Environmental Technology & Innovation 3 (2015) 1–10

Contents lists available at ScienceDirect

Environmental Technology & Innovation

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

Evaluating methods for spatial mapping: Applications for estimating ozone concentrations across the contiguous United States

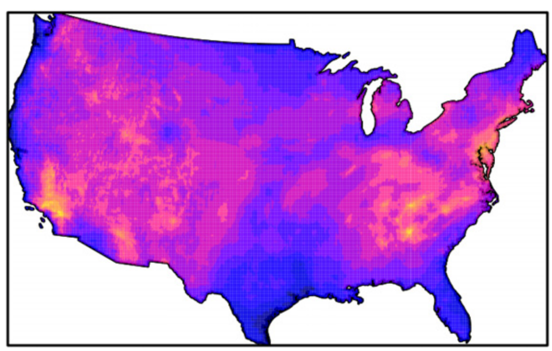
Jesse D. Berman [a,*], Patrick N. Breysse [b], Ronald H. White [c], Darryn W. Waugh [d], Frank C. Curriero [e,f]

[a] School of Forestry and Environmental Studies, Yale University, 301 Prospect Street, New Haven, CT 06511, United States
[b] Department of Environmental Health Sciences, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, United States
[c] R.H. White Consultants, LLC, Silver Spring, MD 20904, United States
[d] Department of Earth and Planetary Science, Johns Hopkins University, Baltimore, MD 21218, United States
[e] Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, United States
[f] Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, United States

HIGHLIGHTS

- Evaluated method performance for predicting and mapping national ozone pollution.
- Compared land use regression, IDW, ordinary and universal kriging for prediction.
- Land use regression models revealed the presence of residual spatial variation.
- Kriging outperformed the other approaches for predicting ozone concentrations.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 12 May 2014
Received in revised form
14 October 2014
Accepted 14 October 2014
Available online 20 November 2014

Keywords:
Ozone
Kriging
Mapping
Method comparison
Spatial prediction

ABSTRACT

Understanding spatial variability of air pollutant concentrations is critical for public health assessments. Our goal is to examine ground-level ozone and comparatively evaluate method performance for predicting and mapping national concentrations across the United States, while assessing the importance of accounting for spatial variability.

Cross-sectional US EPA ozone monitoring data was acquired for three days in 2006, plus environmental covariates of land use, traffic, temperature, elevation, and population. Evaluation of ozone variability was assessed with land use regression (LUR) and spatially explicit kriging models. Ozone concentration was predicted with four approaches, including LUR, inverse distance weighting (IDW), ordinary kriging, and universal kriging, and evaluated with a Monte Carlo cross-validation simulation. Results were mapped for the continental United States.

Temperature, elevation, and distance to major roads were significantly related to ozone concentrations and examination of spatial dependence on LUR models revealed the presence of residual spatial variation. Cross-validation results found kriging outperformed both LUR and IDW in terms of root mean squared prediction error. We demonstrate that national-level ozone is best evaluated using the statistically optimal kriging models, which account for residual spatial variation. Universal kriging was preferred over

* Corresponding author. Tel.: +1 347 628 2444; fax: +1 203 436 9135.
E-mail address: jesse.berman@yale.edu (J.D. Berman).

http://dx.doi.org/10.1016/j.eti.2014.10.003

ordinary kriging by allowing us to assess the significance of environmental covariates both for inference and prediction of ozone concentrations.

Abbreviations

AQS	Air Quality System
BLUP	Best linear unbiased predictor
CV	Cross validation
DEM	Digital elevation model
ESRI	Environmental Systems Research Institute
GLS	Generalized least squares
IDW	Inverse distance weighting
LUR	Land use regression
OK	Ordinary kriging
OLS	Ordinary least squares
ppb	parts per billion (by volume)
RMSE	Root mean squared error
UK	Universal kriging

1. Introduction

Characterizing exposures to air pollutants is critical for epidemiological studies. Ozone air pollution is linked to adverse health outcomes including respiratory related morbidity, such as decreased pulmonary function (Foster et al., 2000; Kim et al., 2011), asthma exacerbations (McDonnell, 1999), respiratory related hospital admissions (Burnett et al., 2001; Choi et al., 2011), and premature mortality (Bell et al., 2004; Jerrett et al., 2009). Ozone exposure presents population wide risks and in particular to susceptible groups, such as children, elderly, and individuals with pre-existing respiratory disease (Kim et al., 2011; Burnett et al., 2001; Bell and Dominici, 2008; Salam et al., 2005). The detection of small relative risks associated with individual exposure to air pollutants necessitate the need for population-level exposure modeling and makes exposure estimation a critical component of health effects studies (Bell et al., 2004; Levy et al., 2005; Berman et al., 2012).

Characterization of air pollution exposure can be a complex process, especially for population-based studies. Ambient concentrations from regulatory networks are common surrogates for individual exposure, (Bell et al., 2004; Levy et al., 2005; Pope et al., 2002; Samet et al., 2000) but monitors are often spatially heterogeneous with limited geographic coverage. Counties without air monitors tend to be rural, older, and have greater poverty levels, leading to increased vulnerability among their population (Bravo et al., 2012). To overcome these limitations, the identification of environmental determinants influencing air pollutants allows for improved inference about variability in pollutant concentrations. When predicting air pollutant exposures, the literature has found factors such as land use, population density, temperature, elevation, and traffic to all be significant in explaining concentrations.

Extrapolating data to unsampled locations (spatial prediction), allows us to create pollution maps of exposure for epidemiology applications, environmental health, and related policy research (Beelen et al., 2009). A popular approach for prediction and identifying environmental determinants of pollutant concentration is land use regression (LUR). First introduced by Briggs et al. (1997), it utilizes a geographic information system (GIS) to combine monitored air pollution data with land use and environmental variables for building regression covariates, and is increasingly popular in both European (Beelen et al., 2009; Briggs et al., 1997; Hoek et al., 2011; Freire et al., 2010) and North American studies (Clougherty et al., 2008; Henderson et al., 2007; Ross et al., 2007; Poplawski et al., 2008; Wilton et al., 2010; Ryan and LeMasters, 2007; Su et al., 2011, 2009). It has been argued that LUR results in better spatial predictions of small area variability compared to alternative approaches, notably at city-level geographies where LUR is most frequently (but not exclusively) applied (Ross et al., 2007; Brauer et al., 2003). Statistically the LUR model is equivalent to multivariate linear regression and trend surface modeling, both which assume a regression error structure of independence (Cressie, 1993; Schabenberger and Gotway, 2005). However, the performance of a statistical regression can be suboptimal in the presence of residual autocorrelation (Cressie, 1993; Schabenberger and Gotway, 2005; Gaffney et al., 2005; Jerrett et al., 2003). While this is appropriately handled through the introduction of spatial covariance functions (Freire et al., 2010; Szpiro et al., 2010; Franklin et al., 2012), it is not universally performed (Hoek et al., 2011; Poplawski et al., 2008; Ryan et al., 2007).

An alternative approach for spatial modeling is kriging. This regression-based method allows the inclusion of covariates either for directly quantifying effects and/or to improve spatial predictions. Kriging also automatically includes a spatially structured residual component to capture and account for spatial variation not explained by the regression covariates (Cressie, 1993; Schabenberger and Gotway, 2005). In a regression framework kriging is known to produce spatial predictions that are statistically optimal and best linear unbiased predictions (BLUPs) (Cressie, 1993). Kriging has been previously used as a tool to assess exposure to air pollutants, including ozone and particulate matter at both the regional city-scale (Künzli et al., 2005; Jerrett et al., 2005) and national-level (Beelen et al., 2009; Sampson et al., 2013).

Despite a variety of spatial methods used in environmental exposure studies, limited papers have comprehensively evaluated prediction approaches for exposure assessment settings (Bravo et al., 2012; Beelen et al., 2009; Marshall et al., 2008; Mercer et al., 2011; Son et al.,