Environmental Research 166 (2018) 487–496

Contents lists available at ScienceDirect

Environmental Research

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

Spatial modeling of daily concentrations of ground-level ozone in Montreal, Canada: A comparison of geostatistical approaches

Yuddy Ramos[a], Weeberb J. Requia[b,*], Benoît St-Onge[c], Jean-Pierre Blanchet[d], Yan Kestens[e,f], Audrey Smargiassi[g,h]

[a] Département de géographie, Université de Montréal, Montréal, Québec, Canada
[b] Department of Environmental Health, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, United States
[c] Département de géographie, Université du Québec à Montréal (UQAM), Montréal, Québec, Canada
[d] Département des sciences de la Terre et de l'atmosphère, Université du Québec à Montréal (UQAM), Montréal, Québec, Canada
[e] Département de médecine sociale et préventive, Université de Montréal, Montréal, Québec, Canada
[f] Centre de recherche du Centre hospitalier de l'Université de Montréal (CHUM), Québec, Canada
[g] Département de santé environnementale et de santé au travail, Université de Montréal Montréal, Québec, Canada
[h] Institut national de santé publique du Québec (INSPQ), Montréal, Québec, Canada

ARTICLE INFO

Keywords:
Air pollution
Ground-level ozone
Kriging with external drift
Geostatistics
Spatial analysis

ABSTRACT

Ground-level ozone (O_3) is a powerful oxidizing agent and a harmful pollutant affecting human health, forests and crops. Estimating O_3 exposure is a challenge because it exhibits complex spatiotemporal patterns. The aim in this study was to provide high-resolution maps (100 × 100 m) of O_3 for the metropolitan area of Montreal, Canada. We assessed the kriging with external drift (KED) model to estimate O_3 concentration by synoptic weather classes for 2010. We compared these results with ordinary kriging (OK), and a simple average of 12 monitoring stations. We also compared the estimates obtained for the 2010 summer with those from a Bayesian maximum entropy (BME) model reported in the literature (Adam-Poupart et al., 2014). The KED model with road and vegetation density as covariates showed good performance for all six synoptic classes (daily R^2 estimates ranging from 0.77 to 0.92 and RMSE from 2.79 to 3.37 ppb). For the summer of 2010, the model using KED demonstrated the best results (R^2 = 0.92; RMSE = 3.14 ppb), followed by the OK model (R^2 = 0.85, RMSE = 4 ppb). Our results showed that errors appear to be substantially reduced with the KED model. This may increase our capacity of linking O_3 levels to health problems by means of improved assessments of ambient exposures. However, future work integrating the temporal dependency in the data is needed to not overstate the performance of the KED model.

1. Introduction

Spatial variation of ground-level ozone (O_3) is strongly influenced by environmental factors, including meteorological conditions, land use, spatial distribution of the population, and the long-range transport (Moral et al., 2012; So and Wang 2003; Verstraeten et al., 2015). For example, high O_3 levels usually occur during the summer months when heat and sunlight are more intense. However, high levels of O_3 can also be observed at mid-latitude sites in the late winter and spring seasons (Ahmadov et al., 2015; Schnell et al., 2009). During the winter, high concentrations of O_3 can be caused by long-range transport (i.e., transport of air pollutants in the atmosphere for a distance greater than 100 km) and buildup of O_3 precursors (Diem, 2004). During the spring, high concentrations of O_3 can be caused by a strong generation of isoprene and terpenes from vegetation (Liudchik et al., 2013) and modulated by enhanced O_3 photochemistry with UV radiation. High O_3 concentrations also occur in non-industrial areas (even several hundred kilometers downwind from urban and industrial areas) and can be found in areas with low population density (Canada-United States Air Quality Committee, 1999). In urban areas, both anthropogenic and natural NOx and VOC are important precursors of O_3 formation, unlike in non-urban areas, where the biogenic VOC emitted from vegetation is the most important precursor of O_3 formation (U.S. EPA, 2013). The intra-urban variations of O_3 levels are also associated with the geographic variation of sources of O_3 precursors and sources of oxidizing compounds such as road traffic-related NOx (Liu and Rossini, 1996).

The assessment of O_3 exposure is a challenge because it exhibits complex spatiotemporal patterns (Adam-Poupart et al., 2014). Studies

* Corresponding author.
E-mail address: wjrequia@hsph.harvard.edu (W.J. Requia).

https://doi.org/10.1016/j.envres.2018.06.036
Received 6 April 2018; Received in revised form 5 June 2018; Accepted 17 June 2018

have suggested that these patterns may be explained by the interrelationship between the environmental factors and the source of emissions. For example, high levels of O_3 in Eastern Canada have been associated with clear sky anticyclonic conditions when high temperatures and west to south-west winds predominate. This facilitates the long-range transport of O_3 and its precursors (McKendry, 1993) from heavily industrial regions located along the shorelines of the Great Lakes (Canada-U.S. border) where higher levels of O_3 occur (Garcia et al., 2010) and favors photochemical oxidation rates.

To improve the exposure assessment of populations, different geostatistical techniques have been used in air pollution studies. Ordinary kriging (one version of kriging) and land-use regression (LUR) are the most common geostatistical approach used in air quality modeling (Jerrett et al., 2005). More complex spatiotemporal models such as Bayesian maximum entropy (BME) have also been used to estimate exposure over large areas (Adam-Poupart et al., 2014). These geostatistical models have specific differences and each one has inherent uncertainties due to the complexity of the atmospheric environment. For example, while kriging and BME are based on interpolation techniques to predict pollutant concentrations at unmonitored sites (Deligiorgi and Philippopoulos, 2011), LUR is based on regression models, where the spatial distribution of independent variables is used to estimate the values at unmonitored areas (Rivera et al., 2012; Ryan and LeMasters, 2007).

These studies have demonstrated that identifying the underlying geographic phenomena (e.g., spatiotemporal patterns) is useful to monitor conditions on the ground, to calculate spatiotemporal changes, to compare populations, and, especially, to communicate actionable data to potentially affected people as well as governmental agencies and policymakers (Adams and Kanaroglou, 2016; Bateman et al., 2013; Kurland and Gorr, 2012; Mitchell, 1999).

In this study, we assessed a multivariable kriging approach known as Kriging with external drift (KED) to predict daily O_3 concentrations at the local scale (100 m grid) in the metropolitan area of Montreal, Canada. These results were compared with those of Ordinary kriging (OK) and with an approach based on the simple mean of O_3 levels at fixed-site monitoring stations. We also compared summer estimates to those obtained from a BME model reported in the literature (Adam-Poupart et al., 2014).

2. Methods

2.1. Study area

The study was carried out in Montreal, Canada, which is located between the parallels of 45.27–45.72° N and 73.98–73.28° W. The area is composed of the following nine census consolidated subdivisions (CCS), representing the Montreal census metropolitan area (CMA): Montreal, Laval, Longueuil, Brossard, La Prairie, Saint-Philippe, Saint-Mathieu, Saint-Constant, and Kahnawake. Montreal has an area of 1309.5 km^2 and a population of 2.48 million people. The Montreal CCS includes the City of Montreal, which is the second largest city in Canada, and is located in the Saint Lawrence Valley at the eastern end of the Windsor-Quebec corridor (located along the north shore of the Great Lakes and Saint Lawrence River), the most industrialized region of Canada (McKendry, 1993) (Fig. 1).

In this study, the climate is categorized as humid continental, with severe winters and with hot and humid summers. It is subject to large continental climate variations. The wind direction is predominantly from the south-west, alternating with northeasterlies along the axis of the Saint Lawrence Valley.

2.2. Study design

This study was conducted over a period of one year, from January to December 2010. The OK, the KED model and the mean estimates were implemented for each of six synoptic weather classes to account for spatial dependencies on meteorological variables and to deal with temporal dependence. The OK and KED models were selected because a preliminary spatial analysis of data indicated that there exists a spatial autocorrelation. So we used interpolation approaches (univariable and multivariable) to consider spatial dependences. In order to compare the results of these models with those of a BME model (developed for the summer season), we also produced estimates for the months of May through September 2010. Table 1 shows the differences among the models used in our study in terms of the spatiotemporal aspects.

2.3. O_3 data

Hourly data of O_3 concentrations were obtained from the National Air Pollution Surveillance (NAPS) network of Environment Canada (Environment Canada, 2013d). O_3 data was measured by 12 NAPS monitors in the study area (Fig. 1), providing a spatially sparse sample (i.e., about 1 station per 100 km^2) of O_3 data with a good temporal coverage. These measurements are performed using gas analyzers operating on ultraviolet light absorption principles (Environment Canada, 2013a).

We modeled O_3 concentrations throughout the year considering two periods category: (i) summer – May through September; and ii) winter, spring and fall – January through April, and October through December. The synoptic class 2 corresponding to the winter season had 11 days that exceeded the criteria of 38 ppb for 8-h average and for the synoptic class 5 corresponding to the spring season had 42 days that exceeded the criteria (in Section 2.5, we present the details of how synoptic classes were defined). We used midday 8-hr averages because the largest values of the 24-h O_3 concentrations period for a given day are normally in this interval and because it corresponds to the Canada-wide standard for O_3 estimation (Conseil Canadien des Ministres de l'Environnement, 2000). O_3 concentration was calculated for the 12 monitors having no more than two hours of missing values on any given day. For this reason, we calculated midday 8-hr averages (09:00–17:00 h) of O_3 levels for 2010.

2.4. Meteorological data

Meteorological measurements were obtained from the National Climatic Data and Information Archive (DAI) of Environment Canada (Environment Canada, 2013b). There were 8 meteorological monitoring stations across the Montreal CMA for the calendar years from January 2008 to December 2010 (Fig. 1). To create spatiotemporal meteorological surfaces (for KED, see below), we calculated midday 8-hr averages for temperature and relative humidity (except for KED, which we describe further in Section 2.7). The same interval as for the O_3 measurements was used because these meteorological variables influence O_3 chemistry and concentration. Also, we used daily average precipitation since only daily total information was available. Precipitations are responsible for the reduction of O_3 concentrations by wet removal of soluble O_3 precursors (Hou et al., 2015). We considered daily average wind speed as representative for long-range transport associated with the distance traveled by the air mass within the 24-h period (Camalier et al., 2007). Hourly measures of the wind speed and directions were also used to create wind roses. Each wind rose shows the frequency and speed of wind blowing from each direction (hourly measures) in a particular distribution (i.e., a synoptic group) for 2010.

Same as we considered for O_3 concentration, meteorological data were calculated for all available monitoring data presenting no more than two hours of missing values on any given day.

2.5. Synoptic weather analysis

Weather patterns evolve in time (Hufty, 1982) influencing the frequency and intensity of air pollutant concentrations including O_3 (Ebi