6% (5.8%, 19.4%) with 95% confidence intervals in parentheses.

4 Conclusions

PM$_{10}$ concentrations were found to have statistically significant associations with asthma incidents in Central Phoenix at the 95% confidence level. For children ages 5–17 a much stronger effect was noted than determined by previous studies, while preschool children ages 0–4 exhibited virtually no effects. Previous studies have reported the risk of adverse health effects as a function of the change of Dailymean PM$_{10}$ from the 25% to 75% percentile; in this study the corresponding change was 36 $\mu$g m$^{-3}$. This increase in PM$_{10}$ is associated with a 13% increase in the probability of asthma incidents among children ages 5–17; gender was insignificant. Note that the concentrations considered are much lower than the 24-h PM$_{10}$ standard of 150 $\mu$g m$^{-3}$.

The lack of a preschool effect could be explained by the difficulty of asthma diagnoses at the younger ages and/or their possible spending of more time indoors.

With only a sparse network of five monitoring sites in a large metropolitan area, Inverse Distance Weighting (IDW) interpolation provided an adequate spatial distribution of urban PM$_{10}$ concentrations efficiently and rapidly. There were differences, but not unacceptably large, between IDW interpolated concentrations and those from the more sophisticated ordinary kriging technique. The kriging interpolation was realized by deploying temporary monitors over a special test period and utilizing concentrations from two permanent sites that began operation in 2007, thus increasing the number of
monitors in central Phoenix from five to eleven. Because the two methods produced comparable distributions, and because the period of analysis had only five monitoring sites, inverse distance weighting was the method of choice for this health analysis.

It has been suggested, based on the present study, to install neural networks for each of the permanent monitors to predict PM$_{10}$ based on several years of prior measurements at each site as well as meteorological predictions for those sites based on a deterministic model (Mammarella et al., 2009; Fernando et al., 2011). The PM$_{10}$ predictions could be interpolated to a census tract grid, which together with some forecaster intervention and examining could be used to issue health warnings for the area. Such a predictive system would exemplify the use of best-available scientific and technological tools for issuance of health warnings. The proposed methodology has been well-received within the USEPA, as evident from the issuance of funding to develop a pilot health warning system.

Appendix A

A1 Definitions of statistics

The following indicators were used for performance evaluation. Here IDW is the value obtained by Inverse Distance Weighting method, OK is the value obtained by Ordinary Kriging method, and \( \overline{IDW} \) and \( \overline{OK} \) are the mean values.
\[ MB = \frac{1}{N} \sum_{i=1}^{N} (IDW_i - OK_i) \quad \text{(Mean Bias)} \quad (A1) \]

\[ MAE = \frac{1}{N} \sum_{i=1}^{N} |IDW_i - OK_i| \quad \text{(Mean Absolute Error)} \quad (A2) \]

\[ RMSE = \sqrt{\frac{\sum_{i=1}^{N} (IDW_i - OK_i)^2}{N}} \quad \text{(Root Mean Square Error)} \quad (A3) \]

\[ NMB = \frac{\sum_{i=1}^{N} (IDW_i - OK_i)}{\sum_{i=1}^{N} OK_i} \times 100 \% \quad \text{(Normalized Mean Bias)} \quad (A4) \]

\[ NME = \frac{\sum_{i=1}^{N} |IDW_i - OK_i|}{\sum_{i=1}^{N} OK_i} \times 100 \% \quad \text{(Normalized Mean Error)} \quad (A5) \]

\[ MFB = \frac{\sum_{i=1}^{N} \left( \frac{(IDW_i - OK_i)}{0.5 \cdot (IDW_i + OK_i)} \right)}{100 \%} \quad \text{(Mean Fractional Bias)} \quad (A6) \]

\[ MFE = \frac{\sum_{i=1}^{N} \left( \frac{|IDW_i - OK_i|}{|IDW_i + OK_i| \cdot 0.5} \right)}{100 \%} \quad \text{(Mean Fractional Error)} \quad (A7) \]

\[ NMSE = \frac{\sum_{i=1}^{N} (IDW_i - OK_i)^2}{\sum_{i=1}^{N} IDW_i \cdot \sum_{i=1}^{N} OK_i} \times 100 \% \quad \text{(Normalized Mean Square Error)} \quad (A8) \]

\[ IA = 1 - \frac{\sum_{i=1}^{N} (IDW_i - OK_i)^2}{\sum_{i=1}^{N} ((IDW_i - OK)^2 + (OK_i - OK)^2)} \quad \text{(IA Index of Agreement)} \quad (A9) \]
Appendix B

Description of the case-crossover method

Here we present a general discussion of the odds ratio, the typical statistic employed in reporting the results from the case-crossover method, and continues with a discussion of conditional logistic regression, and it concludes with the specific mathematical formulation of the odds ratio used in the present work.

The odds ratio is a measure of effect size particularly important in logistic regression. It is defined as the ratio of the odds of an event occurring in one group to the odds of it occurring in another group, or to a sample-based estimate of that ratio. These groups might be men and women, an experimental group and a control group, or any other dichotomous classification. If the probabilities of the event in each of the groups are \( p \) (first group) and \( q \) (second group), then the odds ratio is:

\[
\frac{p/(1-p)}{q/(1-q)} = \frac{p(1-q)}{q(1-p)} \tag{B1}
\]

An odds ratio of 1 indicates that the condition or event under study is equally likely in both groups. An odds ratio greater than 1 indicates that the condition or event is more likely in the first group. And an odds ratio less than 1 indicates that the condition or event is less likely in the first group. The odds ratio must be greater than or equal to zero. As the odds of the first group approaches zero, the odds ratio approaches zero. As the odds of the second group approaches zero, the odds ratio approaches positive infinity.
The increased use of logistic regression in medical and social science research means that the odds ratio is commonly used as a means of expressing the results in some forms of clinical trials, in survey research, and in epidemiology, such as in case-control studies. It is often abbreviated “OR” in reports. When data from multiple surveys is combined, it will often be expressed as “pooled OR”.

Conditional logistic regression has been widely applied to model stratified data, such as case-crossover analysis and case-control analysis (Lin et al., 2005; Neas et al., 1999; Figuerias et al., 2005; Redelmeier and Tibishirani, 1997). Logistic regression is a form of regression used when the response is binary. Let \( x \) be a vector of \( k \) predictor variables (or covariates). Let \( Y \) be a binary response (\( y = 0 \) or \( y = 1 \)). The conditional probability that \( Y = 1 \) is \( P(Y = 1|x) = \rho(x) \). The logit of the logistic regression model is

\[
f(x) = \beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p
\]

(B2)

so the logistic regression model is

\[
p(x) = \frac{e^{f(x)}}{1 + e^{f(x)}}
\]

(B3)

Let consider the system that has a binary response \( y \) and a two-level covariate \( x \), \( x = 0 \) or \( x = 1 \). The logit transformation is defined in terms of \( \rho(x) \) as

\[
f(x) = \ln[p(x)/(1 - p(x))] = \beta_0 + \beta_1 x
\]

(B4)

The conditional logistic regression works in similar fashion as regular logistic regression. It considers the stratification structure in the data specifying which individuals belong to which strata. Suppose there are \( K \) strata with \( n_k \) subjects in the \( k \)th stratum, where \( k = 1, 2, 3, \ldots, K \). There are \( n_{1k} \) case subjects, \( n_{0k} \) control subjects, and \( n_k = n_{1k} + n_{0k} \) (Hosmer and Lemeshow, 2000). The logistic regression model can be shown as

\[
p_k(x) = \frac{e^{\beta_{0k} + \beta' x}}{1 + e^{\beta_{0k} + \beta' x}}
\]

(B5)
where is a nuisance parameter with the contribution of all terms constant within the \(k\)th stratum, \(\beta' = (\beta_1, \beta_2, \ldots, \beta_k)\) are the vector of coefficients of covariates, \(x\).

To analyze case-crossover design using conditional logistic regression, the PROC PHREG from SAS® statistical software package was used. Conditional logistic regression can be performed by the PHREG procedure by using the discrete logistic model and forming a stratum for each matched set. The dummy survival times are needed to be created so that all the cases in a matched set have the same event time value, and the corresponding controls are censored at later times.

From the logistic regression model defined earlier. The odds of \(y = 1\) when the covariates have values \(x_1\) is defined as \(p(1)/(1 - p(1))\) and the odds of \(y = 0\) when the covariates have values \(x_0\) is defined as \(p(0)/(1 - p(0))\). The odds ratio (OR) is defined as the odds for \(x_1\) to the odds for \(x_0\), odds ratios (OR) can be shown to be

\[
\text{OR} = \frac{p(1)/[1 - p(1)]}{p(0)/[1 - p(0)]} = \frac{\left(\frac{e^{\beta_0 + \beta_1}}{1 + e^{\beta_0 + \beta_1}}\right)}{\left(\frac{1}{1 + e^{\beta_0}}\right)} = \frac{\left(\frac{e^{\beta_0 + \beta_1}}{1 + e^{\beta_0 + \beta_1}}\right)}{\left(\frac{1}{1 + e^{\beta_0}}\right)} = \frac{e^{\beta_0 + \beta_1}}{e^{\beta_0}} = e^{(\beta_0 + \beta_1) - \beta_0} = e^{\beta_1}
\]

(B6)

The odds ratio is an associate measure of how much more likely (an odds ratio > 1), unlikely (an odds ratio < 1), or equally likely (an odds ratio = 1) it is for the response to present under \(x_1\) than under conditions \(x_0\).

**Acknowledgements.** The present work was carried out within the Arizona Children’s Health Challenge Grant Project supported by US Environmental Protection Agency (USEPA) in partnership with the Arizona Department of Environmental Quality (ADEQ) and the Arizona Department of Health Services (ADHS).
References

Akinbami, L. and Schoendorf, K.: Trends in childhood asthma: prevalence, health care utiliza-


Centers for Disease Control and Prevention, National Center for Health Statistics: Summary Health Statistics for US Children: National Health Interview Survey 2008, Hyattsville, Mary-
Relationship between particulate matter and childhood asthma

R. Dimitrova et al.

Introduction

Conclusions

References

Tables

Figures


Midwest Research Institute (MRI): Analysis of the fine fraction of particulate matter in fugitive dust, Final Report for Western Governors’ Association Western Regional Air Partnership (WRAP), MPI Project No. 110397, Midwest Research Institute, Kansas City, Missouri, October, 2005.

Midwest Research Institute (MRI): Background document for revisions to fine fraction ration factors used for AP-42 fugitive dust emission factors, for Western Governors’ Association Western Regional Air Partnership (WRAP), MPI Project No. 110397, Midwest Research Institute, Kansas City, Missouri, November, 2006.


Ono, D.: Ambient PM$_{2.5}$/PM$_{10}$ ratios for dust events from the Keeler dunes, Great Baron UAPCD, Bishop, CA, 2005.

Pace, T. G.: Examination of multiplier used to estimate PM$_{2.5}$ fugitive dust emissions from PM$_{10}$, EPA Emission Inventory Conference, Las Vegas NV, http://www.epa.gov/ttnchie1/conference/ei14/session5/pace.pdf, April, 2005.

Peel, J., Tolbert, P., Klein, M., Metzger, K., Flanders, W., Todd, K., Mulholland, J., Ryan, P., and Frumkin, H.: Ambient air pollution and respiratory emergency department visits, Epidemiology, 16, 164–174, 2005.


Table 1. Performance statistics for IDW validation.

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Average</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Bias – MB (µg m(^{-3}))</td>
<td>4.5</td>
<td>-1.5</td>
<td>10.9</td>
</tr>
<tr>
<td>Mean Absolute Error – MAE (µg m(^{-3}))</td>
<td>5.3</td>
<td>0.5</td>
<td>10.9</td>
</tr>
<tr>
<td>Root Mean Squared Error – RMSE (µg m(^{-3}))</td>
<td>6.3</td>
<td>0.6</td>
<td>12.5</td>
</tr>
<tr>
<td>Normalized Mean Bias – NMB (%)</td>
<td>14.4</td>
<td>-0.5</td>
<td>26.8</td>
</tr>
<tr>
<td>Normalized Mean Error – NME (%)</td>
<td>15.8</td>
<td>6.6</td>
<td>27.5</td>
</tr>
<tr>
<td>Mean Fractional Bias – MFB (%)</td>
<td>13.4</td>
<td>-3.7</td>
<td>34.8</td>
</tr>
<tr>
<td>Mean Fractional Error – MFE (%)</td>
<td>15.1</td>
<td>1.5</td>
<td>34.8</td>
</tr>
<tr>
<td>Normalized Mean Square Error – NMSE (%)</td>
<td>0.2</td>
<td>0.001</td>
<td>0.6</td>
</tr>
<tr>
<td>Index of Agreement – IA (%)</td>
<td>82.3</td>
<td>50.9</td>
<td>99.8</td>
</tr>
<tr>
<td>Coefficient of determination – (R^2)</td>
<td>0.68</td>
<td>0.60</td>
<td>0.80</td>
</tr>
</tbody>
</table>
Table 2. Model statistics for each main effect (preschool age group excluded).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Chi square</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dailymean</td>
<td>0.00331</td>
<td>11.7859</td>
<td>0.0006</td>
</tr>
<tr>
<td>Lag1</td>
<td>0.00272</td>
<td>6.7347</td>
<td>0.0095</td>
</tr>
<tr>
<td>Lag2</td>
<td>0.00103</td>
<td>0.9352</td>
<td>0.3335</td>
</tr>
<tr>
<td>Lag3</td>
<td>0.00112</td>
<td>1.2531</td>
<td>0.263</td>
</tr>
<tr>
<td>Lag4</td>
<td>0.00244</td>
<td>5.7373</td>
<td>0.0166</td>
</tr>
<tr>
<td>Lag5</td>
<td>0.00231</td>
<td>4.1032</td>
<td>0.0428</td>
</tr>
<tr>
<td>Lag6</td>
<td>0.00212</td>
<td>4.0581</td>
<td>0.044</td>
</tr>
<tr>
<td>Lag7</td>
<td>0.00168</td>
<td>2.522</td>
<td>0.1123</td>
</tr>
<tr>
<td>Lag8</td>
<td>-0.0001995</td>
<td>0.031</td>
<td>0.8602</td>
</tr>
</tbody>
</table>
Relationship between particulate matter and childhood asthma

R. Dimitrova et al.

Table 3. Case-crossover statistics relating asthma incidents to PM$_{10}$ concentrations in Phoenix metropolitan.

<table>
<thead>
<tr>
<th>Case</th>
<th>Coefficient</th>
<th>St. Error</th>
<th>p-value</th>
<th>Adjusted OR*</th>
<th>95 % CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>All events</td>
<td>0.0025</td>
<td>0.00075</td>
<td>0.0010</td>
<td>1.094</td>
<td>(1.041, 1.147)</td>
</tr>
<tr>
<td>Preschool only</td>
<td>–0.000092</td>
<td>0.00121</td>
<td>0.4457</td>
<td>0.997</td>
<td>(0.911, 1.082)</td>
</tr>
<tr>
<td>Preschool excluded</td>
<td>0.0033</td>
<td>0.00096</td>
<td>0.0006</td>
<td>1.126</td>
<td>(1.058, 1.194)</td>
</tr>
</tbody>
</table>

* Adjusted OR is the odds ratio adjusted by the approximate interquartile range from 25th to 75th percentile of 36 µg m$^{-3}$. 
Fig. 1. 24-h averaged PM$_{10}$ from five permanent monitors in Central Phoenix (abbreviations: CP – Central Phoenix, DC – Durango complex, WF – West 43rd Avenue, WP – West Phoenix, SS1 – Supersite).
Fig. 2. Study domain and location of the monitors: permanent (data used in the health study) and temporary (data used for validation of spatial interpolation method). Abbreviations: CP – Central Phoenix, DC – Durango complex, WF – West 43rd Avenue, WP – West Phoenix, SS1 – Supersite; GR – Greenwood; SP – South Phoenix; CSA - Community Service of AZ; MRV – Maryvale; VGC – Valley Garden Center; WAS – Weaver’s Auto Service.
Fig. 3. 3.24-h average PM$_{10}$ concentration [$\mu$g m$^{-3}$], obtained using two different methods: IDW and OK, 21 December 2007.
Fig. 4. Mean Absolute Error (µg m\(^{-3}\)) for December 2007 between IDW and OK methods – based on daily maps. (Abbreviations: CP – Central Phoenix, DC – Durango complex, WF – West 43rd Avenue, WP – West Phoenix, SS1 – Supersite; GR – Greenwood; SP – South Phoenix; CSA – Community Service of AZ; MRV – Maryvale; VGC – Valley Garden Center; WAS – Weaver’s Auto Service).
Fig. 5. Natural logarithm of the odds ratio (log odds) of asthma incidents versus categories of Daily mean PM$_{10}$ concentrations ($\mu$g m$^{-3}$). Categories are defined by quartiles of PM$_{10}$ and represented by the median PM$_{10}$ within the category.
Fig. 6. Asthma incidences by age within five miles of PM$_{10}$ monitoring sites.
Fig. 7. Asthma incidences by month within five miles of PM$_{10}$ monitoring sites.
Fig. 8. Asthma incidences by the day of the week within five miles of PM$_{10}$ monitoring sites.