H13-53

UNCERTAINTY IN AIR QUALITY FORECASTS CAUSED BY EMISSION UNCERTAINTY

Piotr Holnicki [1], Zbigniew Nahorski [1], Bartłomiej Solarz-Niesłuchowski [1], Marko Tainio [1,2]

[1] Systems Research Institute of the Polish Academy of Science, 01-447 Warsaw, Newelska 6, Poland
[2] National Institute for Health and Welfare (THL), P.O. Box 95, FI-70701 Kuopio, Finland

Abstract: An important application of air pollution models is to support decisions concerning air quality management and emission control. For example, to quantify possible ecological or health benefits of emission abatement, there is a need to estimate the incremental contribution of the respective group of emission sources to ambient concentrations with reasonable accuracy. However, due to significant complexity of such forecasting systems, there exist many sources of imprecision or uncertainty in the modeling of environmental effects of atmospheric pollution (e.g. model conceptual simplifications, model parameters, input data). This paper addresses the problem of uncertainty of emission inventory and impact of this uncertainty on the ambient air pollution concentrations and adverse health effects. The computational experiment implemented for Warsaw Metropolitan Area, Poland, encompasses a one-year forecast with the year 2005 meteorological dataset. The full emission inventory is composed of four categories of sources, characteristic for urban area: a) large point sources, b) intermediate point sources, c) area sources (residential sector), d) linear sources (transportation system). The CALPUFF air dispersion model was used as the main forecasting tool, combined with Monte Carlo statistical techniques to propagate uncertainty of the emission data. Based on the forecasted concentrations of air pollution, sensitivity of the adverse health effect estimates to emission uncertainties is evaluated.

Key words: *Air pollution modeling, particulate matter, uncertainty analysis, adverse health effects.*

INTRODUCTION

Air quality forecasting models and the more complex, integrated assessment systems (IAM) are recently used for supporting decisions concerning air quality management and emission control policy (Warren R.F. and H.M. ApSimon, 1999). The operational models are applied for analysis of air pollution mitigation policies, for example, to indicate where the required air quality limits will be exceeded, and what emission mitigation strategy should be applied to reach the prescribed standards. However, due to the very complex, multidisciplinary structure of such systems, there exist many sources of imprecision and uncertainty in the modeling of environmental effects of atmospheric pollution and also in the resulting regulatory decisions. To assess accuracy of modeling results and a connected decision support process, performance and uncertainty of the model should be evaluated. The most common way of examining the relative agreement between volume-averaged simulations and point measurements is not sufficient because of different spatial scales of these two values. To better characterize the problem, the main sources of variability (temporal, spatial, or inter-individual differences of input data) and uncertainty (lack of information about unknown quantity or imprecise information) should be identified and assessed (Park S.-K. *et al.*, 2006; Sax T. and V. Isakov, 2003; Zimmermann H.-J., 1999). In addition, implementations of operational models of air pollution transport usually involve some specific simplifications or parameterizations and cannot completely characterize complex physical processes, which is the source of conceptual uncertainty also reflected in the final results. Previous studies have revealed (Russel A. and D. Dennis, 2000) that major uncertainties (measurement or estimation error) in air quality models are due to meteorological data and emission inventory.

The problem related to urban air pollution is high in the priorities of environmental concern. Numerous studies of model outputs and measurement data have shown that most of air pollution models poorly describe both temporal and spatial dependencies of pollutant concentrations (ApSimon H.M. *et al.*, 2002; Sportisse B., 2007). Estimation of the urban-scale pollution is a computationally sophisticated modeling problem due to complexity of emission field, but also, due to complicated building orography and wind-field effects. Emission inventory of urban areas usually encompasses different categories of emission sources, and namely: (I) point-wise sources of power or heating plants, (II) other industrial sources, (III) area-type sources of residential sectors, and (IV) emission of urban transportation system. Each kind of source is characterized by specific composition of emitted compounds which can cause numerous adverse environmental effects. Varieties of primary pollutants generate secondary compounds, by means of chemical transformation processes, which may be even more dangerous for the environment. On the other hand, due to high population density, urban air pollution exposure is a crucial factor associated with numerous adverse health effects. In particular, many research results indicate that a considerable harm in public health is caused by particulate matter air pollution, especially PM2.5.

It is known that official emission data are not accurate, due to inventory uncertainties connected with some categories of urban emissions. Emissions of major power plants of energy sector can be treated as relatively accurate because of well specified parameters of combustion process as well as of the fuel used. On the other hand, emission data that characterize residential area or transportation systems are usually based on some aggregated and averaged information related to the fuel consumption and parameters. These categories of emission data do not reflect either the real temporal variability or chemical constitution of polluting compounds and are remarkably uncertain. In complex uncertainty analysis, correlation between some pollutants emitted by a source (Page T. *et al.*, 2003) should also be taken into account.

This paper presents results of uncertainty and sensitivity analysis applied to air quality modeling implemented for Warsaw Metropolitan Area, Poland. The experiment is focused on investigation of urban emission uncertainty, where the main forecasting tool, which links the emission field to the resulting annual average concentrations, is CALPUFF model. Uncertainty analysis is performed using Monte Carlo algorithm (Hanna S. *et al.*, 1998; Moore G.E and R.J. Londergan, 2001) The resulting values of fine particulate matter (PM2.5) concentrations, combined with the population density data, were used to estimated premature mortality due to fine particulate matter air pollution following the methods presented in (Tainio M., 2009).

COMPUTATIONAL EXPERIMENT

The computational test implemented for Warsaw Metropolitan Area, Poland, encompasses one-year forecast with the year 2005 meteorological and emission dataset. Emission field inventory is composed of four categories of sources: a) high point sources of energy sector (16 heating and/or power plants); b) other point sources (1017 industrial sources); c) area sources (877 sources of the residential sector); d) linear sources (1156 sources of the urban transportation system). Emission inventory of the last two groups consists of 100 m x 100 m area emission units. Each category of sources is characterized by specific technological parameters, a set of emitted compounds and correlation between the basic pollutants. The main results of simulation relate to the annual average concentration forecasts of the following primary compounds: SO_2, NO_X, PPM10, PPM2.5, PPM10_R, PPM2.5_R, BAP, Ni, Cd, Pb, PAH (here PPM10_R, PPM2.5_R denote particulate matter raised by the road traffic) and secondary: $SO_4^=$, NO_3^-, HNO3, PM10, PM2.5 (here the total particulate matter concentrations are denoted PMx = PPMx + PPMx_R + $SO_4^=$ + NO_3^-). Resulting concentrations have been computed and assigned to 563 receptor points shown in Fig. 1.

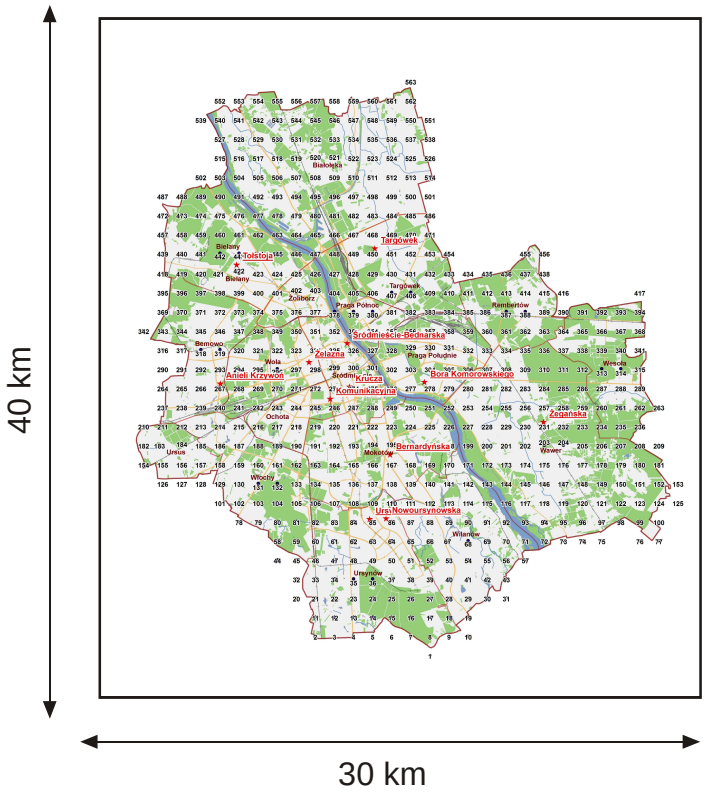

Figure 1. Computational domain; location of receptor points and monitoring stations.

Monte Carlo statistical technique has been utilized to propagate uncertainty of the emission data. Simulations of the pollution dispersion have been performed for 2000 randomly generated sets of input emission data and then utilized by CALPUFF atmospheric transport model. The input uncertainty of emission intensity for each category of sources was individually generated for each pollutant (normal distribution is assumed). Example ranges of emission uncertainty for selected polluting compounds, depending on source category are shown in Table 1. To avoid creating technologically unrealistic sets of emission data (Page T. *et al.*, 2003), the dedicated random number generator takes into account correlation between key compounds for each category of sources.

Table 1. Ranges of input emission uncertainty for the main pollutants in four categories of sources.

Pollutant	High point sources	Other point sources	Area sources	Linear sources
SO_2	± 15%	± 20%	± 30%	± 30%
NO_X	± 20%	± 30%	± 40%	± 40%
PPM10	± 25%	± 30%	± 40%	± 40%
PPM2.5	± 25%	± 30%	± 40%	± 40%

Results of Monte Carlo analysis are recorded in a dedicated database. The data have a form of annual-averaged concentration values of all pollutants at receptor points including the output distribution resulting from the input data uncertainty. Keeping the data in the database allows for easy further analysis of accuracy and uncertainty of modeling performance as well as quantitative evaluation of the impact of specific emission sources to air quality and to harmful environmental effects (including human health). This makes a ground for searching counteraction policies.

SELECTED RESULTS

The data gathered enabled us to perform complex statistical analysis, for example concerning forecast uncertainty of pollutants (including full statistical characterization of uncertainty distribution), relative environmental impact of the selected emission sources or categories of sources (important in supporting decisions related to emission control). Preliminary evaluation of the model performance accuracy can be seen by comparing obtained averaged concentrations with observations by air monitoring system. As an example, Figure 2 presents comparison of averaged concentration of particulate matter (PM10) with measurement values registered at monitoring stations (locations shown in Fig. 1). The dashed lines show ranges of the factor of 2, usually adopted in comparison of modeling and observed atmospheric pollution data.

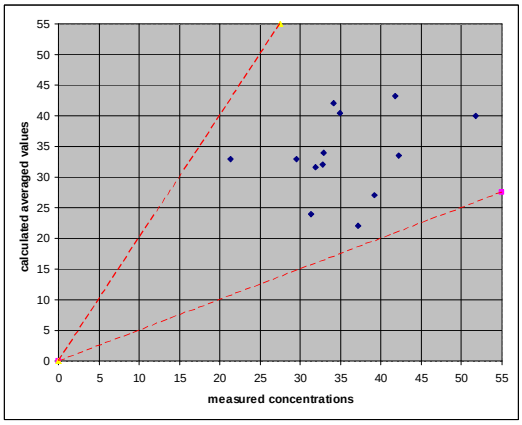

Figure 2. Computed versus observed annual averaged PM10 concentrations [$\mu g \cdot m^3$] in 2005.

The general database has been integrated with ArcMap-ArcView instrumentation, which enables wide possibilities of graphical presentation of results. An example of the available visualizations presented in Fig. 3 covers a map of the selected sub-domain of Warsaw area and relates to the resulting uncertainty distributions of PM10 concentrations. The map directly presents - in a form of the column plots - approximate distributions of the total PM10 concentrations at receptor points. Moreover, it is possible to develop full graphics of uncertainty distribution. Fig. 3 illustrates this case for the receptor No 275, where the extended graphics displays in the vertical layout: cumulative distribution function, distribution density and the standard box plot for the selected receptor and polluting compound.

Figure 3. Annual average concentration (and uncertainty distribution) of PM10 at the receptor points.

Analysis of the results provides general conclusions related to the share of the assumed categories of emission sources in environmental impact and the future decisions. In particular, point sources mainly contribute to SO_2 and NO_X pollution, but the resulting concentration uncertainty is rather small, mainly due to high precision of input emission for this category of sources. Relatively low contribution of the major heating plants is due to the stack height of this group of sources (about 200 m), which causes exporting of the substantial part of pollution outside of the domain. Therefore, the particulate matter pollution is mainly due to area sources (PM10 and PM2.5) and linear sources (PM10). Moreover, the resulting uncertainties of these two pollutants are relatively high and range from ±6% up to ±30%, depending on the receptor location. This is due to high input uncertainty of area and linear sources (compare Table 1).

SENSITIVITY OF HEALTH EFFECTS TO EMISSION UNCERTAINTY

We used the forecasted PM2.5 concentrations to estimate the adverse health effects caused by PM2.5 air pollution in Warsaw. We also compared sensitivity of the mortality estimates to different emissions sources uncertainties to examine, which uncertainties are most important when estimating adverse health effects caused by PM2.5 in Warsaw Metropolitan Area.

Estimation of adverse health effects caused by PM2.5 air pollution followed the methods presented in (Tainio 2009). First we estimated the exposure to PM2.5 in Warsaw Metropolitan Area by calculating the population average exposure to PM2.5. Second, we estimated background mortality in the study area. Third, we estimated exposure-response functions that describe the statistical relationship between exposure and adverse health effects. Finally we combined all the data to estimate the adverse health effects caused by PM2.5 in Warsaw in year 2005. After that, we performed a sensitivity analysis for the mortality estimates. These different phases are described more detailed below.

Population average exposure. Population average exposure describes the average exposure encountered by the population in the given area. The population average exposure was calculated with the following equation:

$$E = \sum_i C_i (Pop_i / Pop) \quad (1)$$

In this equation, E is the exposure for PM2.5 [$\mu g \cdot m^{-3}$], C is PM2.5 the concentration [$\mu g \cdot m^{-3}$], Pop is the number of population. The population data (Pop_i) for Warsaw Metropolitan Area was obtained from European Environment Agency (EEA) (http://www.eea.europa.eu/data-and-maps/data/population-density-disaggregated-with-corine-land-cover-2000-2). The EEA population data included population density in 100 m x 100 m grid over the Europe. The data was transferred with ArcMap to population estimates for the same receptor points, for which the PM concentrations have been estimated (see Figure 1). Population of the study area was estimated to be 1 790 872.

The resulting population average concentrations for different pollutants and sources are presented in Table 2. Total population average exposure to PM2.5 due to local sources was estimated to be approximately 7.0 $\mu g \cdot m^{-3}$. Primary PM2.5 from linear sources contributed approximately half of total population average exposure in the study area.

Table 2. Population average exposure (mean and 95% confidence intervals) to PM2.5 concentrations [$\mu g \cdot m^{-3}$] due to local emissions sources in Warsaw Metropolitan Area. The uncertainty includes only the emission source uncertainty.

Pollutant	High point sources	Other point sources	Area sources	Linear sources	Together
$SO_2 + NO_X$	0.04 (0.04-0.04)	0.03 (0.03-0.03)	0.04 (0.04-0.04)	0.31 (0.31-0.31)	0.41 (0.41-0.41)
PPM2.5	0.02 (0.02-0.02)	0.24 (0.24-0.24)	2.00 (1.98-2.04)	1.44 (1.41-1.46)	3.69 (3.65-3.73)
PPM2.5_R	-	-	-	2.89 (2.85-2.92)	2.89 (2.85-2.92)
Together	0.06 (0.06-0.06)	0.27 (0.26-0.27)	2.04 (2.01-2.07)	4.63 (4.58-4.67)	6.99 (6.94-7.05)

Background mortality. Background mortality was estimated based on Central Statistic Office mortality data for Warsaw for 2007. Non-accidental mortality was estimated to be 16 583 deaths per year.

Exposure-response function. Exposure-response function was obtained from Tuomisto *et al.* (2008). That study used an expert elicitation method to estimate the exposure-response function for PM2.5 air pollution. The exposure-response function is presented with percentage change in mortality caused by 1 $\mu g \cdot m^{-3}$ change in PM2.5 exposure. Based on Tuomisto *et al.* (2008), we assumed that 1 $\mu g \cdot m^{-3}$ change in PM2.5 exposure would change non-accidental mortality 0.62%. Exposure-response function uncertainties were not taken into account because this study focuses on emission uncertainties.

Table 3. Premature mortality estimates for Warsaw Metropolitan Area for 2005. Unit is premature deaths per year.

Pollutant	High point sources	Other point sources	Area sources	Linear sources	Together
$SO_2 + NO_X$	4 (4-4)	3 (3-3)	4 (4-4)	31 (31-32)	42 (42-43)
PPM2.5	2 (2-2)	24 (24-25)	206 (203-209)	148 (145-150)	380 (376-384)
PPM2.5_R	-	-	-	297 (293-300)	297 (293-300)
Together	6 (6-6)	27 (27-28)	210 (206-213)	476 (471-480)	719 (713-724)

Premature mortality. The premature mortality was calculated with the following equation:

$$M = M_b \cdot E \cdot (ER/100) \quad (2)$$

In the equation, M is premature death due to PM2.5 air pollution, M_b is background non-accidental mortality in Warsaw, E is population average exposure to PM2.5 [$\mu g \cdot m^{-3}$], and ER is exposure-response function. The resulting premature death estimates are presented in Table 3. PM2.5 air pollution due to local sources was estimated to cause approximately 700

premature deaths per year in the Warsaw Metropolitan Area. Relative contribution of different sources is the same as in population average exposure because we did not assume any toxicity differences between different PM2.5 emission sources.

Sensitivity analysis. Sensitivity analysis was used to estimate how uncertainty in different input values impact the mortality results. Sensitivity analysis was done by calculating rank-order correlation between the input variables and the mortality estimates uncertainties. The results of the sensitivity analysis are presented in Fig. 4. Emission uncertainties related to linear sources were having highest impact to premature mortality estimates uncertainties. This is due to both high emission uncertainties (see Table 1) and importance of linear sources when estimating adverse health effects for the study area (Table 3). Other uncertainties beside the emission uncertainty were not taken into account in these calculations.

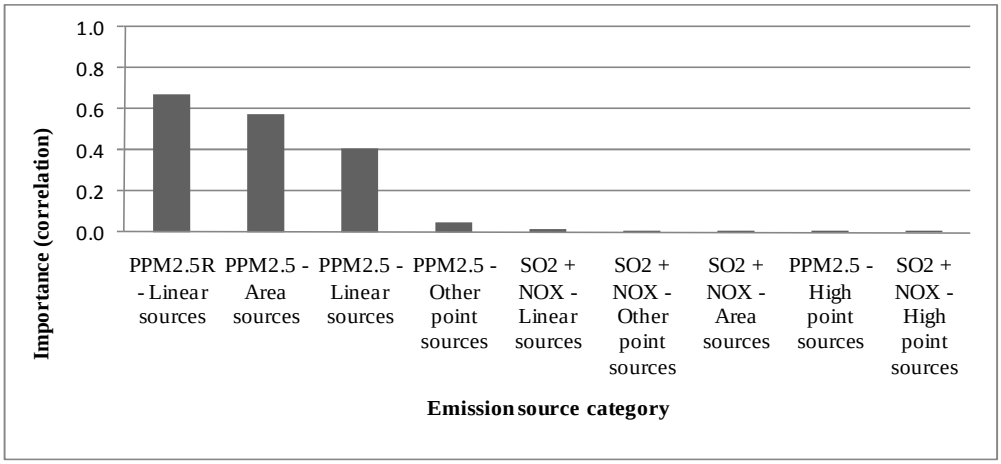

Figure 4. Sensitivity analysis. High importance means the large impact of particular input variable to the uncertainty of the model. Only emission uncertainties were taken into account in these calculations. From the different emission source categories -- primary PM2.5 emission uncertainties related to the linear and area sources were having highest impact to the uncertainty.

Acknowledgements: This research has been supported by the Polish Ministry of Science and Higher Education under Grant N N519316735. The authors would also like to thank Wojciech Trapp for his assistance in preparation of emission inventory.

REFERENCES

ApSimon H.M., R.F. Warren and S. Kayin, 2002: Addressing uncertainty in environmental modeling: a case study of integrated assessment of strategies to combat long-range transboundary air pollution. *Atmospheric Environment*, **36** 5417-5426.

Hanna S., J.C. Chang and M.E. Fernau, 1998: Monte Carlo estimates of uncertainties predictions by a photochemical grid model (UAM-IV) due to uncertainties in input variables. *Atmospheric Environment*, **32**, 3619-3628.

Moore G.E. and R.J. Londergan, 2001: Sampled Monte Carlo uncertainty analysis for photochemical grid models. *Atmospheric Environment*, **35**, 4863-4876.

Page T., J.D. Whyatt, K.J. Beven and S.E. Metcalfe, 2003: Uncertainty in modeled estimates of acid deposition across Wales: a GLUE approach. *Atmospheric Environment*, **38**, 2079-2090.

Park S.-K., C.E. Cobb., K. Wade, J. Mulholland, Y. Hu and A.G. Russel, 2006: Uncertainty in air quality model evaluation for particulate matter due to spatial variations in pollutant concentrations. *Atmospheric Environment*, **40**, S563-S573.

Russel A. and D. Dennis, 2000: NASTRO critical review of photochemical models and modeling. *Atmospheric Environment*, **34**, 2283- 2324.

Sax T. and V. Isakov, 2003: A case study for assessing uncertainty in local-scale regulatory air quality modeling applications. *Atmospheric Environment*, **37,** 3481-3489.

Sportisse B., 2007: A review of current issues in air pollution modeling and simulation. *Computational Geosciences*, **11,** 159-181.

Tainio M., 2009: Methods and uncertainties in the assessment of the health effects of fine particulate matter (PM2.5) air pollution. National Institute for Health and Welfare, Helsinki, 165 pp.

Tuomisto J.T., A. Wilson, J.S. Evans and M. Tainio, 2008: Uncertainty in mortality response to airborne fine particulate matter: Combining European air pollution experts. *Reliability Engineering and System Safety*, **93**, 732–744.

Warren R.F. and H.M. ApSimon, 1999: Uncertainties in integrated assessment modeling of abatement strategies: illustrations with the ASAM model. *Environment Science and Policy*, **2**, 439-456.

Zimmermann H.-J., 1999: An application-oriented view of modeling uncertainty. *European Journal of Operational Research*, **122**, 190-198.