

# Computation of Optimal Weights for Solving the Atmospheric Source Term Estimation Problem

GRÉGORY TURBELIN, SARVESH SINGH, AND JEAN PIERRE ISSARTEL

*LMEE, Univ. Evry, Université Paris-Saclay, Paris, France*

XAVIER BUSCH

*DGA, Maitrise NRBC, Vert-le-Petit, France*

PRAMOD KUMAR

*LMEE, Univ. Evry, Université Paris-Saclay, Paris, France*

(Manuscript received 24 August 2018, in final form 25 March 2019)

## ABSTRACT

In case of a release of a hazardous material (e.g., a chemical or a biological agent) in the atmosphere, estimation of the source from concentration observations (provided by a network of sensors) is a challenging inverse problem known as the atmospheric source term estimation (STE) problem. This study emphasizes a method, known in the literature as the renormalization inversion technique, for addressing this problem. This method provides a solution that has been interpreted as a weighted minimal norm solution and can be computed in terms of a generalized inverse of the sensitivity matrix of the sensors. This inverse is constructed by using an appropriate diagonal weight matrix whose components fulfill the so-called renormalizing conditions. The main contribution of this paper is that it proposes a new compact algorithm (it requires less than 15 lines of MATLAB code) to obtain, in a fast and efficient way, those optimal weights. To show that the algorithm, based on the properties of the resolution matrix, matches the requirements of emergency situations, analysis of the computational complexity and memory requirements is included. Some numerical experiments are also reported to show the efficiency of the algorithm.

## 1. Introduction

In case of an intentional or accidental release of chemical, biological, or radiological (CBR) material into the atmosphere, first responders need accurate methods to localize and quantify the source of the release. For that purpose, the renormalization inversion technique (or renormalized assimilation technique) has been developed. This method, extensively described in earlier papers, for example, Issartel (2005), Issartel et al. (2007, 2012), Kumar et al. (2015, 2017), and Singh et al. (2015, 2017a,b), aims at solving the source term estimation

(STE) problem of reconstructing the sources of atmospheric contaminations from limited sets of concentration measurements. This inverse problem is underdetermined, and the renormalization approach leads to a unique solution without using any prior statistical information regarding the measurements errors and the sought source.

From a theoretical point of view, the renormalized source estimate may be interpreted as a weighted minimal norm solution and computed in terms of a generalized inverse (Turbelin et al. 2014). This inverse may be derived after having computed optimal weights satisfying the so-called renormalizing conditions proposed by Issartel et al. (2007). From a computational point of view, the algorithm used to estimate those renormalizing weights has to be efficient. Indeed, in an operational environment, the time for source estimation is short and the available computational power is often limited.

The contributions of this work are as follows. We present a new compact and efficient algorithm that has not been

---

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JTECH-D-18-0145.s1>.

---

Corresponding author: Grégory Turbelin, [gregory.turbelin@univ-evry.fr](mailto:gregory.turbelin@univ-evry.fr)

DOI: 10.1175/JTECH-D-18-0145.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) ([www.ametsoc.org/PUBSReuseLicenses](http://www.ametsoc.org/PUBSReuseLicenses)).

previously available. This algorithm, based on the properties of the resolution matrix, may be used to compute, in a fast and accurate way, the renormalizing weights. In addition, to show that the algorithm matches the requirements of emergency situations, its complexity has been analyzed. Then a full vectorized version of the algorithm, implemented in MATLAB code, is proposed.

The paper is organized as follows. Section 2 presents the method and its theoretical background in the general framework of nonparametric STE inverse problems. In section 3, we present the new algorithm, and its complexity is analyzed. To illustrate the efficiency of the algorithm, some numerical experiments are reported in section 4.

Note that, throughout the manuscript, vectors and matrices are denoted by boldface symbols and scalars by italic symbols. The superscript “T” denotes the transposition.

## 2. Theoretical background

### a. Problem statement

During an atmospheric dispersion event in a  $p$ -dimensional space and time domain  $\Omega$ , the  $m$  concentration measurements performed by a network of detectors can be assembled into a vector  $\boldsymbol{\mu}$ . A vector space of dimension  $m$ , represented by  $\mathbf{U}$ , and called the measurement space, encompasses all possible measurement vectors that could be observed during the event. For its part, the source of the dispersion may be described by a scalar source function  $s(\mathbf{x})$ , which represents the intensity of the source at coordinate  $\mathbf{x} \in \mathbb{R}^p$ . Note that, usually,  $\Omega$  is the product of a spatial domain  $D$  and of the time interval  $[0, T]$  of the event. Thus, one has  $p = 4$  if  $D$  is three-dimensional. However, we can also consider the simplified case of a continuous release. In this situation time coordinate is ignored and  $p = 3$  or  $2$ , depending on the dimension of  $D$ . When  $\Omega$  is discretized into a grid of  $n$  cells (with  $n \gg m$ ), the source is represented by a vector  $\mathbf{s}$ , with  $s_j = s(\mathbf{x}_j)dx_j$ ,  $j = 1, \dots, n$ , where  $dx_j$  is the volume of the cells. Consequently, each conceivable source lies in a vector space of dimension  $n$ , represented by  $\mathbf{V}$ , called the source space. The characterization of  $\mathbf{s}$  from  $\boldsymbol{\mu}$  is given by an undetermined (since  $n > m$ ) system of equations:

$$\boldsymbol{\mu} = \mathbf{A}\mathbf{s} + \boldsymbol{\varepsilon}, \quad (1)$$

where  $\boldsymbol{\varepsilon} \in \mathbb{R}^m$  is the random errors vector and  $\mathbf{A} \in \mathbb{R}^{m \times n}$  is the sensitivity matrix of the network. This matrix can be expressed as  $\mathbf{A} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_n]$ , where each vector  $\mathbf{a}_j$ , formed by  $m$  rows, contains the measurements generated by a potential unitary release at the point of coordinate  $\mathbf{x}_j$ . During emergency situations, the components of this matrix can be computed, very quickly, by using

analytical dispersion models in a backward mode (Issartel and Baverel 2003). In (1), the matrix  $\mathbf{A}$  defines a linear mapping from  $\mathbf{V}$  into  $\mathbf{U}$ . The source term estimation inverse problem consists in finding an appropriate inverse mapping from  $\mathbf{U}$  back into  $\mathbf{V}$ . To this end, additional assumptions/constraints are required.

### b. The renormalization inversion technique

Many of the existing techniques used to derive a unique and stable solution to the inverse STE problem are nonparametric in nature: they seek to retrieve a source field, that is, the  $n$  components of  $\mathbf{s}$ , by using the measurements and a priori assumptions/constraints about the source. A common assumption consists in considering that the underdetermined system (1) is consistent, that is, in considering that it has an infinite number of solutions all of them satisfying  $\boldsymbol{\mu} = \mathbf{A}\mathbf{s}$  exactly. It is well known that the infinite set of such solutions can be expressed as  $\mathbf{s} = \mathbf{s}_// + \mathbf{s}_\perp$ , where  $\mathbf{s}_\perp$  is any vector orthogonal to the rows of  $\mathbf{A}$ , that is, verifying  $\mathbf{A}\mathbf{s}_\perp = \mathbf{0}$ , and where  $\mathbf{s}_//$  is the only part of  $\mathbf{s}$  contributing to the measurements. This vector is unique and corresponds with the minimum norm solution of the system. According to Ben-Israel and Greville (2003), it is the solution for which the Euclidian norm of  $\mathbf{s}$ , denoted as  $\|\mathbf{s}\|$ , is minimum. As explained by Turbelin et al. (2014), it is defined as

$$\mathbf{s}_// = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\boldsymbol{\mu} = \mathbf{A}^+\boldsymbol{\mu}. \quad (2)$$

In (2), the inverse operator  $\mathbf{A}^+ \in \mathbb{R}^{n \times m}$  denotes the Moore–Penrose pseudo inverse of  $\mathbf{A}$ . It is also called a (1, 2, 3, 4)-generalized inverse, because it verifies the four Penrose equations. First:  $\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}$ ; second:  $\mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+$ ; third:  $(\mathbf{A}\mathbf{A}^+)^T = \mathbf{A}\mathbf{A}^+$ ; and fourth:  $(\mathbf{A}^+\mathbf{A})^T = \mathbf{A}^+\mathbf{A}$ . The solution, as given in (2), has a number of computational advantages, but it has the tendency to put the large components of  $\mathbf{s}$  close to the sampling locations, that is, at locations to which the detectors are most sensitive. To avoid these artifacts, Issartel et al. (2007) proposed a solution minimizing a weighted norm  $\|\mathbf{s}\|_{\mathbf{W}} = \sqrt{\mathbf{s}^T\mathbf{W}\mathbf{s}} = \|\mathbf{W}^{-1/2}\mathbf{s}\|$ , where  $\mathbf{W} \in \mathbb{R}^{n \times n}$  is a diagonal weighting (or renormalizing) matrix. Following entropic criteria, Issartel et al. (2007) have shown that the optimal components of  $\mathbf{W}$  should satisfy the following properties (renormalizing conditions):

$$\begin{aligned} \text{(i) } w_j &> 0; \quad \text{(ii) } \sum_{j=1}^n w_j = m; \quad \text{(iii) } w_j^2 = \mathbf{a}_j^T \mathbf{H}_{\mathbf{W}}^{-1} \mathbf{a}_j \\ &\text{for } j = 1 \dots n, \end{aligned} \quad (3)$$

where  $\mathbf{H}_{\mathbf{W}} = \mathbf{A}\mathbf{W}^{-1}\mathbf{A}^T \in \mathbb{R}^{m \times m}$  is the Gram matrix of the row of  $\mathbf{A}\mathbf{W}^{-1/2}$ . Turbelin et al. (2014) have shown that the corresponding minimum weighted norm solution is obtained as

$$\mathbf{s}_{//w} = \mathbf{W}^{-1} \mathbf{A}^T (\mathbf{A} \mathbf{W}^{-1} \mathbf{A}^T)^{-1} \boldsymbol{\mu} = \mathbf{A}_W^+ \boldsymbol{\mu}. \quad (4)$$

The matrix  $\mathbf{A}_W^+ \in \mathbb{R}^{n \times m}$ , defined as in (4), verifies only the three first Penrose conditions; therefore, it is called a (1, 2, 3)-generalized inverse of  $\mathbf{A}$ . Note that, physically, the optimal renormalizing weights verifying (3) have been interpreted as visibility parameters describing how well the network of sensors monitors the various locations  $\mathbf{x}_j$ : the spatial distribution of the weights (e.g., the maps plotted in section 4) highlights the regions accurately monitored (i.e., the ones characterized by a relatively high value of  $w_j$ ); see Issartel et al. (2007).

c. The model resolution matrix

Substitution of the measurement vector,  $\boldsymbol{\mu} = \mathbf{A}\mathbf{s}$ , into (4) yields the following equation:

$$\mathbf{s}_{//w} = \mathbf{A}_W^+ \mathbf{A} \mathbf{s} = \mathbf{R} \mathbf{s}, \quad (5)$$

where  $\mathbf{R}$  is the “model resolution matrix” (Menke 2012) describing the relationship between the estimated solution

and the original source. This matrix is often used to analyze the quality of inverse solutions: from (5), it can be seen that a unique and perfectly accurate estimate could be obtained if  $\mathbf{R}$  was equal to the identity matrix. However, this situation may not occur in an underdetermined inverse problem ( $\mathbf{R}$  cannot have a rank greater than  $m$ ). For single point source reconstruction, the most we can expect is that the columns of  $\mathbf{R}$ , denoted as  $\mathbf{r}_j \in \mathbb{R}^n$ , with  $j = 1, \dots, n$ , reach their maximal values at the diagonal elements. We now explain the interest of this property and show that, indeed, it holds for the renormalized source solution. If the source is a single point source, one has  $\mathbf{s} = q_s \delta(\mathbf{x} - \mathbf{x}_s)$  (where  $\delta$  is the Dirac delta function) and (5) gives  $\mathbf{s}_{//w} = q_s \mathbf{r}_s$ . Therefore, if the maximum of  $\mathbf{r}_s$  is reached at the main diagonal, the source position corresponds with the position for which  $s_{//w}(\mathbf{x})$  is maximum. Grave de Peralta et al. (2009) called this property “perfect localization of single point source.” By using (4), (5), and the third condition of (3), the resolution matrix for the renormalized source solution is

$$\mathbf{R} = \mathbf{A}_W^+ \mathbf{A} = \begin{bmatrix} w_1 & \dots & w_1^{-1} \mathbf{a}_1^T \mathbf{H}_W^{-1} \mathbf{a}_k & \dots & w_1^{-1} \mathbf{a}_1^T \mathbf{H}_W^{-1} \mathbf{a}_n \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ w_k^{-1} \mathbf{a}_k^T \mathbf{H}_W^{-1} \mathbf{a}_1 & \dots & w_k & \dots & w_k^{-1} \mathbf{a}_k^T \mathbf{H}_W^{-1} \mathbf{a}_n \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ w_n^{-1} \mathbf{a}_n^T \mathbf{H}_W^{-1} \mathbf{a}_1 & \dots & w_n^{-1} \mathbf{a}_n^T \mathbf{H}_W^{-1} \mathbf{a}_k & \dots & w_n \end{bmatrix}. \quad (6)$$

With the Cauchy–Schwarz inequality and the third condition of (3) we can show that  $w_i^{-1} \mathbf{a}_i^T \mathbf{H}_W^{-1} \mathbf{a}_j \leq w_i^{-1} \sqrt{\mathbf{a}_i^T \mathbf{H}_W^{-1} \mathbf{a}_i} \sqrt{\mathbf{a}_j^T \mathbf{H}_W^{-1} \mathbf{a}_j} \leq w_j \forall i, j = 1, \dots, n$ , and

thus, each column of  $\mathbf{R}$  (i.e., each vector  $\mathbf{r}_j$ ) will have a maximum at the diagonal, with a value equal to  $w_j$ . Consequently, without any a priori assumption, in case of observations generated from a continuous point source an estimate  $\hat{\mathbf{x}}_s$  of the location of the source can be obtained by searching the position for which  $s_{//w}(\mathbf{x})$  is maximum. Then the source strength is estimated by dividing this maximum by the value of the weight at this point:

$$\hat{\mathbf{x}}_s = \arg \max_{\mathbf{x} \in \Omega} [s_{//w}(\mathbf{x})] \quad \text{and} \quad \hat{q}_s = s_{//w}(\hat{\mathbf{x}}_s) / w(\hat{\mathbf{x}}_s). \quad (7)$$

When the source is a single point source, without any a priori information, the renormalization technique leads to its perfect characterization. Indeed, one has this property provided that (i) the random noise in (1) is not unacceptably too large and (ii) the number of measurements is larger than the number of unknown source parameters (e.g., for a transient release from a point source, one needs at least six measurements to estimate the three spatial localization parameters, the source intensity, the release starting time, and its duration).

The resolution matrix, as given by (6), has other interesting properties:

- (i)  $\mathbf{R}$  is idempotent and its diagonal elements are equal to the weights.
  - (ii)  $\mathbf{R} \mathbf{W}^{-1}$  is symmetric and its diagonal elements are equal to 1.
- (8)

As a consequence, the optimal weights verifying (3) may also be defined as the diagonal elements of the resolution

matrix  $\mathbf{R}$ , when the diagonal elements of the symmetric matrix  $\mathbf{R} \mathbf{W}^{-1}$  are equal to one. Those properties have

been utilized to define a fast and efficient algorithm for computing the renormalizing weights.

### 3. Computing optimal weights

#### a. The iterative method

To compute the weights at each point  $\mathbf{x}_j$ , Issartel et al. (2007) proposed the following iterative scheme that converges uniformly to the weights satisfying the renormalization conditions:

$$w_j^{\alpha+1}(\mathbf{x}_j) = w_j^{\alpha+1} = w_j^\alpha \sqrt{(w_j^\alpha)^{-1} \mathbf{a}_j^T \mathbf{H}_{\mathbf{W}^\alpha}^{-1} \mathbf{a}_j (w_j^\alpha)^{-1}}$$

with  $w_j^0 = 1$ ,

(9)

where the subscript  $\alpha$  denotes the iteration step. This scheme has been embedded in an algorithm originally implemented in Fortran 77 (J. P. Issartel, unpublished manuscript). This algorithm was effective but very complex (mainly composed of nested loops that consumed computational power) and far from being optimal with respect to speed and memory usage. Here, a very simple (compact) and yet efficient one is proposed. Moreover, we are convinced that this new algorithm better suits the modern languages like Python, R, MATLAB, or even C++ that are nowadays the most popular languages for scientific computations.

#### b. The proposed algorithm

Noticing that the term under the square root in (9) corresponds with the  $j$ th diagonal term of  $\mathbf{R}\mathbf{W}^{-1}$ , an algorithm based on the properties enunciated in section 2c, (8), has been developed. It is graphically represented by the flowchart given in Fig. 1 and requires less than 15 lines of MATLAB code. The algorithm consists of two parts. The first one is the computation of  $\mathbf{H}_{\mathbf{W}^{-1}}$ . It involves the transposition of  $\mathbf{A}$  [computational complexity  $O(mn)$ ], the computation of  $\mathbf{W}^{-1}\mathbf{A}^T$  [ $O(nm)$ ], the matrix product  $\mathbf{H}_{\mathbf{W}} = \mathbf{A}\mathbf{W}^{-1}\mathbf{A}^T$  [ $O(m^2n)$ ], and its inversion [ $O(m^3)$ ]. The second one is the computation of the diagonal of  $\mathbf{R}\mathbf{W}^{-1} = \mathbf{W}^{-1}\mathbf{A}^T\mathbf{H}_{\mathbf{W}}^{-1}\mathbf{A}\mathbf{W}^{-1} = \mathbf{A}_{\mathbf{W}}^+\mathbf{A}\mathbf{W}^{-1}$ . Since the  $j$ th element in the diagonal is the scalar product of the  $j$ th row of  $\mathbf{A}_{\mathbf{W}}^+$  with the  $j$ th column of  $\mathbf{A}\mathbf{W}^{-1}$  (which is also the  $j$ th line of  $\mathbf{W}^{-1}\mathbf{A}^T$ ), this part involves the computation of  $\mathbf{A}_{\mathbf{W}}^+$  [ $O(m^2n)$ ] and the computation of the scalar products [ $O(mn)$ ]. The verification of the convergence criterion is based on the search of the diagonal term farthest from one [ $O(n)$ ; see below]. Consequently, since  $n \gg m$ , the overall computational complexity of the algorithm is  $O(m^2n)$ . This computational complexity has been confirmed by a

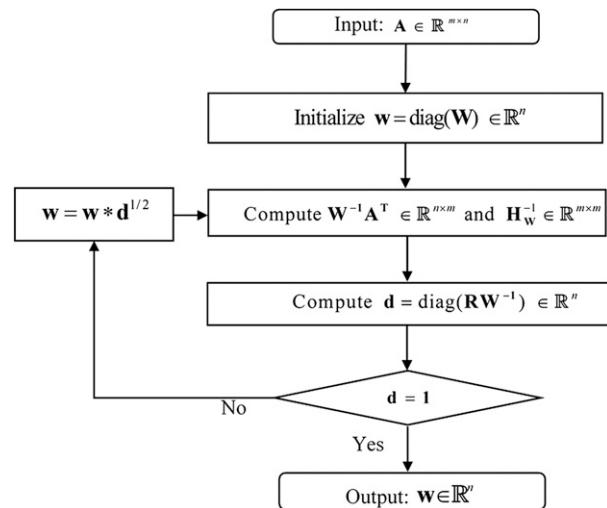


FIG. 1. Flowchart for the proposed algorithm.

linear regression between  $m^2n$  and computational time, see Fig. 2 (left). Regarding the storage requirement, since only the diagonal terms of the “huge matrix”  $\mathbf{R}\mathbf{W}^{-1}$  are computed, the matrices that demand for maximum storage space are the ones of dimensions  $m \times n$  or  $n \times m$ . Hence, the amount of storage requirement is  $O(mn)$ .

A full vectorized version of this algorithm has been implemented in MATLAB code (version 8.4) and is listed in the appendix. The MATLAB function corresponding to the computation of the optimal weights is named “RENORM.” By construction, the output of the function is a vector of dimension  $n$  that contains the weights to be used in (4). The performances of the algorithm have been increased by using high-level MATLAB operators such as “inv( )” and “bsxfun( )” (where “bsx” means “binary singleton expansion”). Note that to satisfy the second condition in (3), the weights are initialized with the value of  $m/n$  instead of the value of one suggested by Issartel et al. (2007). The stopping rule is based on having one (given an acceptance tolerance  $\varepsilon$ ) on the main diagonal of  $\mathbf{R}\mathbf{W}^{-1}$ ; that is,  $\max|d_j - 1| \leq \varepsilon$  for  $j = 1, \dots, n$ . Another stopping criteria based on the second renormalization property  $\sum_{j=1}^n w_j = m$  (or  $\sum_{j=1}^n w_j = \text{tr}(\mathbf{W}) = \text{tr}(\mathbf{R}) = \text{rank}(\mathbf{A}) = m$  can also be used. Note that convergence in sense of the first criteria guarantees convergence in sense of the second one. The numerical tests, using random matrices  $\mathbf{A} \in \mathbb{R}^{m \times n}$  and the first stopping criteria, show that the algorithm always converges, for example, see Fig. 2 (right), with the number of iterations depending on  $\varepsilon$ , see Table 1. With the second criteria, few more iterations are required.

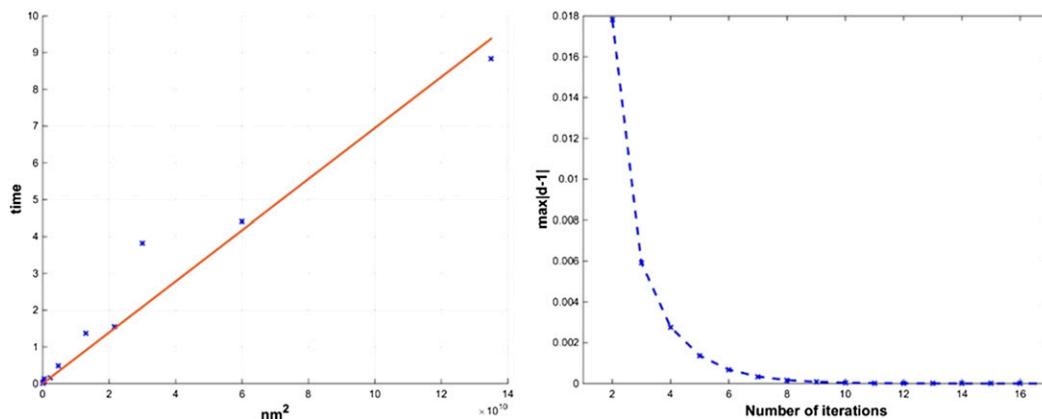


FIG. 2. (left) Linear regression between  $m^2n$  and computation time and (right) convergence for  $m = 100$  and  $n = 150\,000$  with  $\varepsilon = 10^{-7}$ .

### 4. Examples of use

Since the ultimate proof for an STE technique is its practical validation, the renormalization technique has been already tested, and found to be working efficiently, with the concentration measurements obtained from various atmospheric and tracer data experiments. More specifically, it has been recently evaluated, by Kumar et al. (2015, 2017) and Singh et al. (2015, 2017a), by using data from well-known field experiments. Here, the goal is not to perform another evaluation of the method but rather to report some calculations that show the capability of the inverse algorithm proposed in section 3b.

#### a. Application to the DYCE experiments

The first examples use data from a set of dispersion experiments conducted in the Environmental Flow Research Center (ENFLO) meteorological wind tunnel (University of Surrey, United Kingdom), in the frame of the “Dynamic deployment planning for monitoring of chemical leaks using an ad-hoc sensor network” (DYCE) project (Rudd et al. 2012). The experiments involved 11 trials of continuous releases of a tracer (a gas mixture of 1.5% propane in air) from a ground-level source. In all the trials, the propane concentration in the resulting plume was measured by a network of four fast flame ionization detectors (FFIDs) with a frequency response of 200 Hz, over a period of 15 min. The height of the source and samplers was  $0.1H$ , where  $H$  (approximately 1 m) is the boundary layer height. The  $x$  axis was along the wind direction, and the  $y$  axis was across the width of the tunnel. Eleven different network configurations (corresponding with 27 different positions for the FFIDs) had been tested with the same wind speed ( $U_H = 2.5\text{ m s}^{-1}$ ). The different configurations consisted of four sensors arranged either on lines across the plume or in grid-like arrays. The mean distance between the source and sensors  $X_m$  varied from

0.42 to 2.35 m. The maximum spacing, in the crosswind direction, between two sensors was 0.8 m. The spatial scaling between wind tunnel and field observations is  $1/500$ , and the wind speed of  $2.5\text{ m s}^{-1}$ , in the wind tunnel, corresponds to  $10\text{ m s}^{-1}$  in field scale. Therefore, by using the concept of dimensionless time  $UT/H$ , 1 min in the wind tunnel is approximately 2 h at full scale. More details about the experiments and dataset can be found in Rudd et al. (2012) and Turbelin et al. (2018).

For inversion computation, a rectangular domain  $D$  of dimension  $4\text{ m} \times 1\text{ m}$  (wind tunnel scale) has been discretized uniformly as a grid with  $n = 2000 \times 500$  square cells having dimension  $dS = 2\text{ mm} \times 2\text{ mm}$ . The downwind boundary of the computational domain corresponds to the position of the detector farthest from the source in the downwind direction. Two representative trials (5 and 7) have been chosen here for illustration and discussion. The first one involved sensors aligned crosswind ( $X_m = 2.35\text{ m}$ ) and the second a configuration based on a grid-like arrays ( $X_m = 1.825\text{ m}$ ). The actual point source was taken as origin with  $(x_s, y_s) = (0, 0)$ ; its emission rate was  $q_s = 8.92 \times 10^{-7}\text{ kg s}^{-1}$ .

For solving the inverse problem, the sensitivity matrix  $\mathbf{A}$  (with  $m = 4$  lines and  $n = 10^6$  columns), the weighting matrix  $\mathbf{W}$ , and the source estimate  $\mathbf{s}_{//w}$  (both having  $n = 10^6$  nonnull elements) have to be computed. The lines of  $\mathbf{A}$  are obtained from the adjoint of an analytical plume model (Issartel and Baverel 2003). Details of the dispersion model, adjoint computations, and parameterizations are given in Turbelin et al. (2018). Once  $\mathbf{A}$  has been computed, the matrices  $\mathbf{W}$  and  $\mathbf{H}_w^{-1}$  are

TABLE 1. Number of iterations to obtain  $w$  for  $m = 100$  and  $n = 150\,000$ .

Tolerance $\varepsilon$	$10^{-5}$	$10^{-7}$	$10^{-10}$
Average number of iterations	8	15	25

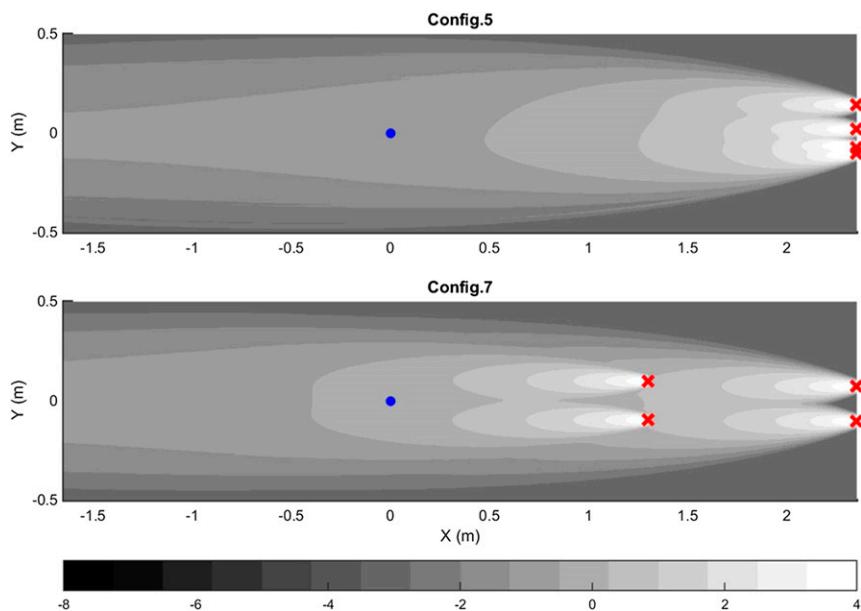


FIG. 3. Classical black and white logarithmic representation of the visibility function  $\varphi(\mathbf{x}) = w(\mathbf{x})/dS$  for DYCE configurations 5 and 7. The FFIDs are marked by red crosses, and the true source is marked by a blue circle.

obtained by using the algorithm described in Fig. 1. Here, the weights are obtained, with an error of  $O(10^{-7})$ , within about 15 iterations and with a CPU time of 15 s (the computation were performed on a desktop Intel Core i7-4790 3.60-GHz CPU and 16 GB RAM). The source estimate  $\mathbf{s}_{//w}$  is computed by using (4) with  $\boldsymbol{\mu}$  taken as the wind tunnel data averaged over 15 min. Then the source location and strength are retrieved from (7).

As illustrations, Fig. 3 shows the visibility function  $\varphi(\mathbf{x}) = w(\mathbf{x})/dS$ , and Fig. 4 shows the  $\mathbf{s}_{//w}$  maps and the estimated location of the source. It can be seen that the true source position is best monitored, with higher local weights, by network configuration 7. Nevertheless, better results are obtained for configuration 5: the point source location is retrieved within 0.13 m (wind tunnel scale) of the true source for configuration 5 and 0.19 m for configuration 7. Therefore, the relative location errors  $E_l$  (Euclidian distance of the retrieved source from the true source divided by  $X_m$ ) are, respectively, 5.8% and 10.5%. Similarly, the relative intensity error  $E_q = |q_s - \hat{q}_s|/q_s$  is 19% and 22.5%, respectively. These results are comparable to those obtained, with a weighted least squares approach, by Rudd et al. (2012) and confirm that a box array configuration poorly performs in comparison with a line configuration.

#### b. Application to the FUSION Field Trials

The second example uses data from the Fusing Sensor Information from Observing Networks (FUSION) Field

Trials 2007, described by Storwald (2007). The experiment involved 84 trials of short-duration releases (with multiple releases including instantaneous and continuous ones) of a passive tracer (propylene) in a flat terrain having a staggered rectangular arrangement of 100 fast-response digital photo ionization receptors in an area  $475 \text{ m} \times 450 \text{ m}$ . The receptors were arranged as an array of 10 rows and 10 columns. The spacing between two receptors was 50 m. The receptors measured the propylene concentration at the frequency of 50 Hz. The height of the source as well as samplers were 2 m above the ground level. The grid was set  $25^\circ$  (toward west) from north corresponding to prevailing wind flow. More details about the dataset are found in Storwald (2007) and Singh and Rani (2014). The representative trial 7, with stationary release conditions, is chosen here for illustration and discussion. This trial involved a point source, located close to the end of the grid, with an emission rate of  $5.53 \text{ g s}^{-1}$  during 10 min. The distance between the source and the center of the grid was  $X_m = 250 \text{ m}$ . Only 19 continuous nonzero concentration measurements have been reported and utilized for the inversion. For computations, we have chosen a domain of  $1200 \text{ m} \times 1200 \text{ m}$ , discretized into  $400 \times 400$  cells (cell size of 3 m). The true point source is located on the ground at cell (200, 200). The sensitivity matrix is computed from the adjoint of an analytical dispersion model proposed by Sharan et al. (1996). Details

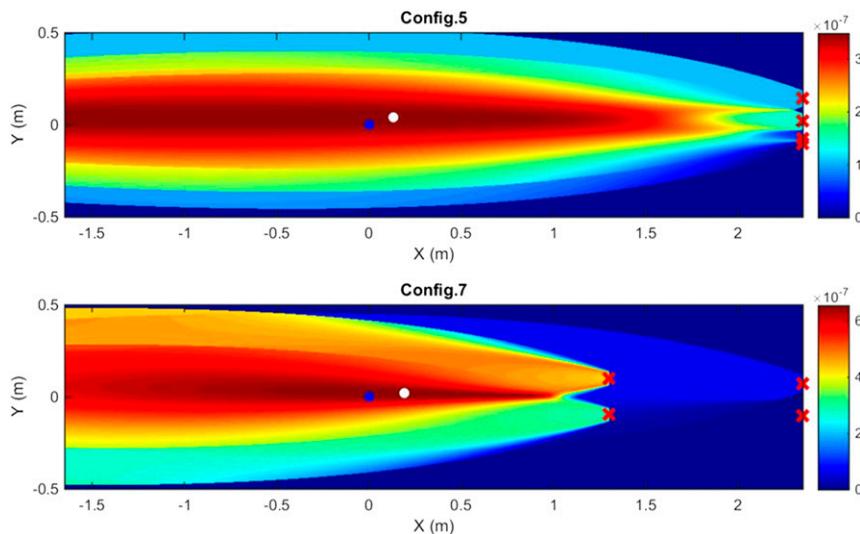


FIG. 4. The  $s_{//w}$  maps for DYCE configurations 5 and 7. The FFIDs are marked by red crosses, the true source is marked by a blue circle and the estimated source is marked by a white circle.

of the dispersion model, adjoint computations, and parameterizations are given in Singh and Rani (2014). The dispersion model, used in a forward mode, was able to describe the observed concentrations with limited model errors (that is with a normalized mean square error of 0.84, a fractional bias of 0.43, and correlation of 0.78). The inputs to the dispersion model, mean wind speed and direction, are provided as 10-min averages (corresponding to the release duration) at 4-m level from the tower located at the grid center.

The convergence of the algorithm, to compute the weights with an error of  $O(10^{-5})$ , is attained within 10 iterations (computational time is 20 s on a laptop; Intel Core i5-3427U 1.80-GHz CPU  $\times$  4 with 8-GB RAM). The source estimate is computed from (4), and the point source location is retrieved by searching the cell associated with maximum of the estimate  $s_{//w}$  in the domain. The strength of the retrieved point source is estimated from (7). As an illustration, Fig. 5 shows the visibility function and Fig. 6, the  $s_{//w}$  maps. It can be seen that the true source position is very well monitored by the network of receptors. As a consequence, good results are obtained: the point source location is retrieved within 5 m of the true source (corresponding to a relative location error of 2%), and the strength is retrieved within a factor of 1.2 to its true value.

**5. Conclusions**

In this paper, we have presented the renormalized data assimilation technique for solving the underdetermined source term estimation problem. The renormalized

source estimate is obtained after computing a (1, 2, 3)-generalized inverse of the sensitivity matrix of the network of sensors. This computation involves a positive diagonal weight matrix  $\mathbf{W}$ , which components satisfy the so-called renormalizing conditions. A compact and computationally efficient iterative algorithm for computing the elements of  $\mathbf{W}$  is proposed. Numerical experiments and comparisons with measurements from wind tunnel and field trials clearly show that the algorithm is efficient in terms of speed and accuracy and can be used in a source term estimation process to estimate the parameters of a hazardous atmospheric release.

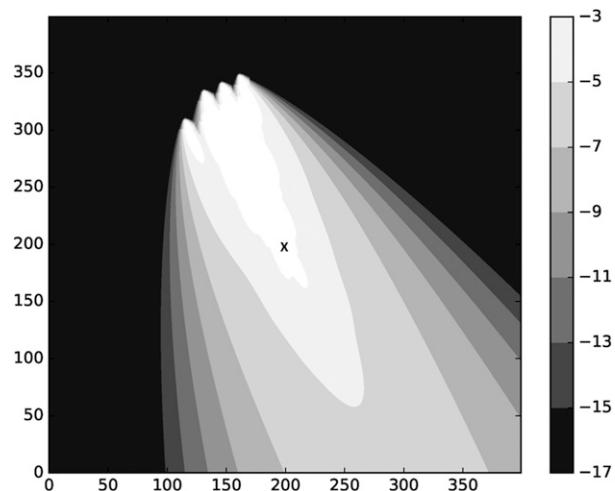


FIG. 5. Classical black and white logarithmic representation of the visibility function  $\varphi(\mathbf{x}) = w(\mathbf{x})/dS$  for FFT07 trial 7. The true source is marked by a black cross.

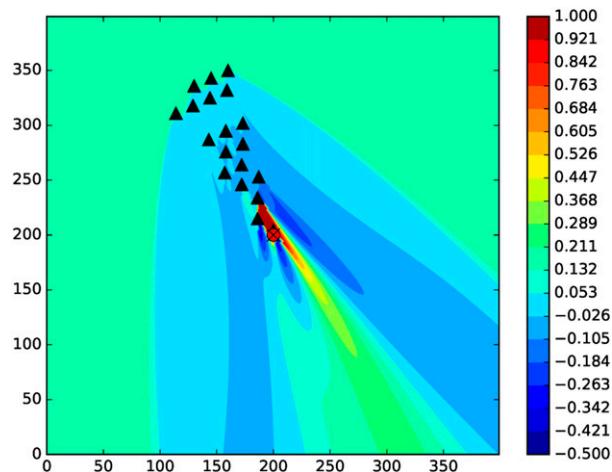


FIG. 6. The  $s/w$  maps for FFT07 trial 7. The FFIDs are marked by black triangles, the true source is marked by a black cross, and the estimated source is marked by a red circle.

It is desirable, in the future, to explore the advantages of this algorithm for reconstructing source terms in real complex scenarios.

## APPENDIX

### MATLAB Code

The following code (MATLAB, version 8.4) is proposed for computing, in a fast and accurate way, the optimal weights introduced in the renormalization inversion technique.

```
function[ w toc_algo iter] = renorm(A,pre)
% renorm: Computes the optimal weights for
% the renormalized assimilation technique
%
% Input parameters:
% A: Sensitivity matrix (m×N)
% pre: Accepted tolerance for the
% convergence criterion
%
% Outputs:
% w: Vector of optimal weights (N×1)
% toc_algo: CPU time to reach convergence
% iter: Number of iterations to reach
% convergence
%
%# Setting the variables
iter=0;
m=size(A,1);
N=size(A,2);
w=(m/N)*ones(N,1); %# Initial weights such
% as sum(w)=m
```

```
d=zeros(N,1);
tic; %# Starts the timer.
tA=A'; %# Transpose of the Weighted
% Sensitivity Matrix
%# While loop
while ~((max(abs(d-1))<=pre)
tAw=bsxfun(@rdivide,tA,w); %# Transpose of
% the Weighted Sensitivity Matrix
IHw=inv(A*tAw); %# Inverse of the Weighted
% Gram Matrix
d=sum(tAw*IHw.*tAw,2); %# Diag. of the
% Weighted Resolution Matrix
d(d==0)=min(d(d~=0)); %# To avoid
% problems with zero values
w=w.*d.^0.5; %# Iteration with new weights
iter=iter+1;
end
toc_algo=toc; %# Elapsed time since tic was
% used.
```

This MATLAB code is available in the online supplemental materials of this paper, along with a test file and a MATLAB script to compute configuration 5 of the DYCE experiments.

## REFERENCES

- Ben-Israel, A., and T. N. Greville, 2003: *Generalized Inverses: Theory and Application*. Springer, 420 pp.
- Grave de Peralta, R., O. Hauk, and S. L. Gonzalez, 2009: The neuroelectromagnetic inverse problem and the zero dipole localization error. *Comput. Intell. Neurosci.*, **2009**, 659247, <https://doi.org/10.1155/2009/659247>.
- Issartel, J.-P., 2005: Emergence of a tracer source from air concentration measurements: A new strategy for linear assimilation. *Atmos. Chem. Phys.*, **5**, 249–273, <https://doi.org/10.5194/acp-5-249-2005>.
- , and J. Baverel, 2003: Inverse transport for the verification of the Comprehensive Nuclear Test Ban Treaty. *Atmos. Chem. Phys.*, **3**, 475–486, <https://doi.org/10.5194/acp-3-475-2003>.
- , M. Sharan, and M. Modani, 2007: An inversion technique to retrieve the source of a tracer with an application to synthetic satellite measurements. *Proc. Roy. Soc.*, **463A**, 2863–2886, <https://doi.org/10.1098/rspa.2007.1877>.
- , —, and S. K. Singh, 2012: Identification of a point of release by use of optimally weighted least squares. *Pure Appl. Geophys.*, **169**, 467–482, <https://doi.org/10.1007/s00024-011-0381-4>.
- Kumar, P., A. A. Feiz, S. K. Singh, P. Ngae, and G. Turbelin, 2015: Reconstruction of an atmospheric tracer source in an urban-like environment. *J. Geophys. Res. Atmos.*, **120**, 12 589–12 604, <https://doi.org/10.1002/2015JD024110>.
- , S. K. Singh, A. A. Feiz, P. Ngae, and G. Turbelin, 2017: Assessment of a CFD model for short-range plume dispersion: Applications to the FUSION Field Trial 2007 (FFT-07) diffusion experiment. *Atmos. Res.*, **197**, 84–93, <https://doi.org/10.1016/j.atmosres.2017.06.025>.
- Menke, W., 2012: *Geophysical Data Analysis: Discrete Inverse Theory*. 3rd ed. Elsevier, 272 pp.

- Rudd, A. C., S. E. Belcher, A. G. Robins, and J. J. Lepley, 2012: An inverse method for determining source characteristics for emergency response applications. *Bound.-Layer Meteor.*, **144**, 1–20, <https://doi.org/10.1007/s10546-012-9712-y>.
- Sharan, M., A. K. Yadav, M. P. Singh, P. Agarwal, and S. Nigam, 1996: A mathematical model for dispersion of air pollutants in low wind conditions. *Atmos. Environ.*, **30**, 1209–1220, [https://doi.org/10.1016/1352-2310\(95\)00442-4](https://doi.org/10.1016/1352-2310(95)00442-4).
- Singh, S. K., and R. Rani, 2014: A least-squares inversion technique for identification of a point release: Application to FUSION Field Trials 2007. *Atmos. Environ.*, **92**, 104–117, <https://doi.org/10.1016/j.atmosenv.2014.04.012>.
- , G. Turbelin, J.-P. Issartel, P. Kumar, and A. A. Feiz, 2015: Reconstruction of an atmospheric tracer source in FUSION Field Trials: Analyzing resolution features. *J. Geophys. Res. Atmos.*, **120**, 6192–6206, <https://doi.org/10.1002/2015JD023099>.
- , P. Kumar, G. Turbelin, and A. R. Rani, 2017a: Uncertainty characterization in the retrieval of an atmospheric point release. *Atmos. Environ.*, **152**, 34–50, <https://doi.org/10.1016/j.atmosenv.2016.12.016>.
- , —, R. Rani, and G. Turbelin, 2017b: Comparative interpretations of renormalization inversion technique for reconstructing unknown emissions from measured atmospheric concentrations. *Earth Space Sci.*, **4**, 184–199, <https://doi.org/10.1002/2016EA000247>.
- Storwald, D. P., 2007: Detailed test plan for the Fusing Sensor Information from Observing Networks (FUSION) Field Trial (FFT-07). U.S. Army Tech. Rep. WDTC-TP-07-078.
- Turbelin, G., S. K. Singh, and J.-P. Issartel, 2014: Reconstructing source terms from atmospheric concentration measurements: Optimality analysis of an inversion technique. *J. Adv. Model. Earth Syst.*, **6**, 1244–1255, <https://doi.org/10.1002/2014MS000385>.
- , —, P. Nge, and P. Kumar, 2018: An optimization-based approach for source term estimations of atmospheric releases. *Earth Space Sci.*, **5**, 950–963, <https://doi.org/10.1029/2018EA000444>.