16th International Conference on
Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes
8-11 September 2014, Varna, Bulgaria

A Fast Reliable Algorithm For Point Source Localization: Application To A New Kitfox Data Set

Grégory Turbelin[1], Sarvesh Kumar Singh[1], Hamza Kouichi[1], Nicolas Bostic[1], Amir Ali Feiz[1], Pierre Ngae[1], Emerson Barbosa[1], Jean-Pierre Issartel[2]

[1]LMEE, Université d'Evry, 40 rue du Pelvoux, 91020 Evry Cedex, France
[2]DGA, Maitrise-NRBC, 5 rue Lavoisier, 91710 Vert le Petit Cedex, France

Abstract: In case of dispersion of hazardous CBRN species in the atmosphere, Source Term Estimation (STE) algorithms can be employed to provide an estimation of the release parameters (localization, intensity,…) by making use of concentrations and meteorological measurements. These algorithms have to be fast, reliable and accurate. One of the fastest methods for solving the source estimation inverse problem is the renormalization technique. It gives a source estimate linear with respect to the observations. This estimate corresponds to a minimum weighted-norm solution which can be expressed by making use of the generalized inverse concept. In this paper, a method for an efficient computation of this inverse solution is proposed. This method is illustrated using continuous release trials conducted during the Kitfox field experiment from September 11th to September 15th 1995, just after the 52 well known releases of August 1995. From the author's knowledge, these smooth desert experiments have never been used for validation purposes and thus can be considerate as a new data set.

***Key words**: Inverse problem, Renormalization, Source estimation, Generalized inverse, Kitfox dataset.*

INTRODUCTION

Significant recent advances in sensing technology have made field measurements less complex and more accurate. As a consequence, real time concentrations and meteorological measurements from array of sensors are now being used, by local authorities, to support rescue teams which respond to accidental or intentional dispersion of hazardous CBRN species in the atmosphere. Currently, these near-live measurements feed into models of atmospheric dispersion to get a clear picture of the situation and provide basics for decisions (COST ES1006, 2012). Successful forecasts from these models rely on an accurate estimation of the source strength and location. But, in most of the cases, these release parameters are not available. To characterize those parameters, several data assimilation techniques exist (Source Term Estimation algorithms). In general, for local scale emergency response, fast, safe and reliable source estimates are obtained from deterministic methods only using the observed data and by adding a minimum of a priori information to the problem. One of them, the renormalization technique, has been tested, and found to be working efficiently, against several Atmospheric and Tracer Data (ATD) sets from research studies from the past 20 years (Issartel et al., 2007, 2011, Sharan et al. 2009, 2012). Recently it has been demonstrated that the source estimate obtained by this approach can be interpreted as a minimal weighted-norm solution to the linear source term estimation problem. As a consequence, it can be written in terms of a weighted generalized inverse (or weighted Moore–Penrose inverse). In this paper, an appropriate method for efficient computations of the weight matrix and generalized inverse operator is proposed. Then, results computed from a series of dispersion experiments conducted over a flat open terrain are provided for illustrative purposes. From our knowledge, those experiments known as the "smooth desert – Kitfox" or "DRI /WRI CO2-II" experiments (Coulombe et al., 1999) have never been used, for validation purpose, during the last 20 years.

INVERSION TECHNIQUE

A continuous release of a tracer may be described, on a grid of N points, defined within the atmospheric domain, by a vector named the source vector and denoted as **s**. A component s_j (j=1…N) of this vector represents the rate of release at the location of coordinates $\mathbf{x_j}=(x_j,y_j,z_j)$. This source vector generates a field of concentrations $C(\mathbf{x_j})$. This field is known only through a finite number of observations $\mu_i=C(\mathbf{x_i})$ performed by a network of m sensors at locations x_i (i=1…m). The STE inverse problem consists in reconstructing $\mathbf{s} \in \mathbb{R}^N$ from the measurements vector $\boldsymbol{\mu} \in \mathbb{R}^m$. Such a problem is necessarily addressed by use of a model which describes the relationship between the measurements and the source vector. In the approach proposed by Issartel (2005), this relationship is represented by an integral equation which can be transformed into a system of linear equations:

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

where $\mathbf{A} \in \mathbb{R}^{N\times m}$ is a sensitivity matrix, the retroplumes matrix, with elements corresponding with the discrete values of adjoint functions a_i, $A_{ij} = a_i(\mathbf{x}_j)$. The retroplumes functions are obtained from physical modeling (dispersion models used in backward mode, Issartel et Baverel, 2003). The estimation of $\mathbf{s}$ from (1) is an ill-posed problem since the system is highly underdetermined (N>m). This under-determination gives rise to the non-uniqueness of the inverse solution. By using a diagonal weight matrix $\mathbf{W} \in \mathbb{R}^{N\times N}$ with elements satisfying the properties: (i) $w_{jj} > 0$ (ii) $\sum_{j=1}^{N} w_{jj} = \mathrm{m}$, the following change of variable can be applied to (1)

$$\boldsymbol{\mu} = \mathbf{A_w}\mathbf{W}\mathbf{s} \tag{2}$$

where $\mathbf{A_w}$ is the modified sensitivity matrix with components $A_{wij} = a_i(\mathbf{x}_j)/w_{jj}$. The estimation of $\mathbf{s}$ from (2) is still an ill-posed problem, but a unique minimum $\mathbf{W}$-weight norm solution can be obtained as

$$\mathbf{s_{//w}} = \mathbf{A_w^T}\mathbf{H_w^{-1}}\boldsymbol{\mu} \tag{3}$$

where $\mathbf{H_w} = \mathbf{A_w}\mathbf{W}\mathbf{A_w^T} = \mathbf{A}\mathbf{W^{-1}}\mathbf{A^T} \in \mathbb{R}^{m\times m}$ is a weighted Gram matrix (superscript "$^{\mathrm{T}}$" denotes the transposition). If the components of the weight matrix $\mathbf{W}$ satisfy a third condition, also called the "*renormalizing condition*", (iii) $diag\left(\mathbf{A_w^T}\,\mathbf{H_w^{-1}}\mathbf{A_w}\right) \equiv 1$, equation (3) provides the unique renormalized estimate. It has been demonstrated that, in case of noiseless observations generated from a single point source, the maximum value of this estimate corresponds to the location of the source. This localization accuracy is due to the condition (iii) which gives to the method optimal properties. Moreover, the optimal weight verifying this condition are currently interpreted as the discrete value of a visibility function $w_{jj} = w(\mathbf{x}_j)$ which characterizes the regions well or poorly monitored by the network. From a theoretical point of view, the optimal elements of $\mathbf{W}$ are the diagonal elements of the square projection matrix $\mathbf{R} = \mathbf{A_w^T}\,\mathbf{H_w^{-1}}\mathbf{A}$, when the diagonal elements of the symmetric matrix $\mathbf{R_w} = \mathbf{A_w^T}\,\mathbf{H_w^{-1}}\mathbf{A_w}$ are equal to one.

COMPUTATION OF THE METHOD

Equation (3) shows that $\mathbf{s}_{//\mathbf{w}}$ can be computed from the observed concentration $\boldsymbol{\mu}$ by using the generalized inverse operator $\mathbf{G} = \mathbf{A_w^T}\mathbf{H_w^{-1}} = \mathbf{W^{-1}}\mathbf{A}^{\mathrm{T}}(\mathbf{A}\mathbf{W^{-1}}\mathbf{A^T})^{-1}$. Since $\mathbf{W}$ is an Hermitian positive definite matrix, $\mathbf{G} \in \mathbb{R}^{N\times m}$ is unique and satisfies the following equations:

$$\text{(a) } \mathbf{AGA} = \mathbf{A} \quad \text{(b) } \mathbf{GAG} = \mathbf{G} \quad \text{(c) } (\mathbf{AG})^{\mathbf{T}} = \mathbf{AG} \quad \text{(d) } (\mathbf{WGA})^{\mathbf{T}} = \mathbf{WGA} \tag{4}$$

$\mathbf{G}$ is known as the $\mathbf{W}$-weighted Moore–Penrose inverse (or $\mathbf{W}$-weighted generalized inverse) of $\mathbf{A}$ often denoted by $\mathbf{G} = \mathbf{A_W^+}$ (Ben-Israel and Greville, 2003). Since $m < \mathrm{N}$ and $\mathrm{rank}(\mathbf{A}) = m$, $\mathbf{G}$ can be obtained by:

$$\mathbf{G} = \mathbf{W^{-1/2}}\left(\mathbf{A}\mathbf{W^{-1/2}}\right)^{+} \tag{5}$$

where the superscript "$^{+}$" denotes the Moore–Penrose inverse (or pseudo-inverse) of a matrix. Thus, to obtain **G,** the pseudo-inverse of the matrix $\mathbf{AW^{-1/2}}$ has to be computed. To that, several efficient methods exist, the most accurate one being based on the well-known Singular Value Decomposition (see for example the *pinv()* function in Matlab). But, to determine $\mathbf{G}$ from (5), the optimal weight matrix $\mathbf{W}$ satisfying the renormalized condition (iii) has first to be computed. To obtain the renormalizing weights, the first idea is to iteratively compute the projection matrices $\mathbf{R}$ and $\mathbf{R_w}$. But the dimension of these matrices is N×N and the storage, and processing time, resources of such matrices rapidly become prohibitive as N increases (if each entry is an 8-byte, 64 bits, double precision floating-point number, the required storage is $8*N*N/1024^3$ Go). As a consequence, to obtain $\mathbf{W}$, only the N diagonal elements of $\mathbf{R}$ have to be computed. By writing the j columns of the matrices $\mathbf{A}$ as vectors denoted as $\mathbf{a}_j \in \mathbb{R}^m$ (j=1…N), the diagonal components of $\mathbf{R}$ and $\mathbf{R_w}$ can be respectively written as

$$R_{jj} = w_{jj}^{-1}\mathbf{a_j}^{\mathrm{T}}\mathbf{H_w^{-1}}\mathbf{a_j} \text{ and } R_{wjj} = w_{jj}^{-2}\mathbf{a_j}^{\mathrm{T}}\mathbf{H_w^{-1}}\mathbf{a_j} \tag{6}$$

In this case, the renormalisation condition is

$$w_{jj}^{-2}\mathbf{a_j}^{\mathbf{T}}\mathbf{H_w^{-1}}\mathbf{a_j} = 1 \text{ or } w_{jj} = \sqrt{\mathbf{a_j}^{\mathbf{T}}\mathbf{H_w^{-1}}\mathbf{a_j}} \text{ for } j = 1\ldots\mathrm{N} \tag{7}$$

Thus $\mathbf{W}$ can be computed as the converged diagonal matrix $\mathbf{W_p}$ (with $\mathbf{W_0}$=m/N*$\mathbf{I_N}$) obtained from the following iterative Algorithm 1. It has been observed that this algorithm converges uniformly to a matrix $\mathbf{W}$ with components satisfying the renormalization condition. Once this optimum weight matrix has been obtained, the generalized inverse can be computed from (5) and, for a given set of measurements, the source estimate $\mathbf{s}_{//\mathbf{w}}$ can be calculated from (3). The location of the maximum component of this vector $s_{//\mathbf{w}}^k$ corresponds to the location of the actual point source and the intensity of the release is given by $s_{//\mathbf{w}}^k / w_{kk}$.

Algorithm 1. Computing the optimal weighted matrix $\mathbf{W}$

Require: Let $\mathbf{A} \in \mathbb{R}^{m\times N}$.
N=columns[$\mathbf{A}$]
m=rows[$\mathbf{A}$]
$\mathbf{W}$=m/N*$\mathbf{I_N}$ (initialization of $\mathbf{W}$)
while d_{min} < =0.99 (definition of the convergence criteria)
    $\mathbf{H}^{-1}=(\mathbf{A}\mathbf{W}^{-1}\mathbf{A}^{\mathbf{T}})^{-1}$ (computation and inversion of the weighted Gram matrix)
        For j=1 to N
            $\mathbf{a}_j=\mathbf{A}(j)$ (writing the j columns of the matrices $\mathbf{A}$ as vectors)
            $\mathbf{d}_j= \mathbf{a_j}^{\mathbf{T}}\mathbf{H}^{-1}\mathbf{a_j}*w_{jj}^{-2}$ (computation of the diagonal elements of $\mathbf{R_w}$, stored in a vector $\mathbf{d}$)
        End for
    d_{min} = min($\mathbf{d}$) (convergence verification)
    $\mathbf{W}=\mathbf{W}*(\mathrm{diag}[\mathbf{d}])^{1/2}$ (definition of a new weight matrix)
end while
return $\mathbf{W}$

THE SMOOTH DESERT KITFOX DATA SET

In late August and early September 1995, a series of dispersion experiments were conducted at the Nevada Test Site. In all the experiments of this series, known as the "KIT FOX series", carbon dioxide (CO_2) was released as a tracer. The gas releases were conducted, during stable and neutral atmospheric conditions, under three different surface roughness configurations: Equivalent Roughness Pattern (ERP), to simulate a refinery tank farm at reduced scale, Uniform Roughness Array (URA) to simulate the surface surrounding a refinery and smooth desert surface configuration. High-concentration and low-concentration CO_2 sensors were laid out in four arrays located downwind from the source and oriented perpendicular of the predicted transport of the cloud. Meteorological conditions were measured by adequate instrumentation both over the flat desert surface and within the dispersion grid. From August 22 to August 31, 70 releases have been conducted under two separate surface configurations: ERP+URA and URA only. Those releases were described in a well-documented report (Western Research Institute, 1998). From criteria recommended by G. Briggs, 52 of them have been selected. Those appropriate releases, and some analyses of the data, have been described by Hanna and Chang (2001). As part of the Modelers' Data Archive (MDA), these 52 experiments have been extensively used for evaluating model performance or for developing new algorithms. From September 11 to 15 1995, 30 releases with the smooth desert configuration were conducted. These experiments, also referred to as the “DRI /WRI CO2-II” experiments, have been described in a separate report (Coulombe et al., 1999), only mentioned by King et al. (2002). From the author’s knowledge, these late experiments have never been used for validation purposes and thus the “smooth desert - Kitfox data set” can be considerate as a new data set. Thanks to J. Chang and S. Hanna, W. Coulombe sends us his report in April 2013.

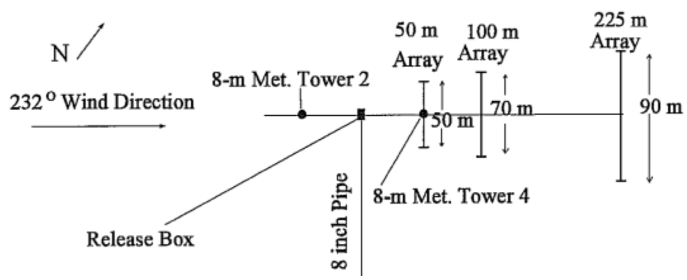

Figure 1: KITFOX, smooth desert surface configuration, from Coulombe et al. (1999)

For these experiments, the sensors were laid out in three arrays of sensors located downwind from the source (at 50m, 100m and 225m) and oriented perpendicular to the centreline of the predicted transport course of the cloud. The meteorological measurements (wind components and temperature at different levels) were made by using a permanent 24-m tower (located southwest of the site) and two 8-m towers located, on the nominal 232° centreline, upwind and downwind of the source location, see Figure 1. The 30 releases were performed under atmospheric stability ranging from neutral (Pasquill D) to very very stable (designated in the report as Pasquill G). A release rate of 1.5 kg/s was used for the 22 short duration releases (20s). For the 8 other experiments, the gas was released continuously over longer periods of time (150-360s) at a rate of 1.5 kg/s (3 experiments) and 1 kg/s (5 experiments). In this study, some of the concentration and meteorological measurements from those 8 continuous release experiments have been used for illustrating the renormalization technique.

Table 1: Characteristics of the continuous release experiments

Test No.	Average wind speed 2m a.g.l. (ms^{-1})	Average wind direction 2m a.g.l. (degree)	Release rate (kgs^{-1})	Release duration mm:ss	Stability Class
9-4	3.5	234	1.527	2:31	D-E
9-7	2.5	229	1.497	3:31	F
9-9	1.9	235	1.438	5:31	F-G
10-5	2.0	232	1.037	5:58	G+
10-6	1.9	198	0.995	5:00	F-G
12-7	1.6	211	1.019	4:59	F-G
13-6	3.0	227	1.114	3:32	E
13-7	2.3	213	1.028	3:00	F

RESULTS & DISCUSSION

In local emergency response applications, adequate results from the available observations are needed as quickly as possible. This tends to favor simplest models, such as the one used in this study to compute the components of the sensitivity matrix **A**. Indeed, use was made of an analytical Gaussian dispersion model (Sharan et al., 1996) in a backward mode (since dense gas effects dominate only in immediate vicinity of the source, it has been assumed that by the time the plume reaches the captors, it should likely be modelled as Gaussian). The values presented in Table 1, and the empirical estimate of Briggs for the lateral and vertical standard deviations of dispersion, have been used as parameters. The observed concentrations from 24 captors of the 50m and 100m arrays have been averaged over the sampling duration in order to obtain the measured concentrations vector $\boldsymbol{\mu}$ (i.e. m=24). The inversion technique described in the previous section has been implemented on a discretized domain of N=300×300 points with a resolution of 1m. With this grid, the computational (CPU) time on a machine Intel® Core™ i5-3427U CPU 1.80GHz, 8Go RAM, involved in estimating the components of **A**, **W** and $\mathbf{s}_{//\mathbf{w}}$ was approximately 30 seconds.

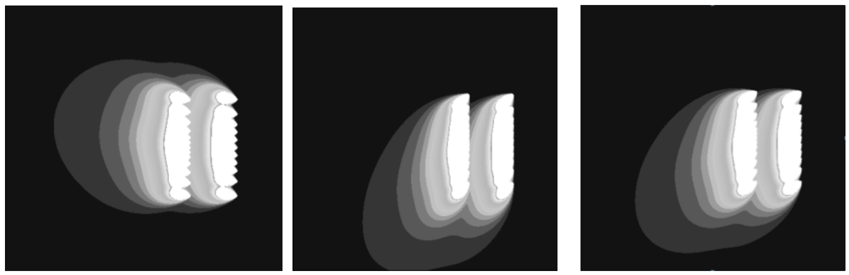

Figure 2: Isopleths of $w(\mathbf{x})$ in the computational domain for cases 9-9,12-7 and 13-7 (from left to right).

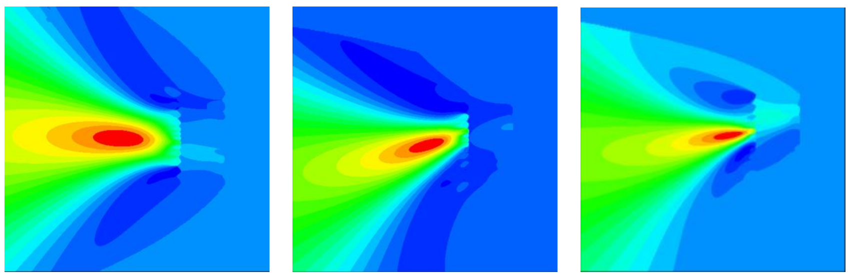

Figure 3: $\mathbf{s}_{//\mathbf{w}}$ for cases 9-9,12-7 and 13-7 (from left to right). The center of the domain denotes the source location.

In every case, the function $w(\mathbf{x})$ is focused at the detectors locations and decreases with increasing downwind distance from the monitoring network. This function is currently interpreted as the visibility of the monitoring network characterizing the regions well or poorly monitored. If the value of the visibility function at a point is nearly zero (black regions in Figure 2), a source at that point will be hardly identified. In this study, the source location lies in a well monitored region of the network, see Figure 2. As a consequence, a sharp maximum region of $\mathbf{s}_{//\mathbf{w}}$ is distinctly observed, see Figure 3. The maxima of $\mathbf{s}_{//\mathbf{w}}$ is unique at position (x_s,y_s). The lateral positions of the source are retrieved within 5 m (i.e $0 \le \Delta y_s/x_m \le 0.08$, where x_m is the mean distance between the source and sensors in the downwind direction). This accuracy is mainly due to the configuration of the network with two arrays across the plume. The estimates for x_s are, on the whole, of lower quality than those for y_s ($0 \le \Delta x_s/x_m \le 0.2$). The longitudinal source locations are placed upstream of the true position, basically because the sensitivity matrix has been derived from a Gaussian model (which includes the use of mean wind speed and direction) associated with empirical dispersion parameters. In practice, the use of empirical parameters will inevitably lead to uncertainty in source term localization. This model error can be reduced by using turbulence-based schemes for lateral and vertical dispersion.

CONCLUSION

For local scale emergency response, fast and reliable source estimates can be obtained from limited concentration measurements by using the renormalization method. The renormalized estimate, given in a discrete form by equation (3), can be obtained by using a weight generalized inverse operator. This operator satisfies criteria used to define good inverses. It can be easily obtained from equation (5) after having computed the pseudo-inverse of a weighted sensitivity matrix. To compute pseudo-inverses, classical algorithm based on a Singular Value Decomposition method can be used. In this paper, an algorithm to compute the optimal weight matrix has been proposed. The components of this matrix are currently interpreted as the discrete value of a visibility function which characterizes the regions well or poorly monitored by the monitoring network. By using this optimal matrix, the maximum value of the source estimate corresponds to the location of the point source. The study has been illustrated using measurements from a new KITFOX data set. By using mean measurements, the method converges onto reasonable estimates: the source is observed to be distinctly located and lies in a highly illuminated region of the monitoring network. The lateral position of the point source is retrieved within 5 m of the true source location but the longitudinal location is placed upstream of the true location. The errors on source localization, observed in this study, can be imputed to the performances of the models (and on the parameters which they are associated) used to obtain the sensitivity matrix.

Acknowledgments. We thank Bill Coulombe for sharing the data and the report. We also thank Steve Hanna, Joe Chang and Bruce King who help us in “tracking” this “forgotten” data set.

REFERENCES

Ben-Israel, A., and T.N.E. Greville, 2003: Generalized Inverses: Theory and Applications, Springer, New-York.

Coulombe, W., J. Bowen, R. Egami, D. Freeman, D. Sheesley, J. Nordin, T. Routh, and B. King, 1999: Characterization of Carbon Dioxide Releases−Experiment Two, DRI Doc. No. 97-7240.F, DRI, P.O. Box 60220, Reno, NV 89506-0220.

COST ES1006, 2012: Background and Justification Document, University of Hamburg.

Hanna, S.R. and J.C. Chang, 2001: Use of the Kit Fox field data to analyze dense gas dispersion modeling issues, *Atmos. Environ*, **35**, 2231-2242, doi: 10.1016/S1352-2310(00)00481-7.

Issartel, J.-P., and J. Baverel, 2003: Inverse transport for the verification of the Comprehensive Nuclear Test Ban Treaty, Atmos. Chem. Phys., 3, 475-486, doi:10.5194/acp-3-475-2003.

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, doi:10.5194/acp-5-249-2005

Issartel, J.-P., 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. A*, **463**, 2863-2886, doi:10.1098/rspa.2007.1877.

King, S. B., J. S. Nordin, D. Sheesley, T. Routh, 2002: Chapter 23: U.S. DOE Hazmat Spill Center Database, in The Handbook of Hazardous Materials Spills Technology, M. Fingas editor, McGraw-Hill.

Sharan, M., M. P. Singh, A. K. Yadav, P. Aggarwal and S. Nigam, 1996: A mathematical model for the dispersion of pollutants in low wind conditions. *Atmospheric Environment*, **30**, 1209-1220, doi:10.1016/1352-2310(95)00368-1.

Sharan, M., J.-P. Issartel, S. K. Singh and P. Kumar, 2009: An inversion technique for the retrieval of single-point emissions from atmospheric concentration measurements. *Proc. Roy. Soc. A*, **465**, 2069-2088, doi:10.1098/rspa.2008.0402.

Sharan, M., J.-P. Issartel and S. K. Singh, 2012: A point-source reconstruction from concentration measurements in low-wind stable conditions. *Q.J.R. Meteorol. Soc*., **138**, 1884–1894, doi:10.1002/qj.1921.

Western Research Institute (WRI), 1998: The 1995 KITFOX Project; Volume-I, Experiment Description and Data Processing, WRI, Laramie Wyoming.