Spatial and Spatio-temporal Epidemiology 000 (2016) 1–11

Contents lists available at ScienceDirect

Spatial and Spatio-temporal Epidemiology

journal homepage: www.elsevier.com/locate/sste

Characterizing the spatial distribution of multiple pollutants and populations at risk in Atlanta, Georgia

John L. Pearce[a,*], Lance A. Waller[b], Stefanie E. Sarnat[c], Howard H. Chang[b], Mitch Klein[c], James A. Mulholland[d], Paige E. Tolbert[c]

[a] Department of Public Health Sciences, College of Medicine, Medical University of South Carolina, 135 Cannon Street, Charleston, SC 29422, United States
[b] Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA, United States
[c] Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, United States
[d] School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, United States

A R T I C L E I N F O

Article history:
Received 31 October 2015
Revised 16 February 2016
Accepted 23 February 2016
Available online xxx

Keywords:
Air pollution
Classification
Cluster analysis
Kohonen map
Geographic information systems (GIS)
Multipollutant

A B S T R A C T

Background: Exposure metrics that identify spatial contrasts in multipollutant air quality are needed to better understand multipollutant geographies and health effects from air pollution. Our aim is to improve understanding of: (1) long-term spatial distributions of multiple pollutants; and (2) demographic characteristics of populations residing within areas of differing air quality.
Methods: We obtained average concentrations for ten air pollutants ($p = 10$) across a 12 km grid ($n = 253$) covering Atlanta, Georgia for 2002–2008. We apply a self-organizing map (SOM) to our data to derive multipollutant patterns observed across our grid and classify locations under their most similar pattern (i.e, multipollutant spatial type (MST)). Finally, we geographically map classifications to delineate regions of similar multipollutant characteristics and characterize associated demographics.
Results: We found six MSTs well describe our data, with profiles highlighting a range of combinations, from locations experiencing generally clean air to locations experiencing conditions that were relatively dirty. Mapping MSTs highlighted that downtown areas were dominated by primary pollution and that suburban areas experienced relatively higher levels of secondary pollution. Demographics show the largest proportion of the overall population resided in downtown locations experiencing higher levels of primary pollution. Moreover, higher proportions of nonwhites and children in poverty reside in these areas when compared to suburban populations that resided in areas exhibiting relatively lower pollution.
Conclusion: Our approach reveals the nature and spatial distribution of differential pollutant combinations across urban environments and provides helpful insights for identifying spatial exposure and demographic contrasts for future health studies.

1. Introduction

Air quality within urban environments involves a mixture of gaseous and particulate concentrations that are affected by a variety of emission sources, local topographies, and meteorological conditions. As such, complex spatial

* Corresponding author.
E-mail address: pearcejo@musc.edu (J.L. Pearce).

http://dx.doi.org/10.1016/j.sste.2016.02.002
1877-5845/

patterning can occur in urban air quality making the variability of such phenomena difficult to characterize as different pollutants often exhibit differential spatial patterns (e.g., ozone vs. nitrogen dioxides). This is a concern for health scientists in the field of air pollution epidemiology who need to identify appropriate spatial contrasts in their exposure assessments of air pollution (Marshall et al., 2008; Hajat et al., 2013). Such challenges, in part, have led investigators performing chronic exposure studies to typically focus on one pollutant at a time (Hoek et al., 2013); however, it is well understood that intercorrelations among various pollutants can be problematic for statistical models designed to estimate individual pollutant risk (Tolbert et al., 2007; Jerrett et al., 2013). Therefore, investigations reporting associations between long-term exposure to air pollution and adverse health generally acknowledge that reported associations are likely the result of a pollutant mixture, not the sole effect of the proxy pollutant (Pope et al., 2004; Lee et al., 2009; Hoek et al., 2013).

In order to improve our understanding of the health effects of long-term exposure to multiple pollutants it is necessary to examine the entire mix of pollutants (Dominici et al., 2010; Vedal and Kaufman, 2011; Levy et al., 2014). However, expanding chronic exposure studies of air pollution to incorporate information on multiple pollutants is expected to be challenging for at least two reasons: (1) measuring/modeling the joint spatial distribution of multiple air pollutants is difficult (Jerrett et al., 2005; Marshall et al., 2008; Riley et al., 2014; Sororian et al., 2014), and (2) characterizing the spatial distribution of multipollutant exposure is complex (Oakes et al., 2014). To further complicate matters, different subgroups within the populations at risk (e.g., those with low socioeconomic status (SES)) may be more intensely exposed to air pollution than others, a situation that may confound estimated associations between air pollution and health (Laurent et al., 2007; Yanosky et al., 2008; Hajat et al., 2013).

Given such challenges, development of approaches that can be useful for investigating the health effects of complex multipollutant exposures are highly desired (Dominici et al., 2010). Recently, many techniques have been presented for characterizing multipollutant exposure (Oakes et al., 2014); however, very few have been applied in spatial settings (Molitor et al., 2011; Austin et al., 2013). Although limited, findings from these studies have noted significant spatial variation in multipollutant exposures within and across cities in the US. Therefore, it is clear more studies are needed to better understand spatial variation of complex exposures as well as heterogeneity in exposure to populations at risk.

In the present study, we use Atlanta, Georgia, as a case study to illustrate a methodological approach for characterizing long-term trends in population exposure to multiple pollutants. Atlanta's air quality issues are well known and several studies have documented associations with health outcomes including asthma, cardiorespiratory morbidity, and preterm births (Alhanti et al., 2015; Chang et al., 2015; Pearce et al., 2015; Winquist et al., 2015). Moreover, a novel set of spatially and temporally resolved multipollutant data is available for the region (Sororian et al., 2014) that will allow us to more closely examine air pollution exposure across a unique and diverse population (Pooley, 2015). Our general objective is to determine whether and to what extent long-term patterns in multipollutant combinations and populations at risk systematically map onto one another in the Atlanta region. We aim to achieve our objective by addressing the following questions of interest:

1. What types of long-term multipollutant combinations occur at locations within our study?
2. What is the spatial distribution of types of multipollutant combinations across our study region?
3. What demographics are associated with areas differentiated by types of multipollutant combinations?

In answering these questions we hope to improve future epidemiologic studies by increasing our understanding of: (1) the long-term geographic patterns of multipollutant air quality across our study region; and (2) the demographic makeup of populations residing in areas that experience distinct long-term multipollutant exposure.

2. Methods

The principal focus of our approach is to identify geographic locations in our study area with similar long-term multipollutant characteristics in order to better understand local, long-term population exposure to ambient multipollutant mixtures. This is achieved in four stages: (1) divide the study area into grid cells, within which it is assumed the spatial distribution of pollution is relatively homogeneous, (2) define a number of multipollutant spatial types that describe the nature of the pollutant attributes of the grid cells, (3) characterize multipollutant geographies by mapping grid assignments to multipollutant spatial types in the study area, and (4) describe the demographic characteristics of the populations residing in locations corresponding to areas defined by the multipollutant spatial types.

2.1. Multipollutant air quality data acquisition

Available data for this study included seven years (2002–2008) of spatially and temporally resolved air pollution concentrations at a twelve kilometer gridded spatial resolution for ten ambient air pollutants obtained for a 31,285 km^2 study area encompassing Atlanta, Georgia (Sororian et al., 2014). This area contained 253 grid cells (Fig. 1). In brief, data at each grid cell are daily concentration estimates obtained from calibrating gridded output from the Community Multi-scale Air Quality (CMAQ) model against measurements from monitoring sites in the study area – a.k.a. 'fusion' data (Sororian et al., 2014). Pollutants available included 1-h maximum carbon monoxide (CO) in ppm, 1-h maximum nitrogen dioxide (NO_2) and nitrous oxides (NO_x) in ppb, 8-h maximum ozone (O_3) in ppb, 1-h maximum sulfur dioxide (SO_2) in ppb, and five 24-h average $PM_{2.5}$ components in $\mu g/m^3$: elemental carbon (EC), organic carbon (OC), nitrate (NO_3), ammonium (NH_4), and sulfate (SO_4). See Table 1 for summary statistics of these data.