Inverse modelling methods for identifying unknown releases in emergency scenarios: an overview

S.K. Singh, M. Sharan, J.P. Issartel

HAL Id: hal-02398161

https://univ-evry.hal.science/hal-02398161

Submitted on 4 Jun 2021

Inverse modelling methods for identifying unknown releases in emergency scenarios: an overview

Sarvesh Kumar Singh[1], Maithili Sharan[1] and Jean-Pierre Issartel[2]

[1]Centre for Atmospheric Sciences, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India Fax: 91-11-26591386

[2]Centre d'Etudes du Bouchet, 5, rue Lavoisier, BP 3, 91710 Vert le Petit Cedex, France

Abstract: The fast and accurate identification of unknown releases helps in an effective implementation of emergency preparedness (including plume exposure assessment) and public safety programs. The identification is based on networking of several receptors, limited a priori information, remotely measured concentrations, use of dispersion model, site description, meteorology, etc. This is considered as an ill-posed inverse problem. An overview of recently proposed/utilised inversion techniques is discussed with their various applications ranging from local scale to regional scale, limitations and inter-comparisons. The studies addressing identification of multiple-point releases are also included. In addition, the current issues, limitations and future scope associated with the inverse modelling are mentioned.

Keywords: Bayesian; inverse modelling; source estimation; unknown releases; variational minimisation.

1 Introduction

1.1 Motivation

The rapid industrial developments and urbanisation lead to an important concern towards risk assessment, security and public safety issues. An attention in these directions increases due to accidental hazards, such as the Bhopal gas leak (1984, India); Chernobyl nuclear disaster (1986, Ukraine), Fukushima nuclear explosion (2011, Japan), etc. The releases are unexpected, highly poisonous and impossible to observe or measure directly on-site. For an effective emergency preparedness program, fast identification of these unknown releases is required. This is treated as an inverse problem (Enting, 1985). The limited available information is site description, meteorology, a priori information about releases and remotely measured concentrations. Therefore, it is essential to develop a methodology for identifying the unknown releases using limited concentration measurements and available a priori information (Figure 1). These methods are called 'inverse modelling techniques' (Rao, 2007; Redwood, 2011).

Figure 1 Correspondence between sources and receptors. 'S' and 'R' denote source and receptor respectively
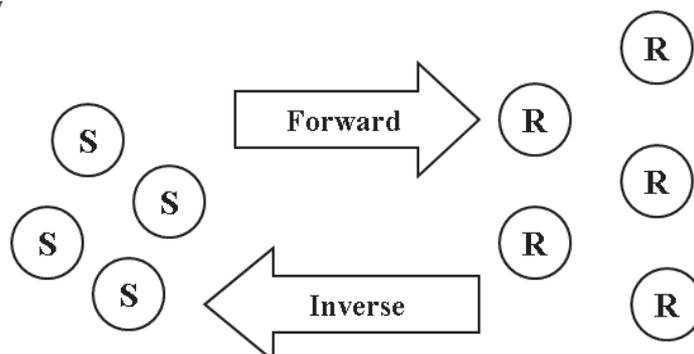
The identification of unknown releases requires an assimilation of concentration measurements, given a suitable atmospheric transport and dispersion model (ATDM) and an error criterion, in such a way that subsequent errors are minimised for the estimated source. The measured concentrations are discrete, limited and obtained as averages in general. The source identification is addressed in a finite dimensional discretised model space; and only a discrete version of the source can be retrieved limited to the resolution of the model space. The discretised model space refers to the discretisation in space as well as in time. Primarily, the source retrieval refers to estimation of origin, strength and release time in the discretised model space irrespective of a priori assumption about nature of the releases. As a priori, the nature of the release is categorised as point, area and volume type. In case of area and volume sources, the emission is assumed to be originated from a large area or volume. On the contrary, a point emission is assumed to originate only from a discrete cell and it is desired to retrieve only that cell along with the mass and time of release. In case of area and volume sources, the patches (regions) with maximum source information are considered as origin; and average flux of emissions estimated over these patches at the time of release can be considered as strength estimate for the area and volume sources. The retrieval of area and volume sources is often termed non-parametric estimations whereas point source retrieval is called a parametric estimation problem subjected to the estimation of location coordinates, height, strength and time of release.

The accuracy of the source retrieval is limited by the resolution and information retrieved in the discretised model space. The major limitations are ill-posedness (caused by imbalance between the number of measurements and degree of freedom in model space), lack of a priori information, model errors, limited and sparse set of concentration measurements, etc. The ill-posedness leads to non-uniqueness (infinite sources can be chosen compatible to the measurements) and instability in the inverse solution. The instability is referred as sensitivity of the retrieved sources towards the random measurement errors. The model errors arise due to representativeness errors, lack of process understanding and turbulence uncertainties. The measurement averaging process causes a loss of high frequency information which is responsible for deviations in the source retrieval. In addition, the loss of information occurs due to unknown structure of random errors from measurement and model during the source retrieval process. Their quantification is important in minimising the uncertainties in the source retrieval and to establish an optimal design of monitoring network (Singh et al., 2015).

1.2 Source-receptor sensitivity

As a first step of inverse modelling, a functional relationship is required between measured concentrations and the unknown source. This is expressed by utilising an ATDM (e.g., advection-diffusion equation or Lagrangian particle dispersion model) governing the transport of pollutant in space and time. The dispersion models provide prediction of concentrations (simulation will often be hindcasts). The source-receptor sensitivity is described in two ways:

1   forward (source-oriented) (e.g., Robertson and Langner, 1998; Krysta et al., 2006)

2   backward (receptor-oriented) (e.g., Issartel and Baverel, 2003).

The source-oriented sensitivity describes the forward transport of pollutants in which the concentrations are simulated from the dispersion model, forward in space and time, under various choices of the source parameters, and later compared with the corresponding receptor measured concentrations (e.g., Thomson et al., 2007, etc.). The parameters, which provide the best fitted concentration measurements, are considered as optimised source parameters. The receptor-oriented relationship is based on the backward propagation of information in the system from the receptors. This is derived from the adjoint of the ATDM and describes the potential sensitivity of the sources with respect to the measured concentrations (e.g., Marchuk, 1995; Pudykiewicz, 1998; Hourdin and Talagrand, 2006; Issartel et al., 2007). Robertson (2004) showed the similarity between back trajectories and adjoint technique using influence functions and examined four different cases as Chernobyl accident, the ETEX-I exercise, the Algeciras accident and an event with heavy smoke noticed in Scandinavia. Back trajectory models estimate back-trajectories using the wind speed and direction data from one or more sites. Back trajectories indicate past paths of infinitesimally small particles of air as they move through time and space (Stohl, 1998). These are obtained by integrating the advection of pollutant backward in time, taking the wind field in the reverse direction. These models are simple and benefit from the time symmetry of tracer transport whereas Eulerian models suffer from numerical diffusion, grid resolution problem near point source emissions and instantaneous mixing of point source emission in the grid. Therefore, this is also known as one of simplest adjoint models for tracing of important source regions attributed to measured tracers (Seibert, 2001). Back trajectory models are quite simple but limited in applications as back-trajectories suffer from over-simplification of the transport-diffusion processes by entirely omitting diffusion and are therefore less useful for long time-integrations (Seibert, 2001).

The model predicted concentration is expressed as a function of input release parameters, such as, $\mathbf{c} = \mathbf{M}(\mathbf{\sigma})$ where $\mathbf{c}$ is the modelled concentration vector of dimension $N \times 1$, $\mathbf{M}$ is the model operator, $\mathbf{\sigma}$ is the discretised source term of dimension $N \times 1$. Here, the bold symbol represents vector and italic symbol represents scalar/constants. The concentration measurement vector $\mathbf{\mu}$ (dimension $m \times 1$) is related to the model predicted concentration $\mathbf{c}$ using an observation operator $\mathbf{H}$, such as, $\mathbf{\mu} = \mathbf{H}(\mathbf{c})$. Therefore, the measured concentrations can be related to the source as

$$\mathbf{\mu} = \mathbf{H}\big(\mathbf{M}(\mathbf{\sigma})\big) = \mathbf{L}(\mathbf{\sigma}) \tag{1}$$

where the operator $\mathbf{L}$ is described as, $\mathbf{L}(\mathbf{.}) = \mathbf{H}(\mathbf{M}(\mathbf{.}))$, includes both observation and model operators. In the present study, the expressions will be written for a linear case. However, in nonlinear case, the formulations will still hold for each stepwise linearisation necessary.

1.3 Cost function

Equation (1) describes a relationship between source and receptors, given a perfect ATDM and no measurement error. In reality, this relationship is associated with model errors, turbulence uncertainties and measurement errors. Therefore, the source estimation is addressed by minimising a cost function (or objective/error function) measuring the residuals between measured concentrations and corresponding model predictions. These

cost functions are commonly defined in terms of least square error (LSE), least absolute error (LAE), maximum likelihood estimation (MLE), etc.

The estimation of source is not trivial since the degree of freedom in model space (N) is imbalanced with the number of available concentration measurements (m). This leads to an ill-posed problem [under-determined ($m < N$) or over-determined ($m > N$)] which further affects the existence, uniqueness, and stability of the computed solution (Figure 2). In case of an under-determined system, additional constraints (like positivity, boundedness, etc.) and a priori information are required to maintain the stability of the solution, whereas in an over-determined system ($m > N$), solution can be obtained without imposing regularisation/constraints. In addition, model errors, turbulence uncertainties, limited concentration measurements and lack of a priori information affect the accuracy and resolution of the source estimation.

Figure 2   A categorisation of estimation problems (see online version for colours)
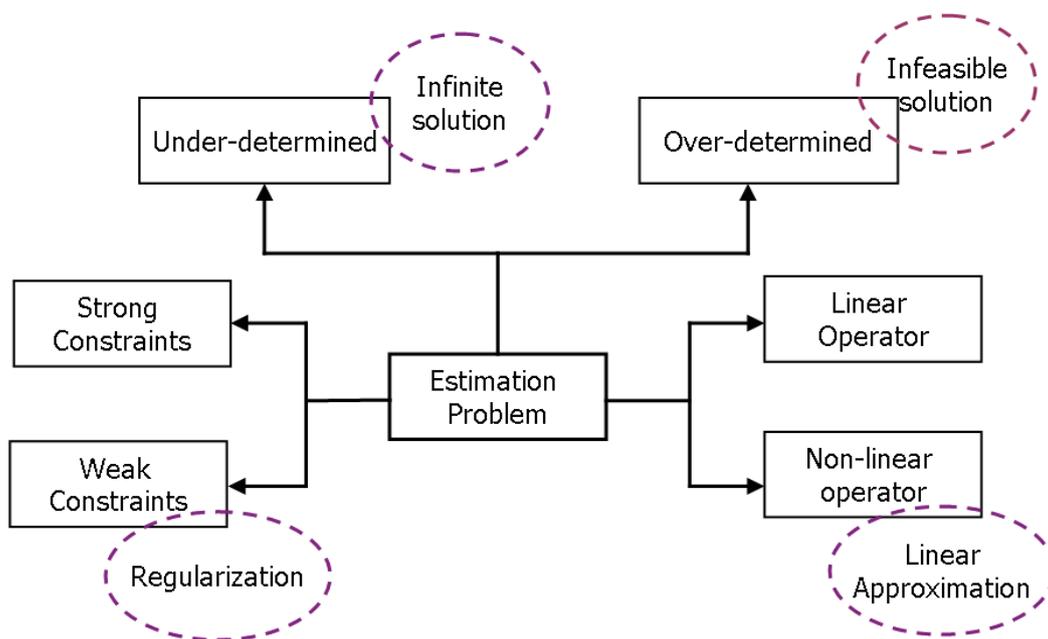
The present study addresses an overview of the inverse modelling technique utilised mainly for the emergency purposes. However, the inversion of large scale emissions and green house gases (GHG) are also discussed briefly. The details of the inversion techniques are included in Section 2 along with their applications carried out in real scenarios. Section 3 gives highlights on identification of multiple point releases. The advantages and limitations of these techniques are discussed in Section 4. Finally, the summary of review and future scope of the inverse modelling are highlighted in Section 5.

2   Inverse modelling techniques

The source information is estimated by minimising a cost function or functional that quantifies the difference between observed and predicted concentrations. Inverse modelling techniques have been categorised within the frameworks of statistical (or Bayesian) inference and optimisation techniques. Statistical methods provide source estimation along with the probability distribution function (pdf) under some confidence

levels and posterior statistics. Comparatively, optimisation techniques can give the simulation results which maximally match the measurement information. The inversion techniques, under these frameworks, are detailed in this section.

2.1 Bayesian inference method

Bayesian probabilistic inferential framework provides a natural means for incorporating both errors (model and observational) and prior (additional) information about the source. The prior information about source is improved to a posteriori when successive sets of concentration measurements are assimilated (e.g., Rodgers, 2000; Tarantola, 2005). The complete state of information available in the system is summarised by a suitably chosen likelihood function. The solution is derived in terms of posterior pdf (e.g., Tarantola, 2005) based on available concentration measurements and a priori information. Following the Bayes theorem, the posterior pdf is written in terms of prior, concentration measurements and likelihood pdf as,

$$P\left(\sigma \mid \mu, \mathbf{I_b}\right) = \frac{P\left(\mu, \sigma \mid \mathbf{I_b}\right)}{P\left(\mu \mid \mathbf{I_b}\right)} = \frac{v\left(\sigma \mid \mathbf{I_b}\right) P\left(\mu \mid \sigma, \mathbf{I_b}\right)}{P\left(\mu \mid \mathbf{I_b}\right)},$$
(2)

where $\mathbf{I_b}$ is background information about concentration measurements and source, $v(\sigma \mid \mathbf{I_b})$ is prior pdf, $P(\mu \mid \sigma, \mathbf{I_b})$ is likelihood function used to quantify the probability of the discrepancy between the concentration measurements and corresponding model predictions. The posterior probability $P(\sigma \mid \mu, \mathbf{I_b})$ describes the source information. The denominator term $P(\mu \mid \mathbf{I_b})$ is pdf of measurements μ and does not depend on σ, and thus, can be considered a constant factor in the inverse problem. The sampling algorithms [e.g., Markov chain Monte Carlo (MCMC), Metropolis-Hastings, etc.] are utilised to approximate the source (or release parameters) from the posterior pdf by drawing random samples (Gilks et al., 1996). The retrieved source is said to be well resolved if the posterior pdf is significantly different from the prior pdf (Tarantola, 2005). The theory is traditionally formulated in a discrete form (Cohn, 1997).

Let σ_b be a background emission vector, obtained from a short range forecast or assumed as constant and $\Delta\sigma = \sigma - \sigma_b$ is the corresponding background error vector. Let $P(\varepsilon)$ be the joint pdf of total error due to measurements and model representativeness and $\mu_b = \mathbf{L}\sigma_b$ is the background measurement vector. The innovation $\Delta\mu = \mu - \mu_b$ is the departure of the concentration measurements from its corresponding predicted value (using background estimates). The substitution allows to obtain as $\Delta\mu = \mathbf{L}\Delta\sigma + \varepsilon$. The $\Delta\sigma$ and ε. are assumed to be independent. In case of under-determined problems, a priori statistical knowledge of the background error $\Delta\sigma$ is introduced with a prior pdf $v(\Delta\sigma)$. Generally, $P(\varepsilon \mid \Delta\mu)$ and $P(\Delta\sigma \mid \Delta\mu)$ are assumed Gaussian with measurement error covariance matrix $\mathbf{Q}$ (dimension $m \times m$) and background error covariance matrix $\mathbf{B}$ (dimension $N \times N$) respectively. Now, the posterior pdf is obtained as,

$$P\left(\Delta\sigma \mid \Delta\mu\right) = \frac{e^{-J(\Delta\sigma)}}{C_0},$$

$$J(\Delta\sigma) = \frac{1}{2}\Delta\sigma^{\mathrm{T}}\mathbf{B}^{-1}\Delta\sigma + \frac{1}{2}\left(\mathbf{L}\,\Delta\sigma - \Delta\mu\right)^{\mathrm{T}}\mathbf{Q}^{-1}\left(\mathbf{L}\,\Delta\sigma - \Delta\mu\right)$$
(3)

where C_0 [equation (3)] is a normalising coefficient and the superscript 'T' denotes transpose. In equation (3), the term $\Delta\sigma^T B^{-1} \Delta\sigma$ corresponds to the departure of source from its background estimates and $(L\Delta\sigma - \Delta\mu)^T Q^{-1} (L\Delta\sigma - \Delta\mu)$ represents the departure from measurement innovations to model predicted innovations. The classical estimate $\Delta\sigma^{est}$ can be computed as,

$$\Delta\sigma^{est} = \left(B^{-1} + L^T Q^{-1} L\right)^{-1} L^T Q^{-1} \Delta\mu = BL^T \left(LBL^T + Q\right)^{-1} \Delta\mu, \qquad (4)$$

where the matrix $K = BL^T(LBL^T + Q)^{-1}$ is often referred as the gain matrix. The solution of equation (4) represents a posteriori maximum likelihood (minimum variance) estimate of the true state of the source. Observation and background errors are assumed to be uncorrelated.

Khemka et al. (2006), Yee (2007) and his co-workers (Keats et al., 2007a, 2007b) implemented Bayesian inference with MCMC sampling for simultaneous determination of all source parameters (conservative and non-conservative scalars) in simple (unobstructed) and complex (e.g., urban) terrain. Zhao and Nehorai (2007) developed a distributed sequential Bayesian estimation method for localising a diffusive source using prior defined as linear combination of polynomial Gaussian density functions. This provides a faster convergence and requires lesser number of sensors in comparison to the Gaussian approximation for the source reconstruction. Source reconstruction in a complex terrain, involving disturbed wind fields and spatial inhomogeneity, is shown by Yee et al. (2008). Senocak et al. (2008) have extended Bayesian inference with MCMC to estimate wind field parameters along with source parameters. Chow et al. (2008) performed source identification using Bayesian inference in urban environments using building resolving simulations.

2.1.1 Maximum entropy on mean

The maximum entropy on mean (MEM) method relies on minimisation of the entropy (information) content that separates the posterior pdf of the source and the errors from the prior pdf (Bocquet, 2005a, 2005b). A duality principle is utilised in minimisation of the entropy function by transforming the problem from infinite dimensional parameter space to finite dimensional observation space.

The total amount of information contained in the source distribution is expressed in terms of an entropy function, called as Kullback entropy function, written as (Bocquet, 2005a),

$$-S = \sum_\sigma P(\sigma | \mu, I_b) \ln\left[\frac{P(\sigma | \mu, I_b)}{v(\sigma | \mu, I_b)}\right], \qquad (5)$$

where S is a vector in the vector space of maximum of entropy functional. Since all the sources compatible with the observations can be written as a linear combination of source and adjoint functions, thus a constraint is defined as (Bocquet, 2005a),

$$\mu = \sum_\sigma P(\sigma | \mu, I_b) L\sigma. \qquad (6)$$

The Kullback entropy is modified by introducing the Lagrange multipliers β_i in the entropy function as (Bocquet, 2005a),

$$-\mathbf{S} = \sum_{\sigma} P(\sigma \mid \mu, I_b) \ln \left[\frac{P(\sigma \mid \mu, I_b)}{v(\sigma \mid \mu, I_b)} \right] + \beta^{\mathrm{T}} \left(\mu - \sum_{\sigma} P(\sigma \mid \mu, I_b) \mathbf{L} \sigma \right). \tag{7}$$

Now using the duality, the problem is reduced in to a function minimising the second entropy (Bocquet, 2005a),

$$\psi = \ln z(\beta) - \beta^{\mathrm{T}} \mu \quad \text{with} \quad z(\beta) = \sum_{\sigma} v(\sigma \mid \mu, I_b) \exp(\beta^{\mathrm{T}} \mathbf{L} \sigma), \tag{8}$$

where $z(\beta)$ is called the partition function. Hence, the source estimate can be obtained by estimating β and pdf, P. The estimator of the source is the average source defined as $\bar{\sigma}$. Alternatively, one can also choose the maximum likelihood estimate of the source as (Bocquet, 2005a),

$$\sigma^{ML} = \arg\min_{\sigma} P(\sigma \mid \mu, I_b) = \arg\min_{\sigma} \left[\ln v(\sigma \mid \mu, I_b) + \beta^{\mathrm{T}} \mathbf{L} \sigma \right]. \tag{9}$$

The application of the method is performed for retrieving the temporal profile of emission with the knowledge of location of emission sources (Bocquet, 2005a, 2005b, 2005c) in European Tracer Experiments [ETEX]-I. The method was extended by Krysta and Bocquet (2007) by adding statistically consistent diagnostic tools, such as convexity, etc., for improving the posteriori and quality of the reconstruction. However in case of Algeciras incident (southern Spain in May 1998), source reconstruction was not successful. The method was then successfully applied to the real data of the ETEX-I (Bocquet, 2007; Krysta and Bocquet, 2007) and ETEX-II (Krysta et al., 2008) with a view to a high-resolution reconstruction of the source without any assumption about the source location. The source was correctly localised in both ETEX-I and ETEX-II. However, the total released mass in ETEX-II was under-estimated, which is attributed to the model representativeness error. Similarly, the method is evaluated for Chernobyl accident (Davoine and Bocquet, 2007). Bocquet (2005c) also explained the origin of the singular behaviour of the adjoint functions at measurement time and locations when grid resolution increases. In general, the source estimation under Bayesian inference (through maximum likelihood or maximum a posteriori) is sensitive to the choice of a priori distribution for the release parameters (or discretised source elements). Bocquet (2008) compared the source retrieval for the choice of a prior distribution as Gaussian and non-Gaussian (semi Gaussian and Bernoulli). Non-Gaussian scheme is meant to handle prior information that cannot always translate onto mean and second-order moments, such as the positivity of the pollutant source which is always the case in a Gaussian framework.

2.2 *Variational minimisation of the cost function*

Variational minimisation of the cost function produces an optimal estimate of the unknown source at analysis time through iterative minimisation of the prescribed cost function (Lewis et al., 2006). In this section, these are presented by categorising into *unregularised* (for over-determined problems; no constraints of background information) and *regularised* using constraints in terms of background information (for under-determined problems). The regularisation transforms an ill-posed problem into

a well-posed problem by imposing constraints based on a priori information and thus, leads to a stable solution of the inverse problem.

2.2.1 Unregularised (classical least-squares)

Classical least-squares estimation is based on minimising sum of squares of the residuals between measured and model predicted concentrations (Lewis et al., 2006). The cost function J can be written as (in matrix notations)

$$J = [\boldsymbol{\mu} - \mathbf{L}\boldsymbol{\sigma}]^{\mathrm{T}} \mathbf{Q}^{-1} [\boldsymbol{\mu} - \mathbf{L}\boldsymbol{\sigma}], \tag{10}$$

where $\mathbf{Q}$ is a symmetric and positive-definite weight matrix of size $m \times m$. The choice of weight matrix depends on statistical knowledge of measurement errors. Here, measurement errors account for both model representativeness and observation errors. The least-squares solution is given as $\hat{\boldsymbol{\sigma}} = (\mathbf{L}^{\mathrm{T}}\mathbf{Q}^{-1}\mathbf{L})^{-1}(\mathbf{L}^{\mathrm{T}}\mathbf{Q}^{-1}\boldsymbol{\mu})$. The least-squares method is applicable only for an over-determined ($m > N$, no exact solutions) inverse problem. The iterative minimisation of function J [equation (10)] requires an initial guess of source term and an incorrect initialisation may lead to an inadequate solution.

Keller and Herrnberger (1997) estimated source strength of a point source in complex dispersion conditions. For a short-range experiment (based on wind tunnel experiment), Krysta et al. (2006) implemented least-squares to retrieve the source strength, wind and dispersion parameters from the concentration measurements. Roberti et al. (2005); Storch et al. (2005, 2007) and Reen and Stauffer (2010) extended applicability of the least-squares for estimating the boundary layer parameters from the concentration measurements. Matthes et al. (2005) and Lushi and Stockie (2010) applied least-squares to estimate the source-strengths of two and four simultaneous releases respectively. Sharan et al. (2012b) proposed two-step algorithm, free from initial guess, to retrieve the location and strength of a point source at local scale. The algorithm is further extended in a least-squares framework to identify the multiple-point releases (Singh et al., 2013). The sensitivity of this algorithm to source reconstruction is discussed by Singh and Rani (2014).

2.2.2 Regularised using weights: minimum-norm weighted least-squares (also called 'renormalisation', Issartel et al., 2007)

The technique is a strategy for linear assimilation of concentration measurements to identify the unknown releases (Issartel, 2003; 2005). The method exploits the natural statistics provided by the geometry of the monitoring network. These statistics are expressed in the form of a weight function derived by using a minimum entropy criterion. This criterion prevents the over-estimation of the available information that would lead to the artefacts especially close to the detectors. These weight functions also serve as a priori information about release apparent to the monitoring network and provide regularisation. The theory utilises the adjoint source-receptor relationship and constructs a source estimate, among vector space of acceptable sources, which describes the possible distribution of the emission sources (Issartel et al., 2007). The method is applicable for both over-determined as well as under-determined problems.

Equation (1) is modified by introducing a weight matrix $\mathbf{W}$ of dimension $N \times N$, such that $\boldsymbol{\mu} = \mathbf{L_w W \sigma}$ where $\mathbf{L_w} = [\ell_{jj} / w_{jj}]$ is the modified sensitivity matrix. The weight matrix is purely diagonal and satisfy following properties:

1    $w_{jj} > 0$

2    $\displaystyle\sum_{j=1}^{N} w_{jj} = m$

3    $diag(\mathbf{L_w^T \, H_w^{-1} \, L_w}) \equiv 1$, where $\mathbf{H_w = L_w W \, L_w^T}$ is the Gram matrix of the weighted sensitivity functions.

The weight functions are computed iteratively using algorithm proposed by Issartel (2005). A minimum norm weighted solution is obtained as

$$\widehat{\boldsymbol{\sigma}} = \mathbf{L_w^T \, H_w^{-1} \, \boldsymbol{\mu}} \tag{11}$$

Equation (11) provides an estimate for the distributed emissions and is seen as a generalised inverse solution to the under-determined class of linear inverse problems (Turbelin et al., 2014).

The technique is utilised by Issartel (2003) to reconstruct the source in case of ETEX-I experiment. Further, this technique is extended and evaluated by Issartel et al. (2007) for identifying the a real sources over Indian region using synthetic satellite measurements. Sharan et al. (2009) extended this technique for identification of a point release using the fact that maximum of the source estimate $\widehat{\boldsymbol{\sigma}}$ will coincide with the location of the release. An evaluation is shown using real concentration measurements from IIT diffusion experiment in convective conditions. Later, Sharan et al. (2012a) proposed an extension of the renormalisation technique for identification of an elevated release with an inversion error estimate. The equivalence of renormalisation technique to optimally weighted least-squares is shown by Issartel et al. (2012). Turbelin et al. (2014) and Singh et al. (2014, 2015) discussed the optimal localisation properties for the retrieved source.

2.2.3 Regularised using constraints

This section addresses the transformation of an under-determined inverse problem to well-posed by adding a priori information, in terms of constraints, to the cost function. Accordingly, the function J [equation (10)] is modified as

$$J = \left[\boldsymbol{\mu} - \mathbf{L}(\boldsymbol{\sigma})\right]^T \mathbf{Q}^{-1} \left[\boldsymbol{\mu} - \mathbf{L}(\boldsymbol{\sigma})\right] + \lambda \phi(\boldsymbol{\sigma}), \tag{12}$$

where $\lambda\phi(\boldsymbol{\sigma})$ is the penalty term, which can protect the model from being overly fitted or is useful for sparse data fitting. $\phi(\boldsymbol{\sigma})$ is the regularisation functional. The regularisation parameter λ imposes a relative weight between the residual and constraint. The choices of both λ and $\phi(\boldsymbol{\sigma})$ are subjective. Commonly, the value of λ is chosen between 0 and 1.

Seibert (2000) and Seibert and Stohl (2000) used a regularisation functional, of the form $\phi(\boldsymbol{\sigma}) = \|\boldsymbol{\sigma}\|^2$ (which imposes an upper bound on the energy of the source distribution $\boldsymbol{\sigma}$), to reconstruct the distribution of emission rates for a CTBT radionuclide monitoring system (Hourdin and Issartel, 2000). Using Tikhonov regularisation and the properties of

L-curve (Hansen and O'Leary, 1993), Kathirgamanathan et al. (2004) estimated emission rate of a non-steady point source. Akecelik et al. (2003) mentioned that Tikhonov regularisation method works well for smooth sources whereas for discontinuous or point sources, a poor solution is obtained leading to oscillations in the vicinity of the sources. Chen and Li (2008) presented a ℓ_p-regularised least-squares method to estimate the location and release rate of single as well as multiple-point sources of instantaneous type or continuous release. Source characterisation by combining the stochastic search and regularised gradient optimisation is addressed by Addepalli et al. (2011).

When background estimate or information is available, the cost function can be formulated as weighted sum of squares of residuals to the observations and background estimates (LeDimet and Talagrand, 1986). In this framework, variational minimisation has two variants:

1 three-dimensional variational assimilation (3D-Var)

2 four-dimensional variational assimilation (4D-Var).

3D-Var performs the analysis within the assimilation window for fixed time and varying spatial dimensions and 4D-Var is an extension of 3D-Var in which observations at times before and after the analysis time are included at varying space and time.

2.2.3.1   3D-Var

The cost function formulated in 3D-Var is similar to equation (3) shown in Bayesian inference (Section 2.1). Using the preconditioning of background error covariance matrix $(\mathbf{B} = \mathbf{U}\mathbf{U}^\mathrm{T})$ such that $\Delta\boldsymbol{\sigma} = \mathbf{U}\mathbf{v}$, equation (3) can be written as (Courtier, 1997)

$$J(\mathbf{v}) = \frac{1}{2}\mathbf{v}^\mathrm{T}\mathbf{v} + \frac{1}{2}\left(\mathbf{L}\,\mathbf{U}\mathbf{v} - \Delta\boldsymbol{\mu}\right)^\mathrm{T}\mathbf{Q}^{-1}\left(\mathbf{L}\,\mathbf{U}\mathbf{v} - \Delta\boldsymbol{\mu}\right) \tag{13}$$

In 3D-VAR, the minimisation of the cost function [equation (13)] using optimisation algorithms (steepest descent, conjugate gradient, etc.) requires derivative information with respect to the control variable. The derivative of J with respect to $\mathbf{v}$ is obtained as

$$\nabla J = \left(\mathbf{I} + (\mathbf{L}\mathbf{U})^\mathrm{T}\mathbf{Q}^{-1}(\mathbf{L}\mathbf{U})\right)\mathbf{v} - \mathbf{U}^\mathrm{T}\mathbf{L}^\mathrm{T}\mathbf{Q}^{-1}\Delta\boldsymbol{\mu}. \tag{14}$$

2.2.3.2   4D-Var

4D-VAR is applicable when the observations are distributed over a time interval $[t_0, t_n]$ (t_0 is the initial time and t_n is the final time) and the source is varying with respect to the time. Therefore, the cost function [equation (13)] is modified by including time dimension as (Elbern et al., 2007)

$$J\left[\mathbf{v}(t_i)\right]$$
$$= \frac{1}{2}\sum_{i=1}^{n}\mathbf{v}(t_i)^\mathrm{T}\mathbf{v}(t_i) \tag{15}$$
$$+ \frac{1}{2}\sum_{i=1}^{n}\left[\mathbf{H}(t_i)\mathbf{M}(t_i)\mathbf{U}\mathbf{v}(t_i) - \Delta\boldsymbol{\mu}(t_i)\right]^\mathrm{T}\mathbf{Q}_i^{-1}\left[\mathbf{H}(t_i)\mathbf{M}(t_i)\mathbf{U}\mathbf{v}(t_i) - \Delta\boldsymbol{\mu}(t_i)\right]$$

The operator $\mathbf{L}$ is decomposed into observation ($\mathbf{H}$) and model operator ($\mathbf{M}$) as $\mathbf{L}\sigma = \mathbf{H}\mathbf{M}\sigma$. The model operator is defined as

$$\sigma(t_{k+1}) = \mathbf{M}_{k+1}\sigma(t_k) = \mathbf{M}_{k+1}\mathbf{M}_k \ldots \mathbf{M}_1\sigma(t_0) \tag{16}$$

The derivative of equation (15) is written as

$$\frac{\partial J}{\partial \mathbf{v}(t_i)} = \mathbf{v}(t_i) + \sum_{i=0}^{n} \mathbf{U}^{\mathrm{T}}\mathbf{M}_{i+1}^{\mathrm{T}}\,\mathbf{H}(t_i)^{\mathrm{T}}\,\mathbf{Q}_i^{-1}\Big[\mathbf{H}(t_i)\mathbf{M}(t_i)\mathbf{U}\mathbf{v}(t_i) - \Delta\boldsymbol{\mu}(t_i)\Big] \tag{17}$$

in which $\mathbf{M}_{i+1}^{\mathrm{T}}$ is called an adjoint model. Using adjoint model, the model state is adjusted towards the observations in an iterative process, starting with an initial guess. ∇J is determined by running the forward model first, and then running the linearised adjoint model backward in time with forcing by the model misfit. All the intermediate values in the forward run have to be stored for their use in the adjoint run (Rao, 2007).

At regional scale, Robertson and Persson (1993) applied four dimensional data assimilation of radiological data for estimating the source intensity using the adjoint technique. Later, source estimation was evaluated with ETEX experiment (Robertson and Langner, 1998). Penenko et al. (2002) discussed methods of sensitivity theory and inverse modelling for estimation of source parameters based on variational principles and adjoint equations. A fast 4D-Var scheme was developed by Bocquet (2012) to retrieve the large parameter fields that are nonlinearly related to concentrations and its application was investigated to Chernobyl accident. The 3D-Var and 4D-Var have been extensively applied for estimating large scale emissions and sources and sinks of GHG using surface and satellite measurements. Several studies are performed for the estimation of emission rates of the trace gases such as CH_4 (Meirink et al., 2008; Bergamaschi et al., 2010), CO_2 (Bousquet et al., 1999; Houweling et al. 1999, 2004), O_3 (Elbern and Schmidt, 2001; Elbern et al. 2007), NOx (Quelo et al. 2005; Muller and Stavrakou, 2005), etc. using their satellite measurements in the atmosphere.

2.3 *Search algorithms for minimising cost function: GA*

Genetic algorithm (GA) is a random search-based optimisation technique, built on a process of natural evolution (e.g., Goldberg, 1989; Haupt et al., 2006). In GA, a set of random population is assumed for the unknowns, which correspond to a set of possible solutions of the cost function (regularised or unregularised). The initial choice of population evolves towards better solutions iteratively. During iterations, every individual of the population is numerically assessed through the objective function and multiple individuals are stochastically selected from the current population (based on their closeness to the measurements). Repeated iterations minimise the value of cost function and provide an estimate for the source term that best accounts for the measured concentrations.

The source identification using GA was investigated by Cartwright and Harris (1993). Haupt et al. (2006) estimated the source emission rate using synthetic data and computed error bounds for the resulting estimates of the calibration factors using a Monte Carlo technique. Thomson et al. (2007) investigated the effect of three different regularisation functionals on source retrieval using a simulated annealing optimisation. Allen et al. (2007a) estimated simultaneously the surface wind direction and the pollutant source

characteristics. Allen et al. (2007b) performed the identification of multiple-point releases by estimating their time of release, source locations, and strengths. Haupt et al. (2009) attempted to improve the wind field in a meandering situation using GA. In an industrial application, Khlaifi et al. (2009) evaluated source identification for three industrial sources using hourly measured SO_2 concentration measurements. Recently, Long et al. (2010) estimated the meteorological parameters (surface wind direction, surface wind speed) in addition to release parameters (location, height, emission rate and time) using concentration measurements. A concept is also proposed for quantifying minimum number of required sensors in source identification. A Lagrangian entity backtracking approach is discussed by Annunzio et al. (2012a) to determine the contaminant source information.

2.4 Kalman filter

The Kalman filter (KF) is a recursive parametric estimation technique for time dependent problems and provides best linear unbiased estimate (minimum variance) of the unknown state. The assimilation process in KF utilises the model predicted and measured concentrations at various times to produce a linear unbiased optimal (minimum variance) estimate $\hat{\sigma}_t$ (Kalman, 1960; Kalman and Bucy, 1961). The formulations of the KF (including analysis and update steps) are given as (Hartley and Prinn, 1993; Cohn, 1997)

- *analysis step*:

$$\sigma_t^a = \sigma_t^f + \mathbf{K}_t \left(\mu_t - \mathbf{L}_t \sigma_t^f \right), \tag{18a}$$

- *update step*:

$$\mathbf{K}_t = \mathbf{B}_t^f \mathbf{L}_t^{\mathrm{T}} \left(\mathbf{L}_t \mathbf{B}_t^f \mathbf{L}_t^{\mathrm{T}} + \mathbf{Q}_t \right)^{-1} = \mathbf{B}_t^a \mathbf{L}_t^{\mathrm{T}} \mathbf{Q}_t^{-1}$$

$$\text{where } \mathbf{B}_t^a = \left(\mathbf{I} - \mathbf{K}_t \mathbf{L}_t \right) \mathbf{B}_t^f. \tag{18b}$$

The subscript '*t*' denotes the time, superscripts '*a*' and '*f*' denote analysis and forecast respectively. The matrix $\mathbf{K}$ is called Kalman gain and $\mathbf{B}^a$ is the analysis (or posteriori) error covariance matrix.

At urban scale, Mulholland and Seinfeld (1995) obtained emission adjustment factors for CO emissions in the South Coast Air Basin of California. Haas-Laursen et al. (1996) used KF to deduce regional emission for chlorofluoro carbon (CFC-11) in a global chemical transport model (Hartley and Prinn, 1993). For estimating halocarbon or methane sources, Hartley and Prinn (1993) and Chen and Prinn (2006) used a global chemistry transport model and KF. Drews et al. (2004, 2005) utilised KF for source term estimation in case of short-range atmospheric dispersion of radioactive materials using off-site radiation monitoring data. Jorquera and Castro (2010) performed an inverse modelling to analyse the urban pollution episodes.

3 Identification of multiple-point releases at urban scales

This refers to a particular context when several point releases have occurred at various locations simultaneously or at different times. The tracers released from several point

sources could be the same or different. This is also considered as important due to security applications [e.g., FFT07 trials (Storwald, 2007)]. Such releases raise the issue of determining

1   number of unknown releases occurred

2   their parameters (locations, height, strengths, release times in case of point sources).

The ambiguity persists as a small source close to the measurement location may give the same concentration as a large source further away. When same tracer is released from various sources, the sampled concentrations will be obtained as merged concentrations. In addition, the sources may be located at the same angle from the receptor but at different distances. The geometry of monitoring network, meteorological conditions, limited concentration measurements, model errors, chemistry involved, etc., are the other factors which limit the release identification.

Yee (2007) proposed a Bayesian approach for identifying the multiple-point releases. Allen et al. (2007b) performed the simultaneous estimation of multiple sources using measurements from a dipole pride (DP) 26 field experiments. Aerodyne Research Inc, Massachusetts, USA, developed an inverse modelling system (based on search algorithm), aerodyne inverse modelling system (AIMS) (Albo et al., 2011) and shown an evaluation by estimating the source term parameters of multiple pollutant sources in FFT07 experiments. Lushi and Stockie (2010) described an inverse Gaussian plume approach for estimating the emission rates of four-point sources in a large lead-zinc smelting operation in Trail, British Columbia. Sharan et al. (2012b) have proposed a two-step least-squares approach for identifying multiple releases using pseudo-real data. Singh et al. (2013) have proposed a weighted least-squares method for multiple-point release identification.

In the above mentioned studies, the multiple-point source estimation is based on the fact that number of releases is known. First, Yee (2008) relaxed this assumption by proposing a reversible jump Markov process method and successfully performed identification of multiple-point sources using noisy synthetic data. Later Yee (2012) proposed a new Bayesian inference technique for the general source reconstruction. Annunzio et al. (2012b) introduced the multi-entity field approximation (MEFA) method for cases involving one or more ground-level point sources. Wade and Senocak (2013) proposed a composite ranking system, based on error (scatter), bias and correlation component, to estimate the number of sources.

4   Issues/limitations in inversion

The inverse problem posed for identifying unknown releases is ill-posed. The estimation process is sensitive for measurement noise, model errors, number of concentration measurements, a priori information, grid discretisation, etc. In under-determined systems, poor conditioning of the matrices cause instabilities in the solution. In particular, representation of point sources poses a big challenge and strong gradient often leads to model representativeness errors (Bocquet, 2005a). In case of multiple-point releases, the complexity of the source retrieval increases with increasing number of releases.

In Bayesian inference, the choice of a priori information is not unique and this, further affects the accuracy of the retrieval. The determination of background and

measurement statistics (or error covariance matrices) is difficult to determine since only limited and averaged concentrations are measured discretely and a history of tracer concentrations may not be available. The use of sampling algorithms can also be tedious and time consuming. In statistical inference methods, there are two important and sensitive parameters in the reconstruction,

1   prior mass scale

2   prior error variance (Bocquet, 2005a).

The prior information in inverse modelling cannot always translate into mean and second-order moments, especially, when specific prior hypotheses on the sources are taken into account such as positivity or boundedness, etc. This leads to non-quadratic cost functions because the underlying assumptions are not Gaussian. However, the minimisation techniques require an initial guess of the releases and their final solution depends on the choice of initial guess. Also, derivative information is required for the use of gradient-based minimisation techniques. GA is associated with a drawback that the estimated solution may not be globally optimal. The KF is associated with limitations in its practical implementation due to sampling errors, spurious correlations, filter divergence, inflation and covariance localisation issues.

The measured concentrations are sampled at distinct point locations for a fixed amount of time and their time average is considered as measured average concentrations. The adjoint models are utilised to describe the potential sensitivity of the discrete cells (unknown emission) with respect to the measurements in the discretised space and time. Since the atmospheric transport is diffusive, strong numerical gradients arise around the discrete cells containing receptors and this leads to the local peaks in the sensitivity matrix (Houweling et al., 2004; Bocquet, 2005c; Issartel et al., 2007). Due to these peaks at measurement cells, the inverse solution is forced to have a maximum only the measurement cells and thus, the source retrieval is not adequate. Also, limited information is provided by the discretely measured concentrations, and thus, the resolution of the retrieved source is limited and relatively lower for a remote region. The measurement errors are assumed commonly as uncorrelated, which certainly may be prevailing with complex correlations in representativeness errors (Issartel et al., 2007).

The choice of a cost function (measuring goodness of a fit), regularisation functional and suitable ATDM also plays an important role in accuracy of the source estimation. The choice of an ATDM depends on the site description, terrain features, nature of release, chemical properties of tracer, availability of meteorological information and turbulence parameters, distance travelled by plume, etc. This affects the computational time associated with the inversion. The arrangement and distribution of monitoring network stations to monitor the unknown releases is also an emerging problem which is responsible for a fast identification and minimising the computational requirements (Korsakissok et al., 2010). This includes both the number of samplers required to identify the releases as well as the design criteria under which the major characteristics of the plume can be captured (Issartel and Gamel, 2010). Furthermore, the availability of dispersion datasets is desired to assess the capability and shortcomings of the inversion algorithms.

5 Summary

The importance of fast and accurate source identification is considered in security and emergency assessment programs. Owing to ill-posedness, limited networking of receptors, limited concentration measurements, model errors, nonlinearity, lack of a priori information, etc., it is difficult to retrieve the source information accurately even if the meteorological information and concentration measurements are provided as error free. These limitations also affect the resolution of the source retrieval. In this review, inversion techniques are discussed based on Bayesian, optimisation and KF framework. These techniques are followed by some advantages and disadvantages and thus, cannot be considered as optimal in all the scenarios without analysing further evaluations and comparisons.

Bayesian inference methods allows to assess (and also incorporate) a priori information about the release and the error statistics. Also, the posterior uncertainty in the source retrieval can be analysed. When first and second moment of the errors and prior information such as positivity, boundedness, etc. are known, the MEM approach can perform significantly better than the optimisation approaches based on minimising a cost function built on a quadratic criterion (Gaussian prior laws for the source and the modelled errors). The method ignores higher order moments and thus, is not believed to perform significantly better than a variational inversion, where first and second moments are already used.

The inversion techniques under optimisation framework provide a source term which is best fitted to the concentration measurements. The minimum norm weighted inversion technique overcomes the requirement of derivative information and background statistics, but requires adjoint source-receptor sensitivities. A key point of this technique is to utilise the natural statistics of the adjoint functions in the form of weight functions and to use them for regularising the inverse problem (Issartel et al., 2007). The proposed Gram matrix provides optimal discrimination to the measurements and thus, can also be considered as measurement covariance matrix in the absence of error statistics (Issartel et al., 2012). The use of search algorithm is advantageous in order to overcome the derivative and covariance statistics requirements and, can deal with large set of parameters. Several other variations of the GA or coupled hybrid GA are proposed like simulated annealing (Thomson et al., 2007), particle swarm optimisation, ant colony optimisation, etc. These algorithms require a further investigation on source retrieval applications (Zheng and Chen, 2010). KF methods are most suited for linear systems with well-conditioned covariance matrices and strong observable relations between the internal state variables and the model outputs (Rao, 2007). There exists several variations of KF called as ensemble KF (EnKF) (Evensen, 2003), ensemble transform KF (ETKF) (Bishop et al., 2001), extended KF (EKF) and unscented KF (UKF) (Redwood, 2011).

Table 1   Advantages and limitations of the inversion methods

Theoretical foundation	Method	Advantages		Limitations	
Statistical inference	Bayesian	1	Estimate the parameters with maximum likelihood probability.	1	Measurement error and background statistics are required a priori and based on assumptions.
		2	Posterior uncertainty in the estimates can be analysed.	2	Use of sampling algorithms is time consuming.
				3	Unrealistic influences arise in the inversion of continuous emissions around observation sites if adjoint source-receptor sensitivity is utilised.
	Maximum entropy on mean	1	Entropy of observations is maximised.	1	Error statistics are required a priori.
		2	Non-Gaussian priors can be accounted.		
	Kalman filter	1	Minimises the variances of the estimates.	1	Requires measurement error and background statistics.
		2	Posterior uncertainty can be determined.	2	Applicable to linear systems with time varying situations.
				3	Sampling errors and spurious correlations lead to filter divergence.
				4	Tuning of inflation and model errors is required.
Optimisation	4D VAR	1	Observations can be assimilated both in space and time.	1	Measurement error and background statistics are required a priori.
		2	Estimates maximally match the observation.	2	Computationally expensive and time consuming.
		3	Derivatives can be computed efficiently with the use of adjoint models.		
	3D VAR or regularised functional	1	Estimates maximally match the observation.	1	Measurement error and background statistics are required a priori.
				2	Derivative information is required.
				3	Observations are assimilated only in space.
				4	Regularisation parameter is not known.
	Minimum weighted norm	1	Minimises the entropy of the source function.	1	Adjoint source-receptor sensitivity is required.
		2	Error statistics are not required.	2	Only linear source reconstruction is possible.
		3	A priori information about releases is determined through the geometry of the monitoring network.		
		4	Weight functions are proposed to deal with the unrealistic influences due to observation sites.		
	Search-based optimisation	1	Derivative information is not required.	1	Locally optimum solution may arise.
		2	A population of solution is obtained.	2	Time consuming.

16

To facilitate the inverse modelling of unknown releases, various sensor-based monitoring networks are deployed by the government agencies, for instance, Defense Threat Reduction Agency (DTRA), US Environmental Protection Agency (USEPA), National Atmospheric Release Advisory Center (NARA), Air Resources Laboratory (ARL), Comprehensive Test Ban Treaty (CTBT), etc. The inverse modelling is sensitive with respect to the observational network (Villani et al., 2010) which needs to be explored. This highlights the need for designing of an optimum network in order to persuade fast identification as well as reduction in the setup cost involved (Korsakissok et al., 2010). Few contributions in this direction are made by Abida et al. (2008) (addressing spatial design of monitoring network in France), Saunier et al. (2009) (principal component analysis-based optimal design of monitoring network), Wu et al. (2010) (network reductions), Wu and Bocquet (2011) (optimal redistribution of ozone monitoring stations over France), Koohkan et al. (2012) (potential of the international monitoring system radionuclide network), etc.

The applicability of inverse modelling requires a further assessment in complex terrain or complex meteorological situations observed in coastal or mountain/hill regions. Data fusion techniques are required to utilise information from multiple sensors and multiple dispersion models in order to evaluate the specific inferences. Advancement in the inversion is needed to minimise the dependence on a priori information about the release. The estimation of prior error and model error statistics is also helpful in making the inversion more efficient and accurate (Winiarek et al., 2012; Koohkan and Bocquet, 2012; Winiarek, 2014). The inverse modelling of moving sources (varying in space or time) is also not yet addressed and requires attention. Similarly, efforts are required to obtain and utilise the concentration measurements from mobile sampling (varying in space and time) for the inverse modelling studies. The availability of dispersion datasets (e.g., ETEX-I, II, 1994; Fusion Field Trials, 2007; Michelstadt, 2013) also helps in analysing and comparing the inversion techniques built on different frameworks. Therefore, real measurements from dispersion experiments are desired to assess the strength and weakness of inversion techniques.

References

Abida, R., Bocquet, M., Vercauteren, N. and Isnard, O. (2008) 'Design of a monitoring network over France in case of a radiological accidental release', *Atmos. Environ.*, Vol. 42, No. 21, pp.5205–5219.

Addepalli, B., Sikorski, K., Pardyjak, E.R. and Zhdanov, M.S. (2011) 'Source characterization of atmospheric releases using stochastic search and regularized gradient optimization', *Inverse Problems in Science and Engineering*, Vol. 19, No. 8, pp.1097–1124.

Akecelik, V., Biros, G., Draganescu, A., Ghattas, O., Hill, J. and Waanders, B.V.B. (2006) 'Inversion of airborne contaminants in a regional model', *Lecture Notes in Computer Science*, Vol. 3993, pp.481–488.

Albo, S.E., Oluwole, O.O. and Miake-Lye, R.C. (2011) 'The aerodyne inverse modelling system (AIMS): source estimation applied to the FFT 07 experiment and to simulated mobile sensor data', *Atmospheric Environment*, Vol. 45, No. 33, pp.6085–6092.

Allen, C.T., Young, G.S. and Haupt, S.E. (2007a) 'Improving pollutant source characterization by better estimating wind direction with a genetic algorithm', *Atmospheric Environment*, Vol. 41, No. 11, pp.2283–2289.

Allen, C.T., Haupt, S.E. and Young, G.S. (2007b) 'Source characterization with a genetic algorithm-coupled dispersion-backward model incorporating SCIPUFF', *Journal of Applied Meteorology and Climatology*, Vol. 46, No. 3, pp.273–287.

Annunzio, A.J., Young, G.S. and Haupt, S.E. (2012a) 'Utilizing state estimation to determine the source location for a contaminant', *Atmospheric Environment*, Vol. 46, pp.580–589.

Annunzio, A.J., Young, G.S. and Haupt, S.E. (2012b) 'A multi-entity field approximation to determine the source location of multiple atmospheric contaminant releases', *Atmospheric Environment*, Vol. 62, pp.593–604.

Bergamaschi, P., Krol, M., Meirink, J.F., Dentener, F., Segers, A., Van Aardenne, J., Monni, S., Vermeulen, A.T., Schmidt, M., Ramonet, M., Yver, C., Meinhardt, F., Nisbet, E.G., Fisher, R.R., O'Doherty, S. and Dlugokencky, E.J. (2010) 'Inverse modelling of European CH_4 emissions 2001–2006)', *Journal of Geophysical Research-Atmosperes*, Vol. 115, Art. No. D22309, No. D22.

Bocquet, M. (2005a) 'Reconstruction of an atmospheric tracer source using the principle of maximum entropy-I: theory', *Quarterly Journal of Royal Meteorological Society*, Vol. 131, No. 610, pp.2191–2208.

Bocquet, M. (2005b) 'Reconstruction of an atmospheric tracer source using the principle of maximum entropy-II: applications', *Quarterly Journal of Royal Meteorological Society*, Vol. 131, No. 610, pp.2209–2223.

Bocquet, M. (2005c) 'Grid resolution dependence in the reconstruction of an atmospheric tracer source', *Nonlinear Processes in Geophysics*, Vol. 12, No. 2, pp.219–234.

Bocquet, M. (2007) 'High-resolution reconstruction of a tracer dispersion event: application to ETEX', *Quarterly Journal of Royal Meteorological Society*, Vol. 133, No. 62, pp.1013–1026.

Bocquet, M. (2008) 'Inverse modelling of atmospheric tracers: non-Gaussian methods and second-order sensitivity analysis', *Nonlinear Processess in Geophysics*, Vol. 15, pp.127–143.

Bocquet, M. (2012) 'Parameter-field estimation for atmospheric dispersion: application to the chernobyl accident using 4D-Var', *Quarterly Journal of Royal Meteorological Society*, Vol. 138, No. 664, pp.664–681.

Bousquet, P., Ciais, P., Peylin, P., Ramonet, M. and Monfray, P. (1999) 'Inverse modelling of annual CO_2 sources and sinks, 1. Method and control inversion', *Journal of Geophysical Research-Atmospheres*, Vol. 104, No. D21, pp.26161–26178.

Cartwright, H.M. and Harris, S.P. (1993) 'Analysis of the distribution of airborne pollution using gas', *Atmospheric Environment*, Vol. 27A, No. 12, pp.1783–1791.

Chen, H. and Li, X.R. (2008) 'Source parameter estimation of atmospheric pollution using regularized least squares', *Proceedings of 11th International Conference on Information Fusion*, June 30–July 3, Cologne, Germany, pp.1–5.

Chen, Y. and Prinn, R.G. (2006) 'Estimation of atmospheric methane emissions between 1996 and 2001 using a three-dimensional global chemical transport model', *Journal of Geophysical Research-Atmospheres*, Vol. 111, No. D10, Art. No. D10307.

Chow, T., Kosovic, B. and Chan, S. (2008) 'Source inversion for contaminant plume dispersion in urban environments using building-resolving simulations', *Journal of Applied Meteorology and Climatology*, Vol. 47, No. 6, pp.1553–1572.

Cohn, S.E. (1997) 'An introduction to estimation theory', *Journal of Meteorological Society of Japan*, Vol. 75, No. 1B, pp.257–288.

Courtier, P. (1997) 'Dual formulation of four-dimensional variational assimilation', *Quarterly Journal of Royal Meteorological Society*, Vol. 123, No. 544, pp.2449–2461.

Davoine, X. and Bocquet, M. (2007) 'Inverse modelling based reconstruction of the chernobyl source term available for long-range transport', *Atmospheric Chemistry and Physics*, Vol. 7, No. 6, pp.1549–1564.

Drews, M., Lauritzen, B. and Madsen, H. (2005) 'Analysis of Kalman filter based method for on-line estimation of atmospheric dispersion parameters using radiation monitoring data', *Radiation Protection Dosimetry*, Vol. 113, No. 1, pp.75–89.

Drews, M., Lauritzen, B., Madsen, H. and Smith, J.Q. (2004) 'Kalman filtration of radiation monitoring data from atmospheric dispersion of radioactive materials', *Radiation Protection Dosimetry*, Vol. 111, No. 3, pp.257–269.

Elbern, H. and Schmidt, H. (2001) 'Ozone episode analysis by four-dimensional variational chemistry data assimilation', *Journal of Geophysical Research-Atmospheres*, Vol. 106, No.D4, pp.3569–3590.

Elbern, H., Strunk, A., Schmidt, H. and Talagrand, O. (2007) 'Emission rate and chemical state estimation by 4-dimensional variational inversion', *Atmospheric Chemistry and Physics*, Vol. 7, No. 14, pp.3749–3769.

Enting, I.G. (1985) 'A classification of some inverse problems in geochemical modelling', *Tellus*, Vol. 37B, Nos. 4–5, pp.216–229.

Evensen, G. (2003) 'The ensemble Kalman filter: theoretical formulation and practical implementation', *Ocean Dynamics*, Vol. 53, No. 4, pp.343–367.

Gilks, W.R., Richardson, S. and Spiegelhalter, D.J. (1996) *Markov Chain Monte Carlo in Practice*, p.486, Chapman and Hall/CRC, London.

Goldberg, D. (1989) *Genetic Algorithms in Search, Optimization, and Machine Learning*, p.412, Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA.

Haas-Laursen, D.E., Hartley, D.E. and Prinn, R.G. (1996) 'Optimizing an inverse method to deduce time-varying emissions of trace gases', *Journal of Geophysical Research-Atmospheres*, Vol. 101, No. D1, pp.22823–22831.

Hansen, P.C. and O'Leary, D.P. (1993) 'The use of the L-curve in the regularization of discrete ill-posed problems', *SIAM Journal on Scientific Computing*, Vol. 14, No. 6, pp.1487–1503.

Hartley, D. and Prinn, R. (1993) 'Feasibility of determining surface emissions of trace gases using an inverse method in a three dimensional chemical transport model', *Journal of Geophysical Research-Atmospheres*, Vol. 98, No. D3, pp.5183–5197.

Haupt, S.E., Lout, A.B., Long, K.J. and Young, G.S. (2009) 'Assimilating concentration observations for transport and dispersion modelling in a meandering wind field', *Atmospheric Environment*, Vol. 43, No. 6, pp.1329–1338.

Haupt, S.E., Young, G.S. and Allen, C.T. (2006) 'Validation of a receptor/dispersion model coupled with a genetic algorithm using synthetic data', *Journal of Applied Meteorology and Climatology*, Vol. 45, No. 3, pp.476–490.

Hein R., Crutzen, P.J. and Heimann, M. (1997) 'An inverse modelling approach to investigate the global atmospheric methane cycle', *Global Biogeochemical Cycles*, Vol. 11, No. 1, pp.43–76.

Hourdin, F. and Issartel, J.P. (2000) 'Sub-surface nuclear tests monitoring through the CTBT xenon network', *Geophysical Research Letters*, Vol. 27, No. 15, pp.2245–2248.

Hourdin, F. and Talagrand, O. (2006) 'Eulerian Backtracking of atmospheric tracers. I: adjoint derivation and parameterization of subgrid-scale transport', *Quarterly Journal of Royal Meteorological Society*, Vol. 132, No. 615, pp.567–583.

Hourdin, F., Talagrand, O. and Idelkadi, A. (2006) 'Eulerian backtracking of atmospheric tracers. II: numerical aspects', *Quarterly Journal of Royal Meteorological Society*, Vol. 132, No. 61, pp.585–603.

Houweling, S., Breon, F-M., Aben, I., Rodenbeck, C., Gloor, M., Heimann, M. and Ciais, P. (2004) 'Inverse modeling of CO2 sources and sinks using satellite data: a synthetic inter-comparison of measurement techniques and their performance as a function of space and time', *Atmospheric Chemistry and Physics*, Vol. 4, No. 2, pp.523–538.

Houweling, S., Kaminski, T., Dentener, F., Lelieveld, J. and Heimann, M. (1999) 'Inverse modelling of methane sources and sinks using the adjoint of a global transport model', *Journal of Geophysical Research-Atmospheres*, Vol. 104, No. D21, pp.26137–26160.

Issartel, J.P. (2003) 'Rebuilding sources of linear tracers after atmospheric concentration measurements', *Atmospheric Chemistry and Physics*, Vol. 3, No. 6, pp.2111–2125.

Issartel, J.P. (2005) 'Emergence of a tracer source from air concentration measurements: a new strategy for linear assimilation', *Atmospheric Chemistry and Physics*, Vol. 5, No. 1, pp.249–273.

Issartel, J.P. and Baverel, J. (2003) 'Inverse transport for the verification of the "comprehensive nuclear test ban treaty" in place of "CTBT"', *Atmospheric Chemistry and Physics*, Vol. 3, No. 3, pp.475–486.

Issartel, J.P. and Gamel, T. (2010) 'Designing a monitoring system for a semi-urban set of buildings', *Proceedings of 13th Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes*, June 1–4, Paris, France, pp.963–967.

Issartel, J.P., Sharan, M. and Modani, M. (2007) 'An inversion technique to retrieve the source of a tracer with an application to synthetic satellite measurements', *Proceedings of Royal Society A*, Vol. 463, No. 2087, pp.2863–2886.

Issartel, J.P., Sharan, M. and Singh, S.K. (2012) 'Identification of a point release by optimally weighted least squares', *Pure and Applied Geophysics*, Vol. 169, No. 3, pp.467–482.

Jorquera, H. and Castro, J. (2010) 'Analysis of urban pollution episodes by inverse modelling', *Atmospheric Environment*, Vol. 44, No. 1, pp.42–54.

Kalman, R. (1960) 'A new approach to linear filtering and linear prediction problems', *Transactions of ASME*, *Series D*, *Journal of Basic Engineering*, Vol. 82, No. 1, pp.35–45.

Kalman, R. and Bucy, R. (1961) 'New results in linear filtering and prediction theory', *Transactions of ASME*, *Series D*, *Journal of Basic Engineering*, Vol. 83, No. 1, pp.95–108.

Kathirgamanathan, P., McKibbin, R. and McLachlan, R.I. (2004) 'Source release-rate estimation of atmospheric pollution from a non-steady point source at a known location', *Environmental Modelling and Assessment*, Vol. 9, No. 1, pp.33–42.

Keats, A., Yee, E. and Lien, F.S. (2007a) 'Bayesian inference for source determination with applications to a complex urban environment', *Atmospheric Environment*, Vol. 41, No. 3, pp.465–479.

Keats, A., Yee, E. and Lien, F.S. (2007b) 'Efficiently characterizing the origin and decay rate of a non-conservative scalar using probability theory', *Ecological Modelling*, Vol. 205, Nos. 3–4, pp.437–452.

Keller, M.H. and Herrnberger, V.R.D. (1997) 'Testing a source rate estimation method by application to a dispersion experiment', *Atmospheric Environment*, Vol. 31, No. 20, pp.3283–3290.

Khemka, A., Bouman, C.A. and Bell M.R. (2006) 'Inverse problems in atmospheric dispersion with randomly scattered sensors', *Digital Signal Processing*, Vol. 16, No. 5, pp.638–651.

Khlaifi, A., Ionescu, A. and Candau, Y. (2009) 'Pollution source identification using a coupled diffusion model with a genetic algorithm', *Mathematics and Computers in Simulation*, Vol. 79, No. 12, pp.3500–3510.

Koohkan, M.R. and Bocquet, M. (2012) 'Accounting for representativeness errors in the inversion of atmospheric constituent emissions: application to the retrieval of regional carbon monoxide fluxes', *Tellus B*, Vol. 64, pp.19047.

Koohkan, M.R., Bocquet, M., Wu, L. and Krysta, M. (2012) 'Potential of the international monitoring system radionuclide network for inverse modelling', *Atmospheric Environment*, Vol. 54, pp.557–567.

Korsakissok, I., Mallet, V. and Sportisse, B. (2010) 'Assessment of the detection abilities of monitoring networks for passive tracers at local and regional scales', *Atmospheric Environment*, Vol. 44, No. 3, pp.408–420.

Krysta, M. and Bocquet, M. (2007) 'Source reconstruction of an accidental radionuclide release at European scale', *Quarterly Journal of Royal Meteorological Society*, Vol. 133, No. 62, pp.529–544.

Krysta, M., Bocquet, M. and Brandt, J. (2008) 'Probing ETEX-II data set with inverse modelling', *Atmospheric Chemistry and Physics*, Vol. 8, No. 14, pp.3963–3971.

Krysta, M., Bocquet, M., Sportisse, B. and Isnard, O. (2006) 'Data assimilation for short-range dispersion of radionuclides: an application to wind tunnel data', *Atmospheric Environment*, Vol. 40, No. 38, pp.7267–7279.

Le Dimet, F.X. and Talagrand, O. (1986) 'Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects', *Tellus*, Vol. 38A, No. 2, pp.97–110.

Lewis, J.M., Lakshmivarahan, S. and Dhall S.K. (2006) '*Dynamic Data Assimilation: A Least Square Approach*, p.680, Cambridge University Press, New York.

Long, K.J., Haupt, S.E. and Young, G.S. (2010) 'Assessing sensitivity of source term estimation', *Atmospheric Environment*, Vol. 44, No. 12, pp.1558–1567.

Lushi, E. and Stockie, J.M. (2010) An inverse Gaussian plume approach for estimating atmospheric pollutant emissions from multiple point sources', *Atmospheric Environment*, Vol. 44, No. 8, pp.1097–1107.

Marchuk, G.I. (1995) *Adjoint Equations and Analysis of Complex Systems*, p.484, Kluwer Academic Publishers, Dordrecht.

Matthes, J., Gröll, L. and Keller, H.B. (2005) 'Source localization by spatially distributed electronic noses for advection and diffusion', *IEEE Transactions on Signal Processing*, Vol. 53, No. 5, pp.1711–1719.

Meirink, J.F., Bergamaschi, P., Frankenberg, C., d'Amelio, M.T.S., Dlugokencky, E.J., Gatti, L.V., Houweling, S., Miller, J.B., Rockmann, T., Villani, M.G. and Krol, M. (2008) 'Four-dimensional variational data assimilation for inverse modelling of atmospheric methane emissions: analysis of SCIAMACHY observations', *Journal of Geophysical Research-Atmospheres*, Vol. 113, No. D17, Art. No. D17301.

Mulholland, M. and Seinfeld, J.H. (1995) 'Inverse air pollution modelling of urban-scale carbon monoxide emissions', *Atmospheric Environment*, Vol. 29, No. 4, pp.497–516.

Muller, J.F. and Stavrakou, T. (2005) 'Inversion of CO and NO_x emissions using the adjoint of the IMAGES model', *Atmospheric Chemistry and Physics*, Vol. 5, No. 5, pp.1157–1186.

Penenko, V., Baklanov, A. and Tsvetova, E. (2002) 'Methods of sensitivity theory and inverse modelling for estimation of source term', *Future Generation Computer Systems*, Vol. 18, No. 5, pp.661–671.

Pudykiewicz, J.A. (1998) 'Application of adjoint tracer transport equations for evaluating source parameters', *Atmospheric Environment*, Vol. 32, No. 17, pp.3039–3050.

Quelo, D., Mallet, V. and Sportisse, B. (2005) 'Inverse modelling of NO_x emissions at regional scale over Northern France. Preliminary investigation of the second order sensitivity', *Journal of Geophysical Research*, Vol. 110, No. D24, Art. No. D24310.

Rao, K.S. (2007) 'Source estimation methods for atmospheric dispersion', *Atmospheric Environment*, Vol. 41, No. 33, pp.6964–6973.

Redwood, M. (2011) 'Source term estimation and event reconstruction: a survey' [online] https://admlc.files.wordpress.com/2014/05/admlc-r6.pdf.

Reen, B.P. and Stauffer, D.R. (2010) 'Data assimilation strategies in the planetary boundary layer', *Boundary-Layer Meteorology*, Vol. 137, No. 2, pp.237–269.

Roberti, D.R., Anfossi, D., Campos Velho, H.F. and Degrazia, G.A. (2005) 'Estimation of emission rate from pollutant sources', *Proceedings of the 5th International Conference on Inverse Problems in Engineering: Theory and Practice*, Cambridge, UK.

Robertson, L. (2004) *Extended Back-trajectories by Means of Adjoint Equations*, Swedish Meteorological and Hydrological Institute, RMK 105, Sweden.

Robertson, L. and Langner, J. (1998) 'Source function estimate by means of variational data assimilation applied to the ETEX-I tracer experiment', *Atmospheric Environment*, Vol. 32, No. 24, pp.4219–4225.

Robertson, L. and Persson, C. (1993) 'Attempt to apply four dimensional data assimilation of radiological data using the adjoint technique', *Radiation Protection Dosimetry*, Vol. 50, Nos. 2–4, pp.333–337.

Rodgers, C. (2000) 'Inverse methods for atmospheric sounding, theory and practice', Series on *Atmospheric*, *Oceanic and Planetary Physics*, p.238, World Scientific Publishing Co., River Edge, NJ.

Saunier, O., Bocquet, M., Mathieu, A. and Isnard, O. (2009) 'Model reduction via principal component truncation for the optimal design of atmospheric monitoring networks', *Atmospheric Environment*, Vol. 43, No. 32, pp.4940–4950.

Seibert, P. (2000) 'Uncertainities in atmospheric dispersion modelling and source determination', *Proceedings Informal Workshop on Meteorological Modelling in Support of CTBT Verification*, Vienna, Austria.

Seibert, P. (2001) 'Inverse modelling with a Lagrangian particle dispersion model: application to point releases over limited time intervals', in S.E. Gryning and F.A. Schiermeier (Eds.): *Air Pollution and its Application XIV*, pp.381–389, NATO, Kluwer Academic/Plenum Publishers, New York, NY.

Seibert, P. and Stohl, A. (2000) 'Inverse modelling of the ETEX-1 release with a Lagrangian particle model', *Proceedings of the 3rd GLOREAM workshop*, September, Ischia, Italy.

Senocak, I., Hengartner, N.W., Short, M.B. and Daniel, W.B. (2008) 'Stochastic event reconstruction of atmospheric contaminant dispersion using Bayesian inference', *Atmospheric Environment*, Vol. 42, No. 33, pp.7718–7727.

Sharan, M., Issartel, J.P. and Singh, S.K. (2012a) 'A point source reconstruction from concentration measurements in low wind stable conditions', *Quarterly Journal of Royal Meteorological Society*, Vol. 138, No. 668, pp.1884–1894.

Sharan, M., Singh, S.K. and Issartel, J.P. (2012b) 'Least square data assimilation for identification of the point source emissions', *Pure and Applied Geophysics*, Vol. 169, No. 3, pp.483–497.

Sharan, M., Issartel, J.P., Singh, S.K. and Kumar, P. (2009) 'An inversion technique for the retrieval of single-point emissions from atmospheric concentration measurements', *Proceedings of Royal Society A*, Vol. 465, pp.2069–2088.

Sharan, M., Singh, M.P., Yadav, A.K., Aggarwal, P. and Nigam, S. (1996) 'A mathematical model for the dispersion of pollutants in low wind conditions', *Atmospheric Environment*, Vol. 30, No. 8, pp.1209–1220.

Singh, M.P., Agarwal, P., Nigam, S. and Gulati, A. (1991) *Tracer Experiments – A Report*, Technical report, CAS, IIT Delhi.

Singh, S.K. and Rani, R. (2014) 'A least-squares inversion technique for identification of a point release: application to fusion field trials 2007', *Atmospheric Environment*, Vol. 92, pp.104–117.

Singh, S.K., Sharan, M. and Issartel, J.P. (2013) 'Inverse modelling for identification of multiple-point releases from atmospheric concentration measurements', *Boundary-Layer Meteorology*, Vol. 146, No. 2, pp.277–295.

Singh, S.K., Turbelin, G., Issartel, J.P., Feiz, A.A., Ngae, P., Barbosa, E. and Kumar P. (2014) 'Analyzing localization features of a weighted least-squares technique in a point source reconstruction', *Proceedings of the 16th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes (HARMO16)*, 8–11 September, Varna, Bulgaria, pp.518–523.

Singh, S.K., Turbelin, G., Issartel, J-P., Kumar, P. and Feiz, A.A. (2015) 'Reconstruction of an atmospheric tracer source in fusion field trials: analysing resolution features', *J. Geophys. Res. Atmos.*, Vol. 120, No. 12, pp.6192–6206, doi: 10.1002/2015JD023099.

Sohn, M.D., Reynolds, P., Singh, N. and Gadgil, A. (2002) 'Rapidly locating and characterizing pollutant releases in buildings', *Journal of Air and Waste Management Association*, Vol. 52, No. 12, pp.1422–1432.

Stohl, A. (1998) 'Computation, accuracy and application of trajectories – a review and bibliography', *Atmospheric Environment*, Vol. 32, No. 6, pp.947–966.

Storch, R.B., Pimentel, L.C.G. and Orlande, H.R.B. (2005) 'Identification of micro-meteorological parameters for the characterization of atmospheric boundary layers', *Proceeding of the 5th International Conference on Inverse Problems in Engineering: Theory and Practice*, Cambridge, UK.

Storch, R.B., Pimentel, L.C.G. and Orlande, H.R.B. (2007) 'Identification of atmospheric boundary layer parameters by inverse problem', *Atmospheric Environment*, Vol. 41, No. 7, pp.1417–1425.

Storwald, D.P. (2007) *Detailed Test Plan for the Fusing Sensor Information from Observing Networks (Fusion) field trial (FFT-07)*, Meteorology Division, West Desert Test Center, US Army Dugway Proving Ground.

Tarantola, A. (2005) *Inverse Problem Theory and Methods for Model Parameter Estimation*, p.342, Society for Industrial and Applied Mathematics, Philadelphia.

Thomson, L.C., Hirst, B., Gibson, G., Gillespie, S., Jonathan, P., Skeldon, K.D. and Padgett, M.J. (2007) 'An improved algorithm for locating a gas source using inverse methods', *Atmospheric Environment*, Vol. 41, No. 6, pp.1128–1134.

Turbelin, G., Singh, S.K. and Issartel, J-P. (2014) 'Reconstructing source terms from atmospheric concentration measurements: optimality analysis of an inversion technique', *J. Adv. Model. Earth Syst.*, Vol. 6, No. 4, pp.1244–1255, doi: 10.1002/2014MS000385.

Villani, M.G., Bergamaschi, P., Krol, M., Meirink, J.F. and Dentener, F. (2010) 'Inverse modelling of European CH_4 emissions: sensitivity to the observational network', *Atmospheric Chemistry and Physics*, Vol. 10, No. 3, pp.1249–1267.

Wade, D. and Senocak, I. (2013) 'Stochastic reconstruction of multiple source atmospheric contaminant dispersion events', *Atmospheric Environment*, Vol. 74, pp.45–51.

Winiarek, V., Bocquet, M., Duhanyan, N., Roustan, Y., Saunier, O., Mathieu, A. (2014) 'Estimation of the caesium-137 source term from the Fukushima Daiichi nuclear power plant using a consistent joint assimilation of air concentration and deposition observations', *Atmospheric Environment*, Vol. 82, pp.268–279.

Winiarek, V., Bocquet, M., Saunier, O. and Mathieu, A. (2012) 'Estimation of errors in the inverse modelling of accidental release of atmospheric pollutant: application to the reconstruction of the cesium-137 and iodine-131 source terms from the Fukushima Daiichi power plant', *Journal of Geophysical Research*, Vol. 117, No. D5, Art. No. D05122.

Wu, L. and Bocquet, M. (2011) 'Optimal redistribution of the background ozone monitoring stations over France', *Atmospheric Environment*, Vol. 45, No. 3, pp.772–783.

Wu, L., Bocquet, M. and Chevallier, M. (2010) 'Optimal reduction of the ozone monitoring network over France', *Atmospheric Environment*, Vol. 44, No. 25, pp.3071–3083.

Yee, E. (2007) 'Bayesian probabilistic approach for inverse source determination from limited and noisy chemical or biological sensor concentration measurements', in Augustus W. Fountain III (Ed.): *Proceeding of SPIE* 6554, *Chemical and Biological Sensing VIII*, 12 pp.

Yee, E. (2008) 'Theory for reconstruction of an unknown number of contaminant sources using probabilistic inference', *Boundary-Layer Meteorology*, Vol. 127, No. 3, pp.359–394.

Yee, E. (2012) 'Probability theory as logic: data assimilation for multiple source reconstruction', *Pure and Applied Geophysics*, Vol. 169, No. 3, pp.499–517.

Yee, E., Lien, F.S., Keats, A. and D'Amours, R. (2008) 'Bayesian inversion of concentration data: Source reconstruction in the adjoint representation of atmospheric diffusion', *Journal of Wind Engineering and Industrial Aerodynamics*, Vol. 96, Nos. 10–11, pp.1805–1816.

Zhao, T. and Nehorai, A. (2007) 'Distributed sequential Bayesian estimation of a diffusive source in wireless sensor networks', *IEEE Transactions on Signal Processing*, Vol. 55, No. 4, pp.1511–1524.

Zheng, X.P. and Chen, Z.Q. (2010) 'Back-calculation of the strength and location of hazardous materials releases using the pattern search method', *Journal of Hazardous Materials*, Vol. 183, Nos. 1–3, pp.474–481.