Atmospheric Environment 119 (2015) 402–414

Contents lists available at ScienceDirect

Atmospheric Environment

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

Assimilation of concentration measurements for retrieving multiple point releases in atmosphere: A least-squares approach to inverse modelling

Sarvesh Kumar Singh [a, *], Raj Rani [a, b]

[a] LMEE, Universite d'Evry-Val d'Essonne, 40 Rue Du Pelvoux, 91080 Courcouronnes, France
[b] Centre for Atmospheric Sciences, IIT Delhi, Hauz Khas, New Delhi 110016, India

HIGHLIGHTS

- A least-squares algorithm for multiple point releases identification is presented.
- A first guess of the release parameters is not required.
- An evaluation and comparison is shown using data from Fusion Field Trials.
- Future applicability and limitations of the inversion algorithm is discussed.

ARTICLE INFO

Article history:
Received 11 May 2015
Received in revised form
18 August 2015
Accepted 21 August 2015
Available online 28 August 2015

Keywords:
Data assimilation
Fusion field trials
Least-squares
Multiple releases
Source reconstruction

ABSTRACT

The study addresses the identification of multiple point sources, emitting the same tracer, from their limited set of merged concentration measurements. The identification, here, refers to the estimation of locations and strengths of a known number of simultaneous point releases. The source–receptor relationship is described in the framework of adjoint modelling by using an analytical Gaussian dispersion model. A least-squares minimization framework, free from an initialization of the release parameters (locations and strengths), is presented to estimate the release parameters. This utilizes the distributed source information observable from the given monitoring design and number of measurements. The technique leads to an exact retrieval of the true release parameters when measurements are noise free and exactly described by the dispersion model. The inversion algorithm is evaluated using the real data from multiple (two, three and four) releases conducted during Fusion Field Trials in September 2007 at Dugway Proving Ground, Utah. The release locations are retrieved, on average, within 25–45 m of the true sources with the distance from retrieved to true source ranging from 0 to 130 m. The release strengths are also estimated within a factor of three to the true release rates. The average deviations in retrieval of source locations are observed relatively large in two release trials in comparison to three and four release trials.

1. Introduction

In the atmospheric dispersion events, fast and accurate identification of unknown releases is one of the major concern to advance the emergency assessment capabilities and to minimize the threat of exposure to the environment. The identification refers to determine the time, origin and strength of the unknown releases. The identification process is mainly governed by three major components: (i) a network of receptors for the rapid detection of the contaminants, (ii) an atmospheric dispersion model for the prediction of contaminant's concentrations in space and time, and (iii) an optimal integration (or assimilation) scheme to assimilate the measured concentrations with the atmospheric dispersion models in order to retrieve the unknown releases. In past years, significant advances are noted in the sensing technology and dispersion models (including simple/obstructed terrain, various atmospheric conditions, etc.). An attention is further required to develop the concentration data assimilation techniques for the retrieval of unknown releases in a fast and consistent manner.

* Corresponding author.
E-mail address: singh@ufrst.univ-evry.fr (S.K. Singh).

http://dx.doi.org/10.1016/j.atmosenv.2015.08.063

The dispersion events might involve one or more releases simultaneously emitting the contaminants. In case of simultaneous releases emitting the same contaminant, the field of plumes may overlap significantly and the sampled concentrations may become the mixture of the concentrations originating from all the releases. The other uncertainties may arise as, (i) the sources are seen from the same angle but are located at different distances, (ii) the receptors near to a weak source will report same concentration as the receptors far away from a strong source, etc. In such cases, it is challenging to separate the influence of each source and to correctly identify each source from a set of merged concentration measurements. In local scale dispersion events, the unknown releases are often formulated as point type and their identification is addressed by estimating a fixed set of parameters, for instance, ground level coordinates of the release location, height, strength, etc,.

Several studies have been carried out addressing the identification of single point release. However, the identification of multiple-point releases is relatively more difficult and has received limited contribution. The approaches were mainly based on the principles of Bayesian inference coupled with sampling algorithms (for example, Markov Chain Monte Carlo (MCMC) etc.) and optimization. The advantage of Bayesian inference lies in estimating the source parameters along with their confidence intervals and posterior statistics whereas optimization techniques provide the parameters which maximally match the measurements. For an efficient computation of source-receptor sensitivity matrix, adjoint of the dispersion model is often suggested (Pudykiewicz, 1998; Yee, 2008; Sharan et al., 2009; etc.). The source parameters are estimated by minimizing the deviations iteratively between the observations and the adjoint solutions (Pudykiewicz, 1998).

In recent works, Yee (2007) have performed a joint estimation of locations and strengths by utilizing a Bayesian probabilistic inference coupled with MCMC sampling assuming that the number of sources are known a priori. The assumptions about number of sources were relaxed by Yee (2008) by utilizing a Metropolis-coupled reversible-jump MCMC method. Lushi and Stockie (2010) have performed the estimation of release rates of the multiple-point emissions by means of a linear least-squares approach in a large lead-zinc smelting operation in Trail, British Columbia. Albo et al. (2011) have developed an Aerodyne inverse modelling systems by combining the tangent linear of SCIPUFF (Second-order Closure Integrated Puff) with a minimization algorithm for characterizing the atmospheric releases (including number of releases, their locations, mass, release times and durations) without requiring initial guess. Sharan et al. (2012) have proposed a least-squares minimization method coupled by an adjoint dispersion model to jointly estimate the locations and strengths of the multiple-point emissions. The method is free from the initial guesses of the release parameters. However, the evaluation is shown with the model generated and pseudo-real datasets. Similarly, Singh et al. (2013) have introduced a weighted least-squares method coupled with an adjoint dispersion model for the joint estimation of locations and strengths of the multiple-point releases. The weights utilize the natural statistics based on the geometry of the monitoring network (Issartel et al., 2007). Annunzio et al. (2012) have proposed a state estimation technique based on Lagrangian approach called as "Multi-Entity Field Approximation" to determine the locations of the multiple-point releases.

Fusion Field Trials (FFT07) refer to a series of short range diffusion tests conducted at Dugway Proving Ground, Utah during September 2007 (Storwald, 2007). The dataset corresponds to the instantaneous/continuous single as well as multiple (two, three and four) point releases. The experiment is designed and distributed widely for evaluating the performance and capability of several source estimation algorithms. In this study, an inversion technique proposed by Sharan et al. (2012) is revisited and modified to efficiently address the retrieval of continuous multiple point releases using real measurements from FFT07 datasets. The objective is to highlight the capability and efficiency of the inversion technique in identifying the parameters (mainly, locations and strengths) corresponding to the continuous multiple point releases in a real scenario.

2. Inversion methodology

The present study deals with a known number of simultaneous point releases emitting a non-reactive tracer from different locations with different release rates. The study is focussed for ground level continuous point releases. This refers to the estimation of locations and release strengths for the known number of simultaneous point releases. The concentrations measured by the samplers are a mixture of the concentrations resulting from each source. The identification of multiple-point releases is addressed from their measured set of merged concentrations. Accordingly, the vertical and time components are ignored in the formulations. Note that the bold symbol denotes a vector/matrix and italic symbol denotes a scalar/constant.

2.1. Source–receptor relationship

The source–receptor relationship is based on an adjoint representation of a dispersion model which provides the potential sensitivity of unknown releases with respect to the measured concentrations (Pudykiewicz, 1998). Let $\boldsymbol{\mu}\in\mathbb{R}^m$ is the vector of measured concentrations and $\mathbf{s}\in\mathbb{R}^N$ is the vector of unknown emissions in a discretized space composed of N cells. The measurements $\boldsymbol{\mu}$ are related to the emissions $\mathbf{s}$ by the use of sensitivity coefficients (or adjoint functions) which describes the propagation of information backward in space from measurements as,

$$\boldsymbol{\mu} = \mathbf{A}\mathbf{s} + \boldsymbol{\varepsilon} \tag{1}$$

where $\mathbf{A}\in\mathbb{R}^{m\times N}$ is the sensitivity matrix and $\boldsymbol{\varepsilon}\in\mathbb{R}^m$ is the noise vector accounting for the instrumental and model errors. Each column vector of matrix $\mathbf{A}=[\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_N]$ denotes sensitivity of a cell with respect to the m measurements. These are obtained as solutions from the adjoint dispersion model with respect to each measurement (see details in section 4).

Let us suppose that the unknown emission vector is composed of p different point releases located at cells $\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_p$ emitting the same tracer with release rates $q_1,q_2,\ldots,q_p$ respectively, such that

$$s_i = q_i\delta(\mathbf{x}-\mathbf{x}_i), \quad i = 1, 2, \ldots, p \tag{2}$$

where $\mathbf{x} = (x,y)$ is a location vector. By putting Eq. (2) into Eq. (1), one obtains,

$$\boldsymbol{\mu} = \mathbf{H}\mathbf{q} + \boldsymbol{\varepsilon} \tag{3}$$

in which the matrix $\mathbf{H}\in\mathbb{R}^{m\times p}$ denotes sensitivity of the p release locations with respect to m measurements and $\mathbf{q}\in\mathbb{R}^p$ denotes vector of unknown release strength. Now, the problem is to estimate vectors $\mathbf{q}$, $\mathbf{x}_1,\ldots,\mathbf{x}_p$ such that the Euclidean norm $\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}$ is minimum.

2.2. Minimization of cost function

An identification of p continuous point releases involves estimation of $3p$ release parameters $(\mathbf{q},\mathbf{x}_1,\ldots,\mathbf{x}_p)$. In this study, an