Atmospheric Environment 169 (2017) 267–277

Contents lists available at ScienceDirect

Atmospheric Environment

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

Minimization of model representativity errors in identification of point source emission from atmospheric concentration measurements

Maithili Sharan [a, *], Amit Kumar Singh [a], Sarvesh Kumar Singh [b]

[a] Centre for Atmospheric Sciences, Indian Institute of Technology Delhi, Hauz Khas, New Delhi, 110016, India
[b] LMEE, Universite d'Evry-Val d'Essonne, 40 Rue Du Pelvoux, 91020 Evry Cedex, France

HIGHLIGHTS

- Minimization of model representativity errors through regression analysis.
- Application of the methodology in the context of two inversion techniques.
- Impact of error minimization using data from low wind stable conditions.
- Significantly improved source reconstruction with the proposed methodology.

ARTICLE INFO

Article history:
Received 18 February 2017
Received in revised form
8 September 2017
Accepted 11 September 2017
Available online 18 September 2017

Keywords:
Inverse modelling
Linear regression
Least square
Model representativity errors
Renormalization
Source identification

ABSTRACT

Estimation of an unknown atmospheric release from a finite set of concentration measurements is considered an ill-posed inverse problem. Besides ill-posedness, the estimation process is influenced by the instrumental errors in the measured concentrations and model representativity errors. The study highlights the effect of minimizing model representativity errors on the source estimation. This is described in an adjoint modelling framework and followed in three steps. First, an estimation of point source parameters (location and intensity) is carried out using an inversion technique. Second, a linear regression relationship is established between the measured concentrations and corresponding predicted using the retrieved source parameters. Third, this relationship is utilized to modify the adjoint functions. Further, source estimation is carried out using these modified adjoint functions to analyse the effect of such modifications. The process is tested for two well known inversion techniques, called renormalization and least-square. The proposed methodology and inversion techniques are evaluated for a real scenario by using concentrations measurements from the Idaho diffusion experiment in low wind stable conditions. With both the inversion techniques, a significant improvement is observed in the retrieval of source estimation after minimizing the representativity errors.

1. Introduction

Model and observations describe different parts of the reality. The model state is an imperfect representation (or modeled version) of the reality. From the perspective of optimal estimation theory, both the model state and the observations are estimates subject to uncertainty (Cohn, 1997). The model state is associated with systematic errors, uncertainties in the initial and boundary conditions and the approximate nature of the model dynamics (Thacker, 2003). The observations might characterize unmodelled processes and contain instrumental inaccuracies. More importantly, the observation operator is not true (model resolution is inconsistent with the nature of measurements) and contains a representativity error present in the observations which is not directly related to errors in the measurement process. Thus, representativity errors can be defined as a residual mismatch, component of the observation errors, due to unresolved scales resulting from different representativity of reality in model and observations (Thacker, 2003; Janjic and Cohn, 2006; Oke and Sakov, 2008). This includes (i) errors due to any physical processes appearing in the observations but not in the model (Anderson et al., 2005; Zaron and Egbert, 2006; Ponte et al., 2007) and (ii) errors in the model equations, either from the missing physics or from discretisation (van Leeuwen, 2015).

* Corresponding author.
E-mail address: mathilis@cas.iitd.ac.in (M. Sharan).

https://doi.org/10.1016/j.atmosenv.2017.09.015
1352-2310/

In the context of air pollution, dispersion models still have large but unknown deficiency. This is one of the major obstacles for source identification problems which has significant importance in emergency response applications (Bocquet, 2005; Issartel et al., 2007; Yee, 2007; Allen et al., 2007; Sharan et al., 2009; Bocquet et al., 2011; Singh and Rani, 2014; etc). The source retrieval is addressed from the tracer's concentrations measured by a given monitoring network along with a data assimilation technique and an appropriate dispersion model. The objective is to minimize the discrepancy between the observations (i.e. measured concentrations) and their corresponding predictions given by the atmospheric dispersion models. The well-known difficulties are sparsity of the observations in time and space, ill-posedness, loss of information due to averaging process, model uncertainties and lack of knowledge about background and observation error statistics.

Note that the tracer concentrations, reported by the fixed monitoring networks, are described as point measurements sampled and averaged over a duration of time. A representativity error arises while dealing with such measurements since the observations represent phenomena that are not resolved by the model. The unresolved variables are continuous and inconsistent to model resolution (Cohn, 1997; Janjic and Cohn, 2006). The data assimilation techniques provide a conditional estimate of the model state given the observations, not of reality given the observations since the representativity errors are unavoidable (Cohn, 1997; Derber and Rosati, 1989; van Leeuwen, 2015). Thus, an additional treatment is required to deal the incompatibilities between the model, data and errors so that model characteristics can be forced towards the observations (Thacker, 2003).

In general, representativity errors are taken into account by inflating the error covariance of the observations, sometimes dependent on position and/or flow regime, or by adding an extra error covariance to the observation-error covariance in the likelihood (Derber and Rosati, 1989). However, a statistical knowledge about model and observation error statistics is not known (often parametrized based on hypothetical assumptions) in source retrieval problems. In an inverse modelling study of carbon monoxide, Mulholland and Seinfeld (1995) have emphasized to adjust the model representativity by tuning the optimal weighting factors in order to avoid steep concentration gradients in stable conditions. Similarly, Krysta et al. (2008) have observed model representativity error in a source reconstruction with ETEX-II data set and claimed a discrepancy scaling factor of seven between the reconstructed and the true mass of releases. Bocquet et al. (2011) and Wu et al. (2011) have derived expressions for the aggregation errors and show to reduce the representativity error related to a too low resolution of the model to a simple addition of a covariance matrix to the measurement error matrix. In real life problems, the model representativity errors are completely inevitable because no forward model can ever incorporate all the physics associated with the problem (Addepalli et al., 2011). Thus, it is desired to develop an approach so that model representativity errors are minimized.

In general, inversion techniques lead to an estimate of the release parameters which describe the observation characteristics with minimum possible errors. However, in case of an imperfect dispersion model or complex meteorological situation (for instance, low wind stable conditions), retrieved release parameters may deviate significantly large from their true values since observation characteristics are difficult to be followed by a simplified model. This is highlighted by Sharan et al. (2012a) in the context of a point source reconstruction in low wind conditions ($U < 2ms^{-1}$, U is mean wind speed). In these conditions, source retrieval becomes complex since (i) the diffusion of pollutant is irregular and indefinite in weak and variable wind (Kumar and Sharan, 2009), (ii) No plume centreline is obvious and the observed concentration distribution is irregular, multi-peaked and non-Gaussian (Sagendorf and Dickson, 1974) and (iii) non-monotonic concentration gradients arises. Thus, the objective here is to propose a methodology to improve the first retrieved estimate of the release parameters by further improving the model representativity. The methodology is tested here for two well known inversion techniques, called renormalization (Issartel et al., 2007) and least-squares (Sharan et al., 2012b), using real measurements from Idaho diffusion experiment (Sagendorf and Dickson, 1974) in low wind stable conditions.

The previous applications (Sharan et al., 2009, 2012a, 2012b) of the inversion techniques and dataset evaluation are mainly involved in predicting a set of required source parameters which, often, differ from their true values. In this article, we analyzed that the deviations between true and predicted source parameters are directly related to the deviations between the observed and predicted concentrations on the receptors. Thus, to retrieve the accurate source parameters, it is important to minimize the corresponding deviations between observed and predicted concentrations. In view of this, a regression based procedure is utilized to improve the model predicted sensitivities (i.e. to minimize the representativity error) or adjoint functions which are utilized in the inversion. The results demonstrate that the inversion results improve significantly using modified model predicted sensitivities.

2. Methodology

The methodology is described here for a continuous point source within the framework of adjoint modelling. The measurements μ_i, $i = 1, 2, \ldots, n$ are generated from a continuous point source, defined as $s(\mathbf{x}) = q_0\delta(\mathbf{x} - \mathbf{x}_0)$ (Eq. (A.8), appendix A) emitting a non-reactive tracer at location $\mathbf{x}_0$ with intensity q_0 where $\delta(.)$ denotes a Dirac delta function and $\mathbf{x} = (x, y)$ is a location vector. The adjoint relationship between measurements and unknown point source is given as (Eq. (A.1)),

$$\mu_i = q_0 a_i(\mathbf{x}_0) \tag{1}$$

in which $a_i(\mathbf{x}_0)$ is the adjoint function which describes sensitivity of i^{th} measurement with respect to location $\mathbf{x}_0$. The adjoint functions are derived as solutions from the adjoint dispersion model (Pudykiewicz, 1998). The identification of a point source refers to the estimation of parameters $\mathbf{x}_0$ and q_0 through an inversion technique by assimilating measurements μ_i with adjoint functions $a_i(\mathbf{x})$.

The methodology is outlined in four major steps (Fig. 1): (i) computation of first estimate of point-source parameters (location $\mathbf{x}_0$ and intensity q_0) using inversion technique, measurements and dispersion model, (ii) development of a regression model to minimize the deviations between observed and model predicted concentrations. Note that these model predicted concentrations are obtained from the dispersion model (in forward mode) using first estimate ($\mathbf{x}_0$ and q_0) of source parameters. (iii) Transformation of regression coefficients into regression curves using curve fitting and modification of adjoint functions using those regression curves, and (iv) application of inversion techniques with measurements and modified adjoint functions.

In the first step, an inversion technique is required for the estimation of the source parameters $\mathbf{x}_0$ and q_0. In this study, the methodology is tested for two inversion techniques, namely renormalization (Issartel et al., 2007) and least-squares (Sharan et al., 2012b). Both the techniques are given in appendices A and B for the sake of completeness. The retrieved estimates ($\hat{\mathbf{x}}_0$ and $\hat{q}_0$) are utilized with a forward dispersion model to predict the concentrations corresponding to the measurements. The second