Atmospheric Environment 44 (2010) 3926–3934

Contents lists available at ScienceDirect

Atmospheric Environment

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

Optimizing emission inventory for chemical transport models by using genetic algorithm

M.J. Li, D.S. Chen*, S.Y. Cheng**, F. Wang, Y. Li, Y. Zhou, J.L. Lang

College of Environmental & Energy Engineering, Beijing University of Technology, Pingleyuan 100#, Beijing 100124, China

ARTICLE INFO

Article history:
Received 20 January 2010
Received in revised form
6 July 2010
Accepted 8 July 2010

Keywords:
Genetic algorithm
Emission inventory
Chemical transport models
CMAQ

ABSTRACT

Air pollutant emission inventory is an important input parameter for chemical transport models (CTMs). Since great uncertainties exist in the emission inventory, further improvements and refinements are required. In this paper, genetic algorithm (GA), a global search and optimization method, was applied to optimize the emission inventory for the Models-3/Community Multiscale Air Quality (CMAQ) model. An emission optimizing system based on GA was developed and embedded to the CMAQ through the design of several core modules, which implemented the basic functions such as emission adjusting, GA population initializing, CMAQ results evaluating and GA operating. Hypothetical and real-data experiments were respectively performed to examine the validity of GA for emission calibrating. GA showed good performance in both experiments and was always able to find the global minimum. The emission optimizing system was then used to calibrate seasonal PM_{10} emission inventories of Beijing. Results revealed that PM_{10} emission in Beijing was underestimated in 2002, an average of 62.74% higher adjustment factor should be imposed on the original emission in target months of different seasons. With the calibrated emission inventories, CMAQ model errors were decreased by 6.46% on average in different seasons. It was concluded that GA was a promising search technique in calibrating emission inputs for CTMs.

1. Introduction

An accurate emission inventory of chemical species plays an important role in chemical transport models (CTMs) because quantitative estimates of emissions are an important input parameter for CTMs, and their uncertainty largely determines the accuracy of the model outputs (Yumimoto and Uno, 2006). Emission inventories have usually been developed by the so-called bottom-up approach, which is based on combinations of activity statistics (such as energy consumption and industrial production) and source-specific or region-specific emission factors (Hao et al., 2002; Streets et al., 2003). In bottom-up emission inventories there are large uncertainties associated with the statistics, emission factors, temporal allocation profiles, and grid allocation factors. Thus these emission inventories require further improvements and refinements using information of observation data and output of numerical models.

Generally, there are mainly two approaches to optimize emission inventory of CTMs. One is the ensemble Kalman filter (EnKF) data assimilation and the other is the four-dimensional variational (4D-Var) data assimilation method. Both of the two approaches have been proved to be useful in various applications. Van Loon et al. (2000) applied EnKF in chemical data assimilation to recover ozone and emissions. His work indicated that it is possible to apply EnKF to atmospheric chemical transport models. Gilliland et al. (2003) used a Kalman filter inverse modeling technique to deduce NH_3 emissions inventories for the eastern United States in 1990. Napelenok et al. (2008) evaluated spatially-resolved NOx emissions using Kalman filter and a controlled scenario with a known solution was used to prove that the methodology can achieve the correct solution.

The 4D-Var method has been widely applied in inverse modeling of CTMs. Elbern and Schmidt (1999) applied 4D-Var to EURAD-CTM2 modeling system and verified 4D-Var implementation by identical twin experiments. Elbern et al. (2001) and Elbern and Schmidt (2001) estimated NOx and VOC emissions from ozone observations, and optimized ozone initial conditions over central Europe using 4D-Var method. Engelen et al. (2004) and Engelen and McNally (2005) estimated atmospheric CO_2 from AIRS radiance data within the ECMWF 4D-Var data assimilation system. Muller and

* Corresponding author. Tel.: +86 10 67396628; fax: +86 10 67391983.
** Corresponding author. Tel.: +86 10 67391656; fax: +86 10 67391983.
E-mail addresses: dschen@bjut.edu.cn (D.S. Chen), chengsy@bjut.edu.cn (S.Y. Cheng).

1352-2310/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.atmosenv.2010.07.010

Stavrakou (2005) estimated CO and NOx emissions using the adjoint version of a global IMAGE model from ground-based and GOME satellite observations. Hakami et al. (2005) estimated black-carbon emissions over eastern Asia using the adjoint STEM-2K1 model. Yumimoto and Uno (2006) applied a 4D-Var data assimilation system to inverse modeling of CO emissions in the eastern Asia. Kurokawa et al. (2009) developed an adjoint inverse modeling system to optimize NOx emissions over eastern China using satellite observations of NO_2 vertical column densities (VCDs) and a priori emissions from the Regional Emission Inventory in Asia (REAS).

As research continues, the merits and flaws of the two approaches were discussed. The adjoint model and gradient of the objective function are required in 4D-Var, but in some complicated models like CMAQ, it is not easy to obtain them. Ensemble Kalman filter (EnKF) has a main limitation (Van Loon et al., 2000) that it assumes model error distribution to be Gaussian, but actually, model errors will not be Gaussian in some specific cases. It is beyond the scope of this study to discuss the advantages and drawbacks of the different methods. A detailed comparison of the relative merits of EnKF and 4D-Var can be found in Lorenc (2003) and Kalnay et al. (2007).

The genetic algorithm (GA), which solves the optimization problem in fundamentally new ways, may be another potential way to optimize emission inventory for CTMs. Genetic algorithm is a search procedure based on the mechanics of natural selection and natural genetics, combining an artificial survival of the fittest with genetic operators abstracted from nature (Holland, 1975; Goldberg, 1989). Based on multiple search directions, GA is easier to escape from a local minimum. As a global optimization method, it has been proved to be useful in parameter estimation of water environmental modeling. Wang (1997) introduced genetic algorithm to calibrate a conceptual rainfall-runoff model with nine parameters using both hypothetical data and real data. Liu et al. (2007) used genetic algorithm to carry out the model calibration in a water quality model. They analyzed the impact of operators on its effectiveness in optimum searching in the GA. However, until now, genetic algorithm has scarcely been used in optimizing the emission inventory for CTMs.

In the present study, a GA optimizing system was developed to optimize air pollutant emission inventory for the Models-3/Community Multiscale Air Quality (CMAQ) model. We presented the establishment of the air pollutant emission inventories in Beijing and introduced the design of the emission optimizing system. This system is tested in both idealized and real data settings, and then applied in optimizing the seasonal emission inventory of Beijing. Finally, PM_{10} concentrations were simulated using the CMAQ model with the calibrated emission inventory, and modeling results were compared with that using the original one.

2. Background

2.1. Overview of the study area

Beijing, the capital city of China, is situated at 40 degrees north latitude and 116 degrees of longitudes (39°08′-41°05′N, 115°25′-117°30′E), with Tianjin on its eastern border and Hebei Province on the other three sides. It covers an area of 16,800 square kilometers, with 62% of the total area being mountainous and other 38% being the plain area. With the rapid economic development, problems of environment have become serious in Beijing. High coal consumption, growing use of vehicles, lacking of dust pollution control, together with its basin-like topography and adverse weather conditions are considered as the main factors leading to Beijing's PM_{10} pollution.

Among all air pollutants, PM_{10} is considered to be the primary pollutant for Beijing (Chen et al., 2007a,b; Wang et al., 2010). Based on the 24 hour average monitoring data published by Beijing Environmental Protection Bureau (BJEPB), we can find that the primary pollutant for Beijing is almost always PM_{10} (http://www.bjepb.gov.cn/). A precise PM_{10} emission inventory is urgently required for air quality simulation and environmental management in Beijing.

2.2. The original emission inventory in Beijing

Emission inventories, lateral boundary conditions and initial conditions are considered to be main sources of uncertainties in atmospheric CTMs (Constantinescu et al., 2007) in addition to meteorological fields. In China, uncertainties in the emission inventory were regarded as the primary source of model errors in air quality forecasts (Xu et al., 2008). The erroneous emission inventory was therefore the major concern of this paper. In our research, efforts have been made to reduce the effects of meteorological fields, boundary conditions and initial conditions. So we assumed the model and observation discrepancies were mainly due to the emission inventories.

In our study, a relatively detailed emission inventory of air pollutants in Beijing was established. The base-year of this emission inventory is 2002. It was developed by the bottom-up approach thus there were large uncertainties in the emission inventory, especially in PM_{10} inventory. The emission data was obtained from the environmental protection administrations of Beijing. The collected emission data in Beijing is available at a municipal level, which includes 16 municipalities. The major pollution sources included in the inventory are industrial, residential, transportation, agriculture and the fugitive emissions. The raw information in the inventory includes 7 species: sulfur dioxide (SO_2), nitrogen oxides (NOx), total suspended particulates (TSP), particulate matter with an aerodynamic diameter less than 10 microns (PM_{10}), carbon monoxide (CO), ammonia (NH_3), and volatile organic compounds (VOCs).

Given close analysis, we found that fugitive PM_{10} emissions (including fugitive industrial emissions, road dust emissions, construction emissions, bare soil emissions, stockpile emissions) account for the major part of the PM_{10} emissions (75%). Among the total fugitive PM_{10} emissions, road dust emission sources, construction emission sources, fugitive industrial emission sources account for 54%, 24% and 15%, respectively. In practical work, fugitive PM_{10} emissions are often hard to measure. Emissions of air pollutants from fugitive emission sources are calculated using empirical models, which are essentially regressions of measured emissions on variables suspected to be important in determining emissions (Venkatram, 2000). Empirical formulas in the AP-42 document published by the USEPA (AP-42, fifth edition) was applied in our work to calculate fugitive emissions in Beijing. Considering the differences of soil types between China and US, some parameters of the formulas have been modified. For example, field experiments were conducted to get some parameters such as the road surface silt loading; some parameters like soil wind erosion index were localized referring to reports by Beijing municipal research institute of environmental protection (2003). For lacking of enough mechanistic bases, the empirical method to obtain the fugitive emissions brings large uncertainties in the PM_{10} emission inventory.

3. Methodology

3.1. MM5-CMAQ model description

The NCAR/Penn State Mesoscale Model (MM5) (Grell et al., 1994) was used to generate the meteorological fields for CMAQ (Byun and Ching, 1999). The CMAQ model is a sophisticated