Chapter 101
Spatial and Temporal Extension of a Novel Hybrid Source Apportionment Model

Cesunica Ivey, Heather Holmes, Yongtao Hu, James A. Mulholland, and Armistead G. Russell

Abstract Exposure assessment and development of control strategies are limited by the air pollutants measured and the spatial and temporal resolution of the observations. Air quality modeling can provide more comprehensive estimates of the temporal and spatial variation of pollutant concentrations, however with significant uncertainties. Source apportionment, which can be conducted as part of the air quality modeling, provides estimates of the impacts of sources on the mixtures of pollutants and contains surrogate estimates for pollutants that are not measured. This study details results using a novel spatiotemporal hybrid source apportionment method employed with interpolation techniques to quantify the impact of 33 $PM_{2.5}$ source categories. The hybrid model, which aims to reduce estimating uncertainties, adjusts original source impact estimates from a chemical transport model at monitoring sites to closely reflect observed ambient concentrations of measured $PM_{2.5}$ species. Daily source impacts are calculated for the contiguous U.S. Two interpolation methods are used to generate the data needed for spatiotemporal hybrid source apportionment. Hybrid adjustment factors are spatially interpolated using kriging, and daily observations are calculated by temporally interpolating available monitoring data. Methods are evaluated by comparing daily simulated concentrations—generated by reconstruction of source impact results—to observed species concentrations from monitors independent of model development. Results also elucidate U.S. regions with relatively higher impacts from specific sources. Monitoring data in this study originated from the Chemical Speciation Network (CSN), EPA-funded supersites, and the Southeastern Aerosol Research Characterization (SEARCH) Network. Results are to be used in health impact assessments.

C. Ivey (✉) • H. Holmes • Y. Hu • J.A. Mulholland • A.G. Russell
School of Civil and Environmental Engineering, Georgia Institute of Technology,
Atlanta, GA, USA
e-mail: civey3@gatech.edu

D. Steyn and R. Mathur (eds.), *Air Pollution Modeling and its Application XXIII*,
Springer Proceedings in Complexity, DOI 10.1007/978-3-319-04379-1_101,

101.1 Introduction

Two types of models are commonly used for source apportionment (e.g. determining impacts of sources on ambient concentrations): receptor models (RM) and chemical transport models (CTM). RMs are statistical models that use observed concentrations and source profiles to determine source impacts. CTMs are source-oriented models that determine impacts by simulating the formation, fate, and transport of pollutants in the atmosphere. Inherent in both approaches are uncertainties and limitations. RMs do not incorporate complex physical and chemical processes into results and are limited by the number of sources that can be resolved. CTMs do not make use of observations and have uncertainties associated with model inputs and modeled processes. This work builds from a previous study on integrating both methods to strengthen source apportionment estimates. This hybrid approach adjusts CTM source impacts to better reflect observations using constrained nonlinear-optimization with effective variance weighting [1]. This work details the application of the hybrid approach along with spatial and temporal interpolation methods to provide daily spatial hybrid source impacts for CONUS for the month of January 2004 (first and last 3 days omitted). The results are beneficial to trans-disciplinary studies that require spatially and temporally dense air quality data (e.g. epidemiological studies that correlate pollutant concentrations and health outcomes) [2, 3].

101.2 Methods

Generating daily, hybrid-kriging source impact spatial fields involves five steps: (1) calculating initial source impact estimates at 36-km resolution for CONUS using CMAQ-DDM3D, (2) generating daily sets of speciated data at monitoring sites over CONUS by temporally interpolating measured concentrations, (3) applying the hybrid source apportionment model for each monitor on each day, (4) spatially interpolating daily sets of hybrid adjustment factors, and (5) finally adjusting original spatial fields of source impacts by applying spatial fields of adjustment factors. Details for each step are discussed below.

The CMAQ-DDM3D model [4] was used to determine source impacts for 41 $PM_{2.5}$ species including total mass at a 36-km resolution. The model apportioned mass to 33 unique source categories, including on- and off-road gasoline and diesel sources, seven biomass burning sources, sea salt, and sources impacted by secondary processes (e.g. biogenic and livestock sources). The hybrid model requires observation data for the monitor and day being analyzed. Observation data from the Chemical Speciation Network (CSN) were used for hybrid-model development. However, CSN data are limited temporally, reporting every 3 or 6 days, hence temporal interpolation is required to generate daily data. Only monitors with speciated data as well as daily Federal Reference Method (FRM) measurements for total mass were used for the temporal interpolation, totaling 55 monitors. Note that concentrations on non-interpolated days were used from all available monitors for hybrid analysis, regardless of daily FRM availability. Observation uncertainties associated with each species concentration are also interpolated using a propagation of error approach. The hybrid method is then applied using interpolated observations for all monitors and for each day in Jan 2004, generating 33 adjustment factors (R's) for each monitor and for each day (Eq. 101.1).

$$X^2 = \sum_{i=1}^{N} \left[\frac{\left[\left(c_i^{obs} - c_i^{sim} \right) - \sum_{j=1}^{J} SA_{i,j}^{base} \left(R_j - 1 \right) \Big) \right]^2}{{\sigma_{obs}}^2 + {\sigma_{sp}}^2} \right] + \Gamma \sum_{j=1}^{J} \frac{ln\left(R_j\right)^2}{{\sigma_{\ln\left(R_j\right)}}^2} \tag{101.1}$$

In Eq. 101.1, c_i^{obs} and c_i^{sim} are observed and CMAQ-simulated species (i) concentrations, respectively, $SA_{i,j}^{base}$ are original CMAQ estimates of source j's impact one species i's concentration, R is the adjustment factor for each source (j), and σ are uncertainties in measurements (σ_{obs}), CMAQ-simulated concentrations (σ_{SP}), and emissions estimates ($\sigma_{\ln(rj)}$). Equation 101.1 is optimized to find the R's that minimize differences in observed concentration and simulated source impacts. Each daily set of R's are spatially interpolated to generate spatial fields of adjustment factors at 36-km resolution. The spatial fields of adjustment factors are then applied to the original CMAQ source impact fields (grid-cell by grid-cell multiplication) to produce hybrid-kriging spatial fields for Jan 2004. Additional data from the Jefferson Street monitor (Atlanta, GA) of the Southeastern Aerosol Research and Characterization (SEARCH) Network were used to evaluate the hybrid spatiotemporal results [5].

101.3 Results

Daily hybrid-kriging spatial fields of source impacts were generated for 33 sources for the month of January 2004 (Fig. 101.1). Source impact fields with significant adjustments include biomass burning (agricultural burning, lawn waste burning, open fires, prescribed burning, wildfires, wood fuel burning, and woodstoves) and dust impact fields. On average for Jan. 2004, for CMAQ-simulated biomass burning spatial fields were reduced by a factor of 2.7. CMAQ-simulated dust spatial fields were reduced by a factor of 5.0. Source impact fields saw fewer changes for stationary diesel sources, non-road sources, and sea salt.

Hybrid-kriging was evaluated with an independent data set by comparing reconstructed concentrations to observations at the Jefferson Street monitor in Atlanta, GA, which reported daily speciated data for January 2004 (Fig. 101.2). The reconstructed, hybrid concentrations show a similar trend as the observations for total $PM_{2.5}$. Initially over-simulated by CMAQ, potassium, silicon, and calcium

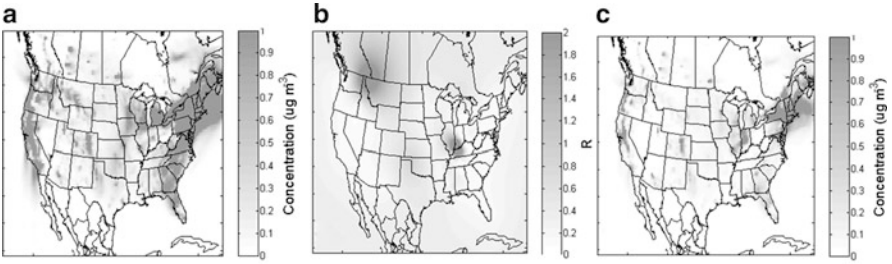

Fig. 101.1 Hybrid adjustment of 36-km CMAQ woodstove PM2.5source impacts on Jan 4, 2004. (**a**) Original CMAQ simulations. (**b**) Spatial field of kriged hybrid adjustment factors (Rwoodstove). (**c**) Hybrid-kriging source impacts

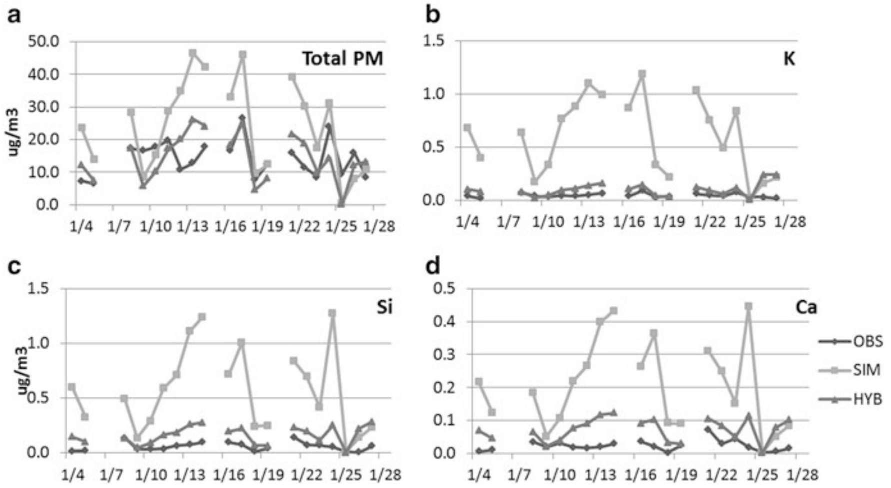

Fig. 101.2 Comparison of observed, CMAQ, and hybrid-kriging concentrations at the Jefferson Street monitor in Atlanta, GA. (**a**) Total PM2.5, (**b**) potassium (K), (**c**) silicon (Si), and (**d**) calcium (Ca)

concentrations were reduced by hybrid adjustment and became more aligned with observations. Potassium is a tracer for biomass burning, and silicon and calcium are tracers for dust. Reducing the impact of these sources greatly improved CMAQ simulations of the tracer species. The monthly average for potassium was reduced from 0.57 to 0.16 $\mu g/m^3$, which is closer to the average observation of 0.06 $\mu g/m^3$. The monthly average for silicon was reduced from 0.60 to 0.10 $\mu g/m^3$ (average observation 0.04 $\mu g/m^3$). The monthly average for calcium was reduced from 0.21 to 0.07 $\mu g/m^3$ (average observation 0.02 $\mu g/m^3$). Traditionally, biomass burning and dust emissions estimates have a high uncertainty [6, 7]. The hybrid source apportionment method takes into account emissions uncertainties, as well as observations, to improve source impact estimates. The spatial and temporal extensions of the hybrid-kriging method provide improved daily, spatially dense source impacts for used in health studies.

Acknowledgments This work was made possible in part by USEPA STAR grants R833626, R833866, R834799 and RD83479901 and by NASA project SV6-76007 under grant NNG04GE15G. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the USEPA and NASA. Further, neither USEPA nor NASA endorses the purchase of any commercial products or services mentioned in the publication. We also acknowledge the Southern Company for their support and thank Eric Edgerton of ARA, Inc. for access to the SEARCH data.

References

1. Watson JG, Cooper JA, Huntzicker JJ (1984) The effective variance weighting for least-squares calculations applied to the mass balance receptor model. Atmos Environ 18(7):1347–1355
2. Thurston G et al (2005) Results and implications of the workshop on the source apportionment of PM health effects. Epidemiology 16(5):S134–S135
3. Darrow LA et al (2011) "Ambient Air pollution and birth weight in full-term infants in Atlanta", 1994-2004. Environ Health Perspect 119(5):731–737
4. Napelenok SL et al (2006) Decoupled direct 3D sensitivity analysis for particulate matter (DDM-3D/PM). Atmos Environ 40(32):6112–6121
5. Hansen DA et al (2003) The southeastern aerosol research and characterization study: 1-overview. J Air Waste Manage Assoc 53(12):1460–1471
6. Choi YJ, Fernando HJS (2008) Implementation of a windblown dust parameterization into MODELS-3/CMAQ: application to episodic PM events in the US/Mexico border. Atmos Environ 42(24):6039–6046
7. Tian D et al (2009) Assessment of biomass burning emissions and their impacts on urban and regional PM2.5: a Georgia case study. Environ Sci Technol 43(2):299–305