Science of the Total Environment 485–486 (2014) 563–574

Contents lists available at ScienceDirect

Science of the Total Environment

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

Evaluation of observation-fused regional air quality model results for population air pollution exposure estimation

Gang Chen [a], Jingyi Li [a], Qi Ying [a,*], Seth Sherman [b], Neil Perkins [c], Rajeshwari Sundaram [c], Pauline Mendola [c]

[a] Zachry Department of Civil Engineering, Texas A&M University, College Station, TX 77843, United States
[b] The EMMES Corporation, Rockville, MD 20850, United States
[c] Division of Intramural Population Health Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Rockville, MD 20852, United States

HIGHLIGHTS

- Ten-year CMAQ simulations for population air pollution exposures of gases and PM
- Inverse-distance weighting is used to create observation-fused concentration fields.
- Differences in population exposure based on five estimation methods were studied.
- Population-weighted average should be used to account for spatial variability.
- Exposure based on raw CMAQ or observations alone might have significant biases.

ARTICLE INFO

Article history:
Received 14 November 2013
Received in revised form 17 February 2014
Accepted 22 March 2014
Available online 17 April 2014

Editor: Xuexi Tie

Keywords:
Community Multiscale Air Quality (CMAQ) model
Data fusing
Inverse distance weighting
Model performance
Exposure
Population weighted average

ABSTRACT

In this study, Community Multiscale Air Quality (CMAQ) model was applied to predict ambient gaseous and particulate concentrations during 2001 to 2010 in 15 hospital referral regions (HRRs) using a 36-km horizontal resolution domain. An inverse distance weighting based method was applied to produce exposure estimates based on observation-fused regional pollutant concentration fields using the differences between observations and predictions at grid cells where air quality monitors were located. Although the raw CMAQ model is capable of producing satisfying results for O_3 and $PM_{2.5}$ based on EPA guidelines, using the observation data fusing technique to correct CMAQ predictions leads to significant improvement of model performance for all gaseous and particulate pollutants. Regional average concentrations were calculated using five different methods: 1) inverse distance weighting of observation data alone, 2) raw CMAQ results, 3) observation-fused CMAQ results, 4) population-averaged raw CMAQ results and 5) population-averaged fused CMAQ results. It shows that while O_3 (as well as NO_x) monitoring networks in the HRRs are dense enough to provide consistent regional average exposure estimation based on monitoring data alone, $PM_{2.5}$ observation sites (as well as monitors for CO, SO_2, PM_{10} and $PM_{2.5}$ components) are usually sparse and the difference between the average concentrations estimated by the inverse distance interpolated observations, raw CMAQ and fused CMAQ results can be significantly different. Population-weighted average should be used to account for spatial variation in pollutant concentration and population density. Using raw CMAQ results or observations alone might lead to significant biases in health outcome analyses.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Investigations of the effects of air pollution on population health are dependent on the quality of available air pollutant exposure estimates (Bell et al., 2004; Laden et al., 2000). Traditionally, exposure levels are estimated based on measurements made at nearby air monitoring stations that are limited in space and time (Bell et al., 2007). However, many of the populations in air pollution epidemiology studies are located in areas without sufficient air quality monitoring activities. In addition to the lack of spatial coverage of standard air quality monitoring networks, air quality measurements are limited by the capability of the analytical instruments and may not be able to provide sufficient temporal resolution or detailed chemical composition information to support detailed health outcome analyses.

Three-dimensional chemical transport models (CTMs) can provide detailed gaseous and particulate matter (PM) concentrations and their source and chemical composition information at one-hour resolution over large areas. The complete spatial and temporal coverage of CTMs makes them an ideal tool to fill in the spatial and temporal gaps in the

* Corresponding author. Tel.: +1 979 845 9709; fax: +1 979 862 1542.
E-mail address: qying@civil.tamu.edu (Q. Ying).

http://dx.doi.org/10.1016/j.scitotenv.2014.03.107
0048-9697/© 2014 Elsevier B.V. All rights reserved.

exposure estimation solely based on air quality measurements at fixed monitors (Bell, 2006; Bravo et al., 2012). One of the challenges in applying CTM-predicted exposure in health outcome studies is to reduce the error in exposure classifications. Uncertainties in the meteorology and emission inputs, the underlying chemical mechanisms and numerical techniques, as well as spatial resolution (i.e. grid size) of the CTM model itself can all lead to errors in the predicted concentrations. Even though previous studies showed that the long-term performance of a three-dimensional regional air quality model for ozone and PM can generally meet the criteria recommended by the United States Environmental Protection Agency (US EPA), systematic biases do exist in the predicted concentrations which could lead to biases in the exposure estimations (Zhang et al., 2014).

Data fusing techniques have been proposed to improve exposure estimations by adjusting the raw CTM model predictions with the ambient observation data (Fuentes and Raftery, 2005; Sahu et al., 2010). However, the effectiveness of these data fusing techniques on predicted air pollutants has only been examined for a small number of species for relatively short simulation periods (such as SO_2 in Fuentes and Raftery, 2005). Many epidemiologic studies would benefit from long-term exposure inputs of multiple air pollutants such as carbon monoxide (CO), nitrogen oxides (NO_x), ozone (O_3), particulate matter with aerodynamic diameters less than 2.5 and 10 μm ($PM_{2.5}$ and PM_{10}), as well as a number of air toxics and PM components to address the public health implications of poor air quality in a more comprehensive manner. The number of available stations for these different species and their spatial and temporal coverage in a given area can be significantly different. No study to date has examined the effectiveness of data fusing techniques on multiple species in the same air quality domain over a long study period.

Recently, an air quality modeling project was carried out to provide air pollutant exposure estimation over a ten-year period (from 2001 to 2010) for the Air Quality and Reproductive Health (AQRH) study with support from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD). The AQRH study is designed to analyze data collected during the Consortium on Safe Labor (CSL) study (Zhang et al., 2010) and the Longitudinal Investigation of Fertility and the Environment (LIFE) study (Buck Louis et al., 2011) to provide improved understanding of the relationship between air quality and various measures of reproductive health. The CSL study collected detailed electronic medical record information on 228,562 deliveries including both mother and baby records from 12 clinical centers comprising 19 hospitals with 15 non-overlapping HRRs across the United States between 2002 and 2008. The LIFE study was a longitudinal study (2005–2009) that closely followed 501 couples trying to conceive for up to 12 cycles or through pregnancy for those who became pregnant.

The objectives of this paper are to: 1) evaluate the model performance of the three-dimensional regional air quality model, 2) evaluate the effectiveness of observation data fusing in improving regional air quality model predictions and 3) compare the differences in the exposure estimations using raw and observation-fused air quality modeling results. In this study, modeled concentrations in the surface layer are used to represent population exposure to air pollution. Exposure assessment in health effect studies typically include time-activity measures and either active monitoring or some assessment of indoor air as well as ambient measures. These methods can definitely increase the accuracy and precision of an individual's exposure but typically the studies rely on volunteers who are willing to provide detailed information on their activity/mobility and only include subjects who accept the additional burden of personal monitoring. Our analyses reflect the ambient exposures of an entire population. While there is some error in exposure estimation at the ambient level, it is balanced with the inclusion of the full population. It may also be helpful to note that regulations are made at the ambient exposure level. Regardless of the individual level exposure, if an effect is observable at the population level, it is actionable.

2. Methods

2.1. The Community Multiscale Air Quality (CMAQ) Model

Among the many publicly available CTMs, the Community Multiscale Air Quality (CMAQ) model (Byun and Schere, 2006) is one of the most widely used regional air quality modeling systems in the United States in recent years (Simon et al., 2012). The CMAQ model has been deployed to evaluate air pollution control measures, test new atmospheric mechanisms and processes that control air pollution, and determine source contributions to air pollutants. The CMAQ model has also been used in a few recent studies (Arunachalam et al., 2011; Chang et al., 2012; Grabow et al., 2012; Tong et al., 2009) to estimate air pollution exposure. In this study, a recent version of the CMAQ model (version 4.7.1) with the SAPRC-99 photochemical mechanism (Carter, 2000) and the fifth generation aerosol model (AERO5) (Foley et al., 2010) was used. The SAPRC-99 mechanism was modified to treat a number of explicit air toxic pollutants and 16 gas phase polycyclic aromatic hydrocarbon (PAH) species. Details of this extended SAPRC-99 mechanism for air toxics and PAHs will be described in a separate manuscript as the current paper focuses on criteria pollutants.

2.2. Inverse Distance Weighting and Observation Data Fusing

Inverse distance weighting (Shepard, 1968) has been used to estimate air pollutant exposure based on monitoring data alone. For a given spatial location, a search of the air monitor list is performed to find stations with available data within a given search radius. The search radius is applied based on the assumption that stations outside the radius have minimal impact on the estimated exposure. Once the stations within a given search radius are identified, the air pollutant concentration based on inverse distance weighting C^{idw} is estimated by Eqs. (1) and (2):

$$C^{idw}(\mathbf{x}) = \sum_{i=1}^{N} \frac{w_i(\mathbf{x})C_i}{\sum_{j=1}^{N} w_i(\mathbf{x})} \tag{1}$$

$$w_i(\mathbf{x}) = d(\mathbf{x}, \mathbf{x}_i)^{-1} \tag{2}$$

where $\mathbf{x}$ is the location where concentration needs to be estimated; N is the total number of monitors within the search radius; C_i is the measured concentration at the ith monitor within the search radius; $\mathbf{x}_i$ is the location of the ith monitor; and $w_i(\mathbf{x})$ is the weighting factor for the ith monitor at location $\mathbf{x}$. The function d calculates the distance between points $\mathbf{x}$ and $\mathbf{x}_i$. Inverse distance weighting works best for estimating exposure where air quality monitors are nearby. More advanced spatial interpolation techniques based on geostatistical methods such as kriging (Bogaert and Fasbender, 2007) have also been attempted for interpolating air monitor data. However, the interpolated fields using kriging are often too smooth and are not likely to represent actual spatial variation of air pollutants. In fact, it is impossible to get very reliable estimates of pollutant concentrations at locations without nearby monitor sites, regardless whether inverse distance weighting or kriging is used for the interpolation (Bailey and Gatrell, 1995).

The CMAQ model can generally represent the spatial variation of the pollutants but the magnitude of the predicted concentrations is subject to biases in the emissions, meteorology and uncertainties due to other model components. In this study, a data fusing technique based on inverse distance weighting is developed to adjust gridded raw CMAQ predictions of major criteria pollutants and components of $PM_{2.5}$ using the observations at nearby air monitors and the raw CMAQ results, as described below.

Download English Version:

https://daneshyari.com/en/article/6330487

Download Persian Version:

https://daneshyari.com/article/6330487

Daneshyari.com