Environmental Modelling & Software 84 (2016) 165–176

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

A parameter estimation and identifiability analysis methodology applied to a street canyon air pollution model

Thor-Bjørn Ottosen [a, c], Matthias Ketzel [b], Henrik Skov [b, c], Ole Hertel [b], Jørgen Brandt [b], Konstantinos E. Kakosimos [a, *]

[a] *Department of Chemical Engineering, Texas A&M University at Qatar, Doha, Qatar*
[b] *Department of Environmental Science, Aarhus University, Roskilde, Denmark*
[c] *Department of Chemical Engineering, Biotechnology and Environmental Technology, University of Southern Denmark, Odense, Denmark*

A R T I C L E I N F O

Article history:
Received 1 December 2015
Received in revised form
27 June 2016
Accepted 27 June 2016

Keywords:
Uncertainty
Sensitivity
OSPM
Data splitting
Exploratory data analysis
Matlab

A B S T R A C T

Mathematical models are increasingly used in environmental science thus increasing the importance of uncertainty and sensitivity analyses. In the present study, an iterative parameter estimation and identifiability analysis methodology is applied to an atmospheric model – the Operational Street Pollution Model (OSPM®). To assess the predictive validity of the model, the data is split into an estimation and a prediction data set using two data splitting approaches and data preparation techniques (clustering and outlier detection) are analysed. The sensitivity analysis, being part of the identifiability analysis, showed that some model parameters were significantly more sensitive than others. The application of the determined optimal parameter values was shown to successfully equilibrate the model biases among the individual streets and species. It was as well shown that the frequentist approach applied for the uncertainty calculations underestimated the parameter uncertainties. The model parameter uncertainty was qualitatively assessed to be significant, and reduction strategies were identified.

1. Introduction

A few decades ago, the use of mathematical models was mainly limited to the use internally in the scientific community, meaning that the model users to a larger extend had an explicit or implicit understanding of the model uncertainty and sensitivity.[1] Today, mathematical models are often routinely used by engineers, consultants, and planners as well as scientists for environmental regulation and to assess consequences of abatement strategies. This development supports the need for explicit uncertainty and sensitivity analyses to facilitate the communication among model stakeholders.

Within air pollution modelling there has been a growing number of publications on uncertainty and sensitivity analyses in recent years (for a review see Hanna (2007)). Walker et al. (2003) defined six uncertainty categories, based on the location of the uncertainty, of which some have been studied within air pollution modelling. Model technical uncertainty (e.g. Franke et al. (2007)) and model input uncertainty (Manomaiphiboon and Russell, 2004; Hanna et al., 2007; Bei et al., 2012) have been studied previously, however, model parameter uncertainty has received comparatively little attention (e.g. Marsik and Johnson (2010)).

Vardoulakis et al. (2002) studied the local sensitivity of the Operational Street Pollution Model (OSPM®) to marginal changes in ten model parameters for an artificial dataset (parallel and perpendicular wind directions and a constant wind speed). Silver et al. (2013) analysed the applicability of a dynamic parameter estimation (Parameter estimates change along with changes in data) scheme to OSPM for planning and forecasting. Secondarily, Silver et al. (2013) showed in a preliminary application of static parameter estimation (One set of parameters are estimated for all data points) that such an approach can be informative. Silver et al. (2013) used between one and four years of data for five streets. The parameter estimation scheme in Silver et al. (2013) was applied to one model parameter and five multiplicative adjustment factors.

---

* Corresponding author.
E-mail addresses: t.ottosen@worc.ac.uk (T.-B. Ottosen), mke@envs.au.dk (M. Ketzel), hsk@envs.au.dk (H. Skov), oh@envs.au.dk (O. Hertel), jbr@envs.au.dk (J. Brandt), k.kakosimos@qatar.tamu.edu (K.E. Kakosimos).

[1] The analysis of change in model output of marginal changes in one model parameter at a time. This is opposed to *global sensitivity analysis* which is the analysis of change in model output of large changes of several model parameters at a time.

http://dx.doi.org/10.1016/j.envsoft.2016.06.022
1364-8152/

This study can thus be classified as being somewhere between analysis of model input and model parameter uncertainty. The choice of model parameters and multiplicative adjustment factors was based on the developers experience with the model. A natural question is thus whether a more systematic approach would yield better results?

Brun et al. (2002, 2001) developed a systematic parameter estimation and identifiability analysis methodology, which has been applied among others in lake modelling (Omlin et al., 2001), river modelling (Meier and Reichert, 2005; Anh et al., 2006), modelling of waste-water treatment plants (Sin and Vanrolleghem, 2007b), forest modelling (De Pauw et al., 2008), surface hydrology modelling (Freni et al., 2009; Muñoz et al., 2014), and material science (Martinez-Lopez et al., 2015) but has not been applied within atmospheric science before.

To analyse the applicability of, and to explore the potential advantages of, applying this methodology to a model within atmospheric science, the methodology of Brun et al. (2002, 2001) was applied to the Operational Street Pollution Model (OSPM®). The application utilizes more years of data and more parameters compared to the analysis performed in Silver et al. (2013).

This paper explains the appropriate data preparation techniques, reports the results of the application of, and explores the advantages of this methodology through exploratory data analysis of the results.

The working principles behind OSPM are described in Section 2. The model input, the measurements, and the methodologies for data preparation, parameter estimation, and identifiability analysis are likewise explained in Section 2. The results and discussion of the various sub-analyses performed in the present study are presented in Section 3. The conclusions are subsequently presented in Section 4.

2. Model description and methods

The applied methodology, as illustrated on Fig. 1, consists of running parameter estimation and identifiability analysis in an iterative series. This is done until convergence between the obtained parameters and the identifiability of the parameters is achieved. Following the steps outlined in Fig. 1, the model definition has been done in Hertel and Berkowicz (1989b,a,c); Berkowicz et al. (1997) as briefly described in Section 2.1. The experimental layout is defined by the data available through the Danish national air quality monitoring programme and as part of the prior analysis, data preparation has been performed.

2.1. Model description

OSPM is a model for vehicle induced urban street pollution. The model is designed to take differences in atmospheric conditions and types of street into account. The main characteristics of OSPM are:

- The applied version of OSPM consists of emissions calculated with COPERT IV (EMEP/EEA, 2009) and a dispersion model running in series. To limit the scope of the present study the focus is on the parameters related to the dispersion model.
- OSPM models the resulting hourly averaged pollution concentrations, of a specific species, at the side of the street. This is calculated as a sum of a direct contribution (C_{dir}) and a recirculating contribution (C_{rec}) plus a background concentration. The direct contribution is modelled using a simplified Gaussian plume model with a top hat distribution applied to the emission plume. The recirculating contribution is modelled using a trapezium shaped box model (Hertel and Berkowicz, 1989b; Berkowicz et al., 1997; Berkowicz, 2000).

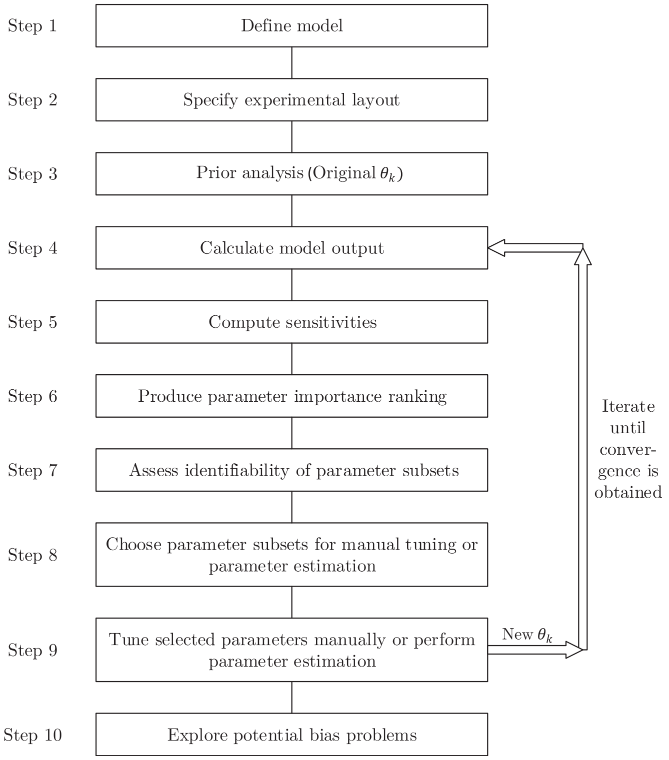

Fig. 1. Schematic representation of the parameter estimation and identifiability analysis methodology applied in the present study (figure based on Brun et al. (2002)). θ_k is the model parameter vector.

- The wind direction, especially for low wind speeds, cannot be assumed constant over a full hour. To account for this, a numerical wind direction averaging procedure is implemented in the model (Hertel and Berkowicz, 1989c).
- The model also contains an algebraic expression for traffic produced turbulence. The expression depends on the number of cars in the street, their respective driving speeds, and the traffic composition (Hertel and Berkowicz, 1989c).
- Most traffic pollutants are assumed to be inert on the time scale of the residence time in a street canyon. However, the conversion of NO to NO_2 in the presence of ozone happens faster. It is therefore included in the model in the form of an algebraic chemical conversion scheme (Hertel and Berkowicz, 1989a; Palmgren et al., 1996). The majority of the parameters of the chemical conversion module are left out of the subsequent parameter estimation to limit the scope of the study.

A total of 16 model parameters have been identified in the model. These are briefly summarized in Table 1. A more detailed description of the model can be found at www.au.dk/ospm or in Berkowicz et al. (1997); Ottosen et al. (2015).

2.2. Model inputs

The concentration and meteorology input data come from the Danish national air quality monitoring programme (Ellermann et al., 2013). In this programme hourly air quality measurements have been performed since 1994. Measurements are performed in two streets in Copenhagen (Jagtvej and H. C. Andersens Boulevard (hereafter referred to as HCAB)) and in one street in respectively Aarhus, Aalborg, and Odense. A map of the streets can be found in