A measure theoretic approach to linear inverse atmospheric dispersion problems

Niklas Brännström¹ and Leif Å Persson¹,²

¹ Swedish Defence Research Agency, FOI, SE-901 82 Umeå, Sweden
² Department of Mathematics and Mathematical Statistics, Umeå University, SE-901 87 Umeå, Sweden

E-mail: niklas.brannstrom@foi.se

Received 17 September 2014, revised 2 December 2014
Accepted for publication 30 December 2014
Published 30 January 2015

Abstract
Using measure theoretic arguments, we provide a general framework for describing and studying the general linear inverse dispersion problem where no a priori assumptions on the source function has been made (other than assuming that it is indeed a source, i.e. not a sink). We investigate the source-sensor relationship and rigorously state solvability conditions for when the inverse problem can be solved using a least-squares optimization method. That is, we derive conditions for when the least-squares problem is well-defined.

Keywords: atmospheric dispersion, inverse problem, measure theory

(Some figures may appear in colour only in the online journal)

1. Introduction

All atmospheric dispersion models share the goal of spatio-temporal forecasting of concentrations of pollutants released into the atmosphere. There are many applications, e.g. planning where a factory should be located (to reduce the risk in case of an accident) which requires mainly a local model, or e.g. forecasting which regions that would be affected by a nuclear power plant incident (like the Fukushima disaster) which mainly requires a regional or global model. An equally natural question to ask is: given that we have detected a pollutant somewhere, can we deduce where the source was located? If not before, this inverse problem became very important in the wake of the Chernobyl accident. In that case the radioactive pollution triggered sensors in Europe before any news of the accident was released. Pinpointing the location of the source could be done by guessing the location, strength, and time of the accident and running the dispersion model forward to see whether it would give the observed measurements. Unless the guess is an educated one this can be a costly process. The
alternative is to solve the inverse problem. Having a solution to the inverse problem, that is, an estimate of the parameters in the source function, enables subsequent forward dispersion modelling to gain a much better understanding of the current state of affairs (a better situation analysis). Alternatively the source estimate may be a crucial part of forensic work, for example trying to calculate the amount of leaked radioactive substances following the accidents in Chernobyl [1] and Fukushima [2] or pinpointing nuclear test sites [3].

A number of methods to solve the inverse problem have been suggested. In addition to the two main contenders optimization algorithms and Bayesian statistics there are methods like footprint analysis, e.g. the survey article [4], influence area [5] and [6], directly inverting the problem and trying to overcome any issues associated with ill-conditioning, see e.g. [7]. Often the methods are designed to bear only on a subclass of inverse dispersion problems by a priori conditioning on the number of sources, the type of source, or the dispersion model employed. In the Bayesian approach to the inverse problem the source is estimated from a so called a posteriori probability distribution function which is obtained by calculating a likelihood function and weighing it with any a priori information that one has at hand (see e.g. [8] for an introduction to general Bayesian inverse problems, and [9] for an early reference). This method avoids the pitfalls of ill-conditioning which are often associated with directly inverted problems and adds the benefit of allowing uncertainties in models and measurements to be handled in a tractable fashion. In a series of papers the Bayesian approach has been adapted to bear on inverse dispersion problems: in [10] the case with one source with unknown position and unknown but constant source strength was treated. [7] deals with the case where there is a known number of sources in given locations but where the source strengths are unknown (there is also an interesting comparison of the results to those obtained with a directly inverted model where the problems of ill-conditioning have been alleviated by singular value decomposition). This study was generalized in [11] and [12] to cover the situation where there is an unknown number of sources in unknown positions, where the only assumption on each source is that during emission the source strength is constant. The case with an unknown number of sources is much harder than working with a fixed number of sources as the dimension of the parameter space is unknown. In [11] and [12] this problem was overcome by using the method reversible jump Monte Carlo Markov chain [13] to sample from the posterior probability distribution function with an unknown number of dimensions (the dimension is one of the parameters that needs to be estimated). In [14] a recursive method is proposed to deal with same issue.

Under the umbrella of the optimization method we find all the various ways of setting up the inverse dispersion problem so that its solution is given as the solution of a least-squares fitting problem. As for the Bayesian method, the body of literature mostly covers the case where it is a priori known that there is only one single source, see e.g. [15–17], and [18]. There are exceptions, e.g. in [19] the least-squares method presented in [18] is generalized to cover an unknown number of point sources, and in [21] the space–time has been discretized and the optimal source term is constructed by forming a union of ‘box-sources’ (the smallest resolution is given by the grid box, so ‘box-source’ seems the appropriate term instead of point source).

One of the referees brought to our attention a series of papers by Butler et al notably [20], where measure theory is used on finite-dimensional source parameter spaces to handle set-valued solutions to inverse problems, as well as measure theoretic formulations of stochastic inverse problems. Their approach is complementary to ours; their analysis is restricted to specific finite-dimensional parametrizations of the sources, but also provides more detailed information on the structure of the solution space, as well as numerical methods.

In this paper we are developing a framework for describing inverse dispersion problems. The framework relies on using measure theoretic ideas and methods to study the general linear inverse dispersion problem without making a priori assumptions on e.g. the number of
sources, their emission patterns or their distribution in the spatio-temporal domain. As such, the framework is non-parametric but since the term non-parametric seems to be overloaded we refrain from using it to describe the framework. We begin by setting up the linear inverse problem and then we present a one-dimensional toy problem that motivate the use of measures rather than probability densities. Then we turn to the problem of determining under which conditions a given set of sensor data can be generated by a source chosen from a given class of sources. As a warm-up we consider linear combinations of base source measures in both the invertible case and the over determined case. The arguments are based on finding appropriate cones in the space of positive measures (describing the source) and in the space of sensor measurements. We then build on this to generalize the analysis to the case where the source is chosen from a closed cone of measures (we dispense with the assumption of having a finite number of base sources) resulting in theorem 11. While theorem 11 is certainly interesting in its own right, explaining when a measurement can be realized, the analysis also allows for a derivation of the main result of the paper: conditions under which the least squares optimization problem is well-defined, theorem 18. In addition to these results we also characterize the set of measurements when a source is approximated by a sequence of instantaneous point sources, theorem 12.

The measure theoretic approach that is presented in this paper introduces a machinery which we believe will be useful in future studies where rigorous results on general linear inverse dispersion problem are sought. Indeed, while not solving any particular inverse dispersion problem, the method is not hampered by any peculiarities that a given set of parameters could have introduced.

2. Setting of the problem, the dispersion model and its adjoint

The atmospheric dispersion problem that we are interested in can be formulated in terms of a transition probability $p(t, x; t^*, x^*)$, where $(t^*, x^*)$, $(t, x) \in T \times V$, where $T \subset \mathbb{R}$ is a time interval and $V \subset \mathbb{R}^3$ is a spatial domain. The transition probability density expresses the probability for a particle released at the time–space point $(t^*, x^*)$ to reside in the time–space point $(t, x)$ for $t \geq t^*$. We note that $p = 0$ when $t < t^*$. The particles whose dispersion is governed by this transition probability is assumed to originate from a source $S$. The source $S$ is assumed to be a positive measure on $T \times V$ (that is, the word 'source' is used in the strict sense; no sinks are considered in this paper). In this way the total mass $M$ released from the source is given by integrating the source measure $S$ over its support

$$M = \int_T \int_V dS(t^*, x^*). \quad (1)$$

The quantity that is usually desired as output from a dispersion model is the concentration of the pollutant in a given space–time point. Since $S$ has its support on $T \times V$ and the transition probability describes the dynamics of the released substance the concentration $c(t, x)$ is obtained by weighing all released particles (released at some $(t^*, x^*)$ with $t^* < t$) with the probability that they have been transported from $(t^*, x^*)$ to $(t, x)$

$$c(t, x) = \int_T \int_V p(t, x; t^*, x^*)dS(t^*, x^*). \quad (2)$$

While $c(t, x)$ is the predicted concentration at the space–time point $(t, x)$ the sensor may not have the resolution to make an ideal measurement from the concentration field $c(t, x)$, indeed the sensor may perform some form of averaging in both space and time to yield the sensor
response $\tilde{c}(t, x)$. We assume that the averaging process in the sensor can be described by a probability measure $S^*$ (usually referred to as the sensor-filter function) on $T \times V$, and hence we express the sensor response as

$$\tilde{c} = \int_T \int_V c(t, x) dS^*(t, x) \equiv \langle S^*, c \rangle.$$ 

According to Fubini’s theorem ([22], p 164), this can be written as

$$\langle S^*, c \rangle = \int \int \int p d\left( S \times S^* \right) = \langle S, c^* \rangle,$$

where

$$c^*\left(t^*, x^*\right) = \int_T \int_V p\left(t, x; t^*, x^*\right) dS^*(t, x).$$ (3)

In the case where $S$ and $S^*$ are given by square-integrable space–time densities $dS(t, x) = s(t, x) dx dt$, $dS^* = s^*(t, x) dx dt$, then $c$ and $c^*$ are also square-integrable space–time densities, and

$$\tilde{c} = \left( c, s^* \right) = \left( c^*, s \right).$$ (4)

where the inner product is defined by $(f, g) = \int \int f(t, x) g(t, x) dx dt$. Therefore, $c^*$ is called the adjoint concentration. We consider here a more general situation where the inner product $(\cdot, \cdot)$ is replaced by the duality product $\langle \cdot, \cdot \rangle$, and sources $S$ and measurements $S^*$ are allowed to have singular parts, for example combinations of instantaneous or continuous point, line or area sources. Then $c$ and $c^*$ defined by (2) and (3) are still functions, but they may have singular points in space and time. For example, if $S = \delta_{(t^*, x^*)}$ then $c(t, x) = p(t, x; t^*, x^*)$ has a singularity at $(t^*, x^*)$, and likewise, if $S^* = \delta_{(t, x)}$ then $c^*(t^*, x^*) = p(t, x; t^*, x^*)$ has a singularity at $(t, x)$. When $c$ is not continuous, $\langle S^*, c \rangle$ is not defined for all probability measures $S^*$; $c$ must be integrable with respect to $S^*$. For example, if $c$ has a singularity at $(t, x)$, then $c$ is not integrable with respect to $S^* = \delta_{(t, x)}$; this measure can only be applied to functions which are continuous at $(t, x)$. Instantaneous point measurements and instantaneous point sources are not allowed at the same space–time point.

We note that equation (3) describing the adjoint concentration field is evolving backwards in time: we may view the transition probability as moving adjoint particles released by $S^*$ backwards in time and space. The main advantage of using the adjoint representation in inverse dispersion modelling is computational efficiency. This is a well-documented fact, see for example [23]. We also remark that the adjoint concentration field $c^*$ is independent of the source function $S$, and the concentration field $c$ is independent of the sensor-filter function $S^*$.

3. Source-receptor relationship

The dispersion problem predicts how a pollutant from a source spreads in the atmosphere. From an abstract point of view this problem can be seen as a problem of mapping of measures: the source $S$ can be viewed as a measure in the spatio-temporal domain $T \times V$ that is being mapped via the dispersion equations into a scalar function $c$ (the concentration), from which we make measurements represented by a probability measure $S^*$, defining the averaging of the concentration function $c$. From this level of abstraction the adjoint version of the
problem is very similar. In this case the adjoint equations maps a probability measure \( S^* \) on \( T \times V \) representing a measurement in a sensor to a scalar function \( c^* \) (adjoint ‘concentration’) from which we can make ‘adjoint measurements’ using a source measure \( S \) acting on the adjoint ‘concentration’ \( c^* \). (Depending on the scaling of the problem the adjoint ‘concentration’ \( c^* \) may not be a proper concentration dimensionally.) In light of this, asking questions about the sensor response in the forward problem or asking questions about the source in the inverse problem are very similar. Based on this observation we therefore propose to adopt a measure theoretic approach and we develop a mathematical framework for studying the inverse problem. While we are omitting the analysis of the forward problem in this paper we note that treating this problem is completely analogous. Studying the problem in this generality will not allow us to solve any particular inverse dispersion problem, but it will allow us to draw general conclusions about whole classes of problems. One particular advantage of this approach hence lies in the fact that we avoid difficulties that may be associated with a particular problem and its parameters—of course, these will have to be addressed when the particular problem is to be solved.

4. One-dimensional example to motivate the use of measure theory

As a model example, consider a stationary one-dimensional diffusion on the unit interval with absorbing boundary conditions

\[
- \frac{c}{x}(x) = S(x), \quad x \in [0, 1], \quad c(0) = c(1) = 0.
\]

The solution \( c(x) \) is a concave function; using the integral formula of Blaschke and Pick [24] the solution can be written

\[
c(x) = \int_0^1 \frac{y(1-y)}{\sqrt{3}} \hat{\phi}(x, y) S(y)dy,
\]

where

\[
\hat{\phi}(x, y) = \begin{cases} 
\sqrt{3}x/y & \text{if } 0 \leq x \leq y, \\
\sqrt{3}(1-x)/(1-y) & \text{if } y \leq x \leq 1.
\end{cases}
\]

This formula is also valid if \( S \) is a unit point mass at a fixed point \( y \), in which case the concentration profile is

\[
c(x) = \frac{y(1-y)}{\sqrt{3}} \hat{\phi}(x, y) = \min(x(1-y), y(1-x)) \equiv f(x, y)
\]

and in case of a point measurement at a fixed point \( x \) we have \( c^e(y) = f(x, y) \). Given a finite number of measurement points \( x_1, \ldots, x_m \) a vector of measured values \( \hat{c} = (c_1, \ldots, c_m) \) is the result of a smooth density \( S(y) \) if and only if the points \( (0, 0), (x_1, c_1), \ldots, (x_m, c_m), (1, 0) \) lie on the graph of the smooth concave function \( c(x) \) given by the formula above. Likewise, \( \hat{c} \) is the result of a point source \( S \) at \( y \) if and only if the same points lie on the graph of a function \( \lambda f(\cdot, y) \) for some \( \lambda > 0 \).

3 The basis functions \( \hat{\phi} \) are normalized so that \( \int_0^1 \hat{\phi}^2(x, y)dx = 1 \).
We want the set of measurement vectors $\tilde{c}$ to be closed, so that we can determine the closest measurement vector from any given vector. Taking a sequence $c_j$ of smooth convex functions converging pointwise towards $f(\cdot, x_k)$ for some $1 < k < m$ we conclude that the vector $\tilde{c} = (\tilde{c}(x_1, x_k), \ldots, \tilde{c}(x_m, x_k))$ should be included. The points $(0, 0), (x_1, \tilde{c}_1), \ldots, (x_k, \tilde{c}_k)$ are collinear, and likewise $(x_{k+1}, \tilde{c}_{k+1}), \ldots, (x_m, \tilde{c}_m), (1, 0)$, and the only concave function containing these points in its graph is $f(\cdot, x_k)$ so $\tilde{c}$ must come from a point source at $x_k$. Hence, point sources must be allowed. Since any measure can be locally approximated (by weak convergence of measures) by a sequence of finite linear combinations of point sources, it is natural to allow sources given by finite measures.

5. Linear combinations of sources

The purpose of this section is to characterize all possible measurement values obtainable when $S$ is a linear combination with positive coefficients of a given finite number of base sources. In other words, we now investigate under which condition there exists a measure $S$ which will produce the concentration measurements exactly. Finding a source $S$ reproducing a value $\tilde{c}$ for a measurement $S^c$ is easy; simply take an arbitrary source that gives a positive measurement value and scale the source properly. Let us try the same idea for several measurements $S^c_1, \ldots, S^c_n$ by considering sources $S_j, j = 1, \ldots, n$ and forming the linear combination $S = \sum_{j=1}^n \lambda_j S_j$ with $\lambda_j \geq 0$ (we only consider $\lambda_j \geq 0$ since we want all $S_j$ to contribute as sources, were some $\lambda_j$ allowed to be negative the corresponding ‘source’ $S_j$ would act as a sink, even if $S$ could still be positive). Given the measured values $c_1, \ldots, c_m \geq 0$ we get the linear system of equations

$$\sum_{j=1}^n a_{ij} \lambda_j = c_i, \quad \text{where } a_{ij} = \langle S_j, c^*_i \rangle$$

and we denote $A = (a_{ij})$, which is sometimes called the source–receptor matrix. Assume first that $A$ is invertible, i.e., $m = n$ and the measurement vectors $\langle S_j, c^*_i \rangle, j = 1, \ldots, n$ are linearly independent. Then, since $A$ is an invertible non-negative matrix (by non-negative matrix we mean a matrix where all elements are non-negative and this is denoted $A \geq 0$), the inverse $A^{-1}$ contains nonpositive elements on row $i$ if $A$ contains off-diagonal positive elements in column $i$ (see the remark below for justification). Hence the condition that $\lambda_i \geq 0$ gives a linear constraint

$$-\sum_{j \in J^+_i} (a^{-1})_{ij} \tilde{c}_j \leq \sum_{j \in J^-_i} (a^{-1})_{ij} \tilde{c}_j,$$

where $J^+_i$ denotes the set of column indices $j$ for which $(a^{-1})_{ij} > 0$ and $J^-_i$ denotes the set of column indices $j$ for which $-(a^{-1})_{ij} > 0$.

Remark 1. Suppose that $A$ is invertible and $A \geq 0$. The row vectors $A_i^{-1}$ of $A^{-1}$ and the column vectors $A_j$ of $A$ satisfy $A_j^{-1} A_j = \delta_{ij}$. Suppose that $A_i$ contains $k$ positive components, e.g., $A_i = a_1 e_1 + \ldots + a_k e_k$ with $a_l > 0, l = 1, \ldots, k$, where $e_i$ are the standard orthonormal basis vectors in $\mathbb{R}^n$. If $j \neq i$ and $A_j^{-1} = \beta_1 e_1 + \ldots + \beta_k e_k$ then either $\beta_1 = \ldots = \beta_k = 0$ or $\beta_i < 0$ for some $1 \leq l \leq k$. In the former case we have $A_j^{-1} = \beta_{k+1} e_{k+1} + \ldots + \beta_k e_k$. There can be at most $n - k$ such $A_j^{-1}$s since the $A_j^{-1}$s are linearly independent. Hence there are at least $k - 1$ column vectors $A_j^{-1}$s with $j \neq i$ that
contain negative elements. Therefore, both the positive and negative parts \((A^{-1})^+_{\mu} = \max\left(0, (A^{-1})_{\mu}\right)\) and \((A^{-1})^-_{\mu} = \max\left(0, -(A^{-1})_{\mu}\right)\) are nonzero, and the non-negativity conditions \(\lambda = A^{-1}\epsilon \geq 0\) give the linear constraints

\[
(A^{-1})^+ \epsilon \leq (A^{-1})^- \epsilon.
\]

The general case requires more work: \(A\) is not necessarily positive or invertible, it may even be a non square matrix. However, it may be solved as a minimization problem. Indeed, the problem of finding the ‘best’ non-negative solution \(x \in \mathbb{R}^n, x \geq 0\) to the linear system \(Ax = b\), where \(A \in \mathbb{R}^{m \times n}\) and \(b \in \mathbb{R}^m\) are given, can be formulated as a quadratic minimization problem, namely: find \(x \in \mathbb{R}^n\) and \(z \in \mathbb{R}^m\) minimizing the goal function \(d(x, z) = \|z\|_2^2/2\), with the equality constraints \(Ax - b - z = 0\) and the inequality constraints \(x \geq 0\). The Lagrangian for this problem is

\[
L(x, z, \mu, \eta) = d(x, z) + \mu^T (Ax - b - z) - \eta^T x,
\]

where \(\mu \in \mathbb{R}^m\) and \(\eta \in \mathbb{R}^n\) are the dual variables. Here \(^T\) denotes the transpose. The Karush–Kuhn–Tucker (KKT) conditions give necessary (and in this case also sufficient) conditions for optimality (see [25], p 321 and p 448 ff). In this case the KKT conditions say that \((\hat{x}, \hat{z})\) is an optimal point if and only if there are \((\hat{\mu}, \hat{\eta})\) such that the following holds: (i) the stationarity conditions \(\nabla_x L = A^T \hat{\mu} - \hat{\eta} = 0\) and \(\nabla_z L = \hat{z} - \hat{\mu} = 0\), (ii) the primal constraints \(A\hat{x} - b - \hat{z} = 0, \hat{x} \geq 0\), (iii) the dual constraint \(\hat{\eta} \geq 0\), and (iv) the complementary slackness condition \(\hat{\eta}_j \hat{x}_j = 0\) for \(j = 1, \ldots, n\). The dual constraint on \(\eta\) occurs because it corresponds to a primal inequality constraint; there is no constraint on the dual variable \(\mu\), because of the corresponding primal equality constraint. The KKT system can be solved by a linear program \(\min \sum_i u_i + \sum_j v_j\) with constraints \(A^T \mu - \eta = u, Ax - b - \mu = v, x \geq 0, \eta \geq 0, u \geq 0, v \geq 0\), where \(x, \eta, u \in \mathbb{R}^n, z, \mu, v \in \mathbb{R}^m\), using a modification of the simplex method, where the remaining complementary slackness condition is enforced by a restricted basis entry rule in the simplex algorithm, see [26]. The Lagrangian dual objective function \(q(\mu, \eta) = \inf_{x,z} L(x, z, \mu, \eta)\) is finite for \((\mu, \eta)\) such that \(\mu^T A - \eta = 0\) (otherwise it has the value \(-\infty\)). For such \((\mu, \eta)\) the minimum occurs for \(z = \mu\) and any \(x \in \mathbb{R}^n\), so \(q(\mu, \eta) = L(x, \mu, \mu, \eta) = -\|\mu\|_2^2/2\). The Lagrangian dual problem is to maximize \(q(\mu, \eta) = -\|\mu\|_2^2/2\) with the constraints \(\mu^T A - \eta = 0, \eta \geq 0\). Since \(d(x, z)\) is convex we have strong duality, i.e., \(\max q(\mu, \eta) = q(\hat{\mu}, \hat{\eta}) = d(\hat{x}, \hat{z}) = \min d(x, z)\). By the KKT conditions derived above we have the optimal value \(-\|\hat{\mu}\|_2^2/2 = \|\hat{\mu}\|_2^2/2\). There are now two mutually exclusive cases: (i) the optimal value is 0, \(\hat{\mu} = \hat{z} = 0\) and \(Ax = b\) has a solution \(x \geq 0\), and (ii) the optimal value is \(> 0\). \(\hat{\mu} = \hat{z}, \|\hat{\mu}\|_2^2 = -z^T b > 0\) and \(Ax = b\) does not have a solution \(x \geq 0\). Considering the directional derivative of \(q(\mu, \eta)\) in the feasible direction \(\nu(v^T A \geq 0)\) we see that

\[
\frac{d}{d\mu} q(\mu, \eta) \bigg|_{\mu=0} = -\nu^T b,
\]

so \(\mu = 0\) is optimal in the dual problem (i.e., \(Ax = b\) has a solution \(x \geq 0\)) if and only if \(-\nu^T b \leq 0\) for all feasible \(\nu(v^T A \geq 0)\). This is the content of the famous Farkas’ lemma, see e.g. [27], p 56.
The simplex method and Farkas’ lemma have an instructive geometrical interpretation: the column vectors of $A$ generates a polyhedral cone $\{Ax : x \in \mathbb{R}^m\}$. If the optimal value is 0 then $b$ belongs to the cone $\kappa_A$ and we can find $x$ such that $Ax = b$, while if the optimal value is $>0$ then $b$ lies outside the cone and the optimal solution $\hat{x}$ is the point on the boundary of the cone minimizing the ‘distance’ $d$ between $Ax$ and $b$, see figure 1.

When the optimal value is zero it means that the measurements $\bar{c}$ can be realized exactly by a linear combination of the given sources, and hence the cone represents all possible measurements obtainable by linearly combining the given base sources.

A linear combination of Dirac measures is particularly interesting since these are extremal elements in the convex sense, and if $S = \delta_{r^*} \cdot \delta^*$, then $\langle S, c^* \rangle = c^*(r^*, x^*)$, so the measurement values obtained from a linear combination of Dirac measures consist of the polyhedral cone generated by the values of $c^*$ at the support points of the Dirac measures. This is generalized to arbitrary positive measures below, see theorem 12.

6. Measurements of arbitrary sources

The purpose of this section is to characterize all possible measurement values obtainable when $S$ is picked from a more general closed cone of positive measures. For this purpose, we define the measurement operator with respect to the given adjoint function $c^*$:

**Definition 2.** Given (i) a set $S$ of positive measures and (ii) a finite subset of non-negative continuous functions $c^* = (c_1^*, \ldots, c_m^*)$ on $T \times V$, we define

$$H_{c^*}(S) = \left\{ \left\langle S, c_1^* \right\rangle, \ldots, \left\langle S, c_m^* \right\rangle \right\} \in \mathbb{R}_+^m$$

for all $S \in \mathcal{S}$.

The results in the previous section shows that if $S$ is a finite positive cone (generated by the given sources $S_1, \ldots, S_n$) then the image $H_{c^*}(S)$ is a polyhedral cone in $\mathbb{R}_+^m$. In this section

![Figure 1. The standard simplex in $\mathbb{R}^3$ and its intersection with a cone in the positive octant.](image)
we drop the assumption on having a finite number of base sources and investigate whether we still can draw similar conclusions about the measurement values (i.e., the image of $H_\ast$).

We need to impose some structure (restrictions) on the set of source measures to perform the analysis, in particular we will make use of the notions of tightness and compactness.

**Definition 3.** A set of positive measures $\mathcal{S}$ on $T \times V$ is said to be **uniformly tight** if for each $\varepsilon > 0$ there is a compact set $K_\varepsilon \subset T \times V$ such that $S(K_\varepsilon^c) < \varepsilon$ for all $S \in \mathcal{S}$, where the set $K_\varepsilon^c$ denotes the complement of $K_\varepsilon$ in $T \times V$.

Loosely speaking the definition says that the mass\(^4\) contained in the complement of the compact set $K_\varepsilon$ can be made arbitrarily small, that is, nearly all mass is contained in the compact set $K_\varepsilon$, which intuitively means that the conceivable sources are not allowed to release ‘too much mass too far away and too long ago’. Measures can be constructed with approximation methods, and to show that approximations converges to the sought solution, we need appropriate compactness properties, in this case the following:

**Definition 4.** A set of positive measures $\mathcal{S}$ on $T \times V$ is said to be **weakly relatively compact** if for any sequence of measures $(S_j)_{j=1}^{\infty}$ in $\mathcal{S}$ there is a subsequence $(j_k \to \infty)$ when $k \to \infty$ such that $S_{j_k}$ is weakly convergent when $k \to \infty$, i.e., there is a measure $S$ (not necessarily in $\mathcal{S}$, unless $\mathcal{S}$ is weakly closed) such that $\int f \, dS_{j_k} \to \int f \, dS$ when $k \to \infty$, for all bounded continuous functions $f$.

The notion of compactness and tightness are related:

**Theorem 5.** (Prokhorov’s theorem) A set of positive measures $\mathcal{S}$ is weakly relatively compact if and only if $\mathcal{S}$ is uniformly tight and all $S \in \mathcal{S}$ have uniformly bounded total masses.

**Proof.** See [28], p 394–6. □

**Example 6.** The tightness condition is necessary. Consider a sequence of Dirac measures $S_j$ at discrete space–time points $(t_j, x_j)$ diverging to infinity. Then the $S_j$’s are not tight since any compact set is eventually avoided by $(t_j, x_j)$, but all $S_j$’s have mass 1 and hence uniformly bounded. There can be no weakly convergent subsequence $S_{j_k}$, since that would mean that $f(t_{j_k}, x_{j_k})$ is convergent for all bounded continuous functions $f$.

We are now in the position to show that if we consider source measures in the weak closure of a set of measures that are uniformly tight and has uniformly bounded total masses then the attainable measurement values $H_\ast(\mathcal{S})$ constitutes a closed and bounded set.

**Lemma 7.** If $\mathcal{S}$ is uniformly tight and has uniformly bounded total masses, then $H_\ast(\mathcal{S})$ is a compact subset of $\mathbb{R}_m^n$, where $\overline{\mathcal{S}}$ denotes closure of $\mathcal{S}$ with respect to weak convergence of measures.

\(^4\) Here the word mass refers to the value of the measure on the set, not physical mass.
Proof. Assume that $y_j \in H_{c^*}(S)$, i.e., there are measures $S_j$ such that $y_j = H_{c^*}(S_j)$, and assume that $y_j \to y$ when $j \to \infty$. By Prokhorov’s theorem, there is a subsequence $j_k \to \infty$ when $k \to \infty$ and a measure $S \in \mathcal{S}$ such that $S_{j_k} \to S$ weakly when $k \to \infty$, which implies that $H_{c^*}(S_{j_k}) \to H_{c^*}(S)$ when $k \to \infty$. Hence $H_{c^*}(S) = y$, so $H_{c^*}(\mathcal{S})$ is closed. Moreover, $H_{c^*}(\mathcal{S})$ is bounded since $c^*$ is bounded and $\mathcal{S}$ has uniformly bounded total masses. □

Example 8. (Single instantaneous point sources) Let $D \subset T \times V$ be an open subset of the space–time domain, and let $\mathcal{S}$ be the set of instantaneous point sources in $D$ with mass $M > 0$. Then $\mathcal{S}$ is the set of instantaneous point sources in $\overline{D}$ (the closure of $D$ in $T \times V$) with mass $M$, and $H_{c^*}(\mathcal{S}) = \left\{ Me^\ast(t, x); (t, x) \in \overline{D}\right\}$. Hence the attainable measurement values for $\mathcal{S}$ is a surface in $\mathbb{R}^m$ parametrized over the four-dimensional domain $\overline{D}$.

In lemma 7 tightness and uniformly bounded total masses implies compactness of $H_{c^*}(\mathcal{S})$, in order to sharpen this statement by replacing the implication by equivalence we introduce a particular kind of tightness adapted to $c^*$, indeed we consider compact sets constructed from level sets of $c^*$.

Definition 9. A set of positive measures $\mathcal{S}$ is said to be uniformly $c^*$-tight if for every $\varepsilon > 0$ there are $\varepsilon_1, \ldots, \varepsilon_m > 0$ and a compact set $K_\varepsilon \subset T \times V$ such that $K_\varepsilon \subset \bigcup_j \left\{ c_j^* \geq \varepsilon_j \right\}$ and $S(K_\varepsilon^c) < \varepsilon$, where the set $K_\varepsilon^c$ is the complement of $K_\varepsilon$ in $T \times V$.

Example 10. If $\mathcal{S}$ consists of measures supported on $\bigcup_j \left\{ c_j^* \geq \varepsilon \right\}$ for some $\varepsilon > 0$, then $\mathcal{S}$ is uniformly $c^*$-tight.

By imposing the stronger assumption (yet natural for the problem we are studying) of $c^*$-tightness we sharpen the result in lemma 7 by having implication in both directions.

Theorem 11. Assume that $\mathcal{S}$ is uniformly $c^*$-tight. Then $\mathcal{S}$ has uniformly bounded total masses if and only if $H_{c^*}(\mathcal{S})$ is a compact subset of $\mathbb{R}^m$.

Proof. Clearly, if $\mathcal{S}$ has uniformly bounded total masses then $H_{c^*}(\mathcal{S})$ is bounded, since $c^*$ is bounded and continuous. If $H_{c^*}(\mathcal{S}) = c$ (componentwise), then take $\varepsilon, \varepsilon_1, \ldots, \varepsilon_m > 0$ and $K_\varepsilon$ such that $K_\varepsilon \subset \bigcup_j \left\{ c_j^* \geq \varepsilon_j \right\}$ and $S(K_\varepsilon^c) < \varepsilon$ for all $S \in \mathcal{S}$. Then for all $S \in \mathcal{S}$ we have $\sum_j \varepsilon_j S(K_\varepsilon) \leq \sum_j \varepsilon_j S \left( c_j^* \geq \varepsilon_j \right) \leq \sum_j \int c_j^* dS \leq \sum_j \varepsilon_j$ so the total mass of $S$ is $S(K_\varepsilon^c) + S(K_\varepsilon) \leq \varepsilon + \sum_j \varepsilon_j / \sum_j \varepsilon_j$. □
Theorem 12. Assume that \( K \subset T \times V \) is compact and all \( c_j^+ \geq \varepsilon \) on \( K \) for some \( \varepsilon > 0 \), and let \( \mathcal{S} \) the set of all positive finite measures on \( K \). Then \( H^c(\mathcal{S}) \) is the closure of the convex conical hull of \( c^+(K) \).

Proof. \( \mathcal{S} \) is a weakly closed set since \( K \) is compact. Moreover, every \( S \in \mathcal{S} \) is the weak limit of a sequence of discrete \( S_j \) supported in \( K \), i.e., \( S_j = \sum_{k=1}^{N_j} c_{jk} \delta_{jk} \) where \( \delta_{jk} \) are Dirac measures supported at suitable space–time points \( (t_{jk}, x_{jk}) \in K \), and \( c_{jk} > 0 \) and \( \sum_{k=1}^{N_j} c_{jk} = \int dS \) for \( k = 1, \ldots, N_j \) and \( j = 1, 2, \ldots \). Also, \( H^c(S_j) = \langle S_j, c^+ \rangle = \sum_{k=1}^{N_j} c_{jk} c_j^+ (t_{jk}, x_{jk}) \), so \( H^c(\mathcal{S}) \) is in the conical hull of \( c^+(K) \), and \( H^c(S_j) \to H^c(S) \) when \( j \to \infty \). This proves that \( H^c(\mathcal{S}) \) is included in the closure of the convex conical hull. Conversely, given a point \( y \) in the closure of the conical hull, there is a sequence \( S_j \) of discrete measures of the above form such that \( H^c(S_j) \to y \). Since all \( c_j^+ \geq \varepsilon \) on \( K \), the masses of the \( S_j \)'s must be uniformly bounded, and since they are supported on the compact set \( K \), they form a tight set of measures. By Prokhorov's theorem there is a subsequence \( j_k \to \infty \) when \( k \to \infty \) and a measure \( \mu \in \mathcal{S} \) such that \( \mu \to S \) weakly, and hence \( H^c(S) \to H^c(S) \) when \( k \to \infty \). Hence \( y = H^c(S) \), so \( y \in H^c(\mathcal{S}) \), which proves that the closure of the convex conical hull is included in \( H^c(\mathcal{S}) \).

7. Cones of measures

In this section, as a preamble to the next section on the least squares solution, we give a technical lemma on the closedness of cones generated by closed bounded sets of measures.

To reach the desired result we have to introduce an additional condition on the generating set, namely a lower bound on the mass of \( S \).

Definition 13. A set \( \mathcal{S} \) of positive measures is said to have uniformly positive total masses if there is a constant \( M > 0 \) such that the total mass of \( S \) is \( \geq M \) for all \( S \in \mathcal{S} \).

Lemma 14. Assume that \( \mathcal{S} \) is a set of positive measures on \( T \times V \), and let \( C = \text{cone}(\mathcal{S}) \), the positive cone generated by \( \mathcal{S} \). Then \( \text{cone}(\mathcal{S}) \subseteq C \). Moreover, if \( \mathcal{S} \) have uniformly positive total masses, then \( \text{cone}(\mathcal{S}) = C \).

Proof. The first statement follows from the fact that if \( S_j \in \mathcal{S} \) and \( S_j \to S \) weakly, then \( \lambda S_j \to \lambda S \) for all \( \lambda \geq 0 \). To prove the second statement, assume that \( \mu \in C \), and take \( \lambda_j S_j \in C \) with \( \lambda_j \geq 0 \), \( S_j \in \mathcal{S} \) and \( \lambda_j S_j \to \mu \) weakly. Since the \( S_j \)'s have uniformly bounded masses from below, the \( \lambda_j \)'s are uniformly bounded, and hence there is a subsequence \( j_k \to \infty \) such that \( \lambda_{j_k} \to \lambda \) when \( k \to \infty \). Hence \( S_j \to \mu/\lambda \) weakly, so \( \mu/\lambda \in \mathcal{S} \), i.e., \( \mu \in \text{cone}(\mathcal{S}) \).

The following example shows that the lower bound on the masses in \( \mathcal{S} \) is necessary for the second statement.
Example 15. Let $\delta = \in \mathcal{S} \times (0, 1 \setminus 0 \subseteq \mathcal{S}. Then \lambda \delta \in \mathcal{S}$ for some $\lambda > 0$, so there is a sequence $x_j \downarrow 0$ such that $x_j \delta \to \lambda \delta_0$ weakly. Hence $x_j f(x_j) \to \lambda f(0)$ for all continuous functions, which is a contradiction since we can have $f(0) \neq 0$. We conclude that $\delta_0 \not\in \mathcal{S}$.

8. Least squares solutions to inverse problems

In addition to characterising the set of measurements, lemma 14 enables us to determine when the least squares inverse problem is well-defined (theorem 18 below). We begin by defining the least squares solution to the inverse problem.

Definition 16. Let (i) a set of adjoint plumes $c^b = (c_1^*, \ldots, c_m^*)$ on $\mathcal{T} \times \mathcal{V}$ be continuous and bounded, (ii) $c = (c_1, \ldots, c_m)$, where $c_i \geq 0$ for $i = 1, \ldots, m$, be a set of measurement values and (iii) $C$ a weakly closed cone of positive measures on $\mathcal{T} \times \mathcal{V}$, then a least square solution to the inverse problem in $C$ is a measure $\tilde{\mu} \in C$ such that

$$\| \tilde{\mu} - H^*(\mathcal{S}) \| = \min_{\mu \in C} \| \tilde{\mu} - H^*(\mu) \|,$$

where $\| \cdot \|$ denotes the Euclidean norm in $\mathbb{R}^m$.

Collecting the results from the previous sections we are now in a position to show when the least squares inverse problem is well defined. Let $\mathcal{S}$ be a set of positive measures on $\mathcal{T} \times \mathcal{V}$. We assume that $\mathcal{S}$ has uniformly positive total masses and let $\mathcal{S} = \text{cone}(\mathcal{S})$, then by lemma 14 it follows that $\mathcal{S}$ is weakly closed. We furthermore assume that $\mathcal{S}$ is tight and has uniformly bounded total masses which by lemma 7 implies that the image of the cone $\kappa = H^*(\mathcal{S})$ is a closed positive cone in $\mathbb{R}^m$. We now prefer to conduct the analysis of the least squares inverse problem on the generating set alone and we therefore need the following lemma justifying that it suffices to solve a minimization problem on the generating set.

For a single ray, we have an analytical formula for the closest point, namely

$$\pi_c(z) = (x \cdot z)z / || z ||^2,$$

the closest point from $x$ on the ray $\{z: \lambda > 0\}$. Therefore we can minimize over a generating set rather than over the full cone:

Lemma 17. Assume that $\kappa$ is a positive cone in $\mathbb{R}_+^m$ generated by a set $B \subseteq \mathbb{R}_+^m \setminus \{0\}$, and assume that $x \in \mathbb{R}_+^m \setminus \kappa, x \neq 0$. If

$$y \in \kappa \text{ and } \| y - x \| = \inf_{w \in \kappa} \| w - x \|$$

then there is a $z \in B$ such that $y = \pi_c(z)$, and

$$\| \pi_c(z) - x \| = \inf_{w \in B} \| \pi_c(w) - x \|.$$

Conversely, if $z \in B$ satisfies (13) then $y = \pi_c(z) = \| \pi_c(z) \| z / || z ||$ satisfies (12).
Proof. Minimizing over rays we have
\[
\inf_{w \in \mathcal{B}} \| w - x \| = \inf_{w \in \mathcal{B}} \| \pi_c(w) - x \|.
\]
Moreover, for any \( y \in \kappa \) there are \( z \in \mathcal{B} \) and \( \lambda > 0 \) such that \( y = \lambda z \), and \( \pi_c(y) = \pi_c(z) \). Consequently, for such \( y \) and \( z \), if either \( \| y - x \| = \inf_{w \in \mathcal{B}} \| w - x \| \) or \( \| \pi_c(z) - x \| = \inf_{w \in \mathcal{B}} \| \pi_c(w) - x \| \) holds we have
\[
\| y - x \| = \inf_{w \in \mathcal{B}} \| w - x \| = \inf_{w \in \mathcal{B}} \| \pi_c(w) - x \| = \inf_{w \in \mathcal{B}} \| \pi_c(w) - x \| = \| \pi_c(z) - x \|.
\]
In view of the previous lemma we see why we insisted on introducing the assumption on uniform positive total masses: it is important that the generating set does not contain the origin. Now, finally, we have come to the point where we can state, and easily prove, the main theorem:

**Theorem 18.** Assume that the set of measures \( \mathcal{S} \) is uniformly tight, and weakly closed, with uniformly bounded and uniformly positive total masses. Let \( \mathcal{C} \) be the positive cone generated by \( \mathcal{S} \). Then \( \mathcal{S} \) is weakly compact, and \( \mathcal{C} \) is weakly closed. Moreover, there is a solution \( \hat{S} \) to the least squares inverse problem (10) on \( \mathcal{C} \), given by
\[
\hat{S} = \pi_c \left( H_{\kappa} \left( \hat{S} \right) \right) \frac{H_{\kappa} \left( \hat{S} \right)}{\| H_{\kappa} \left( \hat{S} \right) \|},
\]
where \( \hat{S} \) is a solution to the following least squares problem on \( \mathcal{S} \):
\[
\| \pi_c \left( H_{\kappa} \left( \hat{S} \right) \right) - \hat{c} \| = \min_{\hat{S} \in \mathcal{S}} \| \pi_c \left( H_{\kappa} \left( \hat{S} \right) \right) - \hat{c} \|,
\]
where the ray projector \( \pi_c \) is given in (11).

Proof. The set \( \mathcal{S} \) is weakly relatively compact by theorem 5 and hence weakly compact since it is assumed to be weakly closed. The cone \( \mathcal{C} \) generated by \( \mathcal{S} \) is weakly closed by lemma 14. The set \( \mathcal{B} = H_{\kappa}(\mathcal{S}) \) is compact by lemma 7, and \( \mathcal{B} \subseteq \mathbb{R}_+^m \) since \( \mathcal{S} \) has uniformly positive total masses. Let \( \kappa = H_{\kappa}(\mathcal{C}) \). Then \( \kappa \) is the positive cone generated by \( \mathcal{B} \), and \( \kappa \) is closed because the mapping \( H_{\kappa} \) is continuous. Since \( \mathcal{B} \) is compact, there is a \( z \in \mathcal{B} \) such that \( \| \pi_c(z) - \hat{c} \| = \min_{\hat{S} \in \mathcal{S}} \| \pi_c(w) - \hat{c} \| \). By the second statement in lemma 17, \( y = \| z \| z \) satisfies \( \| y - \hat{c} \| = \inf_{w \in \kappa} \| w - \hat{c} \| \) and \( y \in \kappa \) since \( \kappa \) is closed. Finally, we take \( \hat{S} \in \mathcal{S} \) such that \( H_{\kappa} \left( \hat{S} \right) = z \); then \( \hat{S} \) satisfies (15) and \( \hat{S} \) given by (14) has \( H_{\kappa} \left( \hat{S} \right) = y \) and \( \hat{S} \) is a solution to (10).

Note that the solution is not necessarily unique, unless \( \mathcal{S} \) is a convex set of positive measures, in which case \( \mathcal{C} \) is a closed convex cone of positive measures and \( \kappa = H_{\kappa}(\mathcal{C}) \) is a closed convex cone in \( \mathbb{R}_+^m \). Note also that it suffices to find a minimizer in the generating set \( \mathcal{S} \), and compute the scaling afterwards.

**Example 19.** Let \( \mathcal{S} \) be the set of single instantaneous point sources in a compact set \( K \subseteq T \times V \). This is a uniformly tight, weakly closed set of measures with uniformly bounded and uniformly positive total masses, representing instantaneous point sources of unit mass.
The positive cone $\mathcal{C}$ generated by $\mathcal{S} = \mathcal{S}$ represents all instantaneous point sources supported in $K$. Hence $B = H_\kappa(\mathcal{S}) = \mathcal{C}(K)$, the image of $K$, is a basic set for the closed cone $\kappa = H_\kappa(\mathcal{C})$. Note that neither of the cones are convex; only single instantaneous point sources, not linear combinations of different ones, are included.

**Example 20.** Let $\mathcal{S}$ be the set of single continuous point sources with spatial support in a compact set $K \subset V$ and unit total mass, i.e.

$$S = q(t) \, dt \otimes \delta_\ast(x)$$

where $q$ is a non-negative continuous function with $\int_T q(t) \, dt = 1$, and $x^\ast \in K$. Then the weak closure $\overline{\mathcal{S}}$ of $\mathcal{S}$ consists of all

$$S = \mu(\, dt) \otimes \delta_\ast(x)$$

where $\mu$ is a probability measure on $T$. Note that $\overline{\mathcal{S}}$ includes temporally singular measures, for example discrete sums of instantaneous point sources $S = \sum_k \lambda_k \delta_\ast(x)$ with $\sum_k \lambda_k = 1$. This kind of singular measures must be included in order to obtain a closed cone $H_\kappa(\mathcal{C})$, and thereby a well-posed minimization problem.

**9. Conclusion**

We have presented a measure theoretic framework for studying the adjoint dispersion problem. This framework and the accompanying measure theoretic machinery enabled us to derive results for general linear inverse dispersion problems without making prior assumptions on the number of sources, their emission patterns and so on. Indeed, in our modus operandi the notion of number of sources is not even a well-defined concept. We investigated when a given set of sensor data can be realizable from a linear combination of source measures chosen from some subset of all positive measures. Then we shifted the view from working with a fixed set of measurement values, to asking (and answering) the question: if the source is chosen from a closed cone of positive measures, what are the possible measurement values that this source can produce? Finally we used the framework to derive necessary and sufficient conditions for the existence of a solution to the inverse least-squares problem.

We conclude that the framework presented in this paper is a powerful tool for stating and proving results on linear inverse atmospheric problems in their full generality. The framework is not limited to proving the results that we have presented here, indeed our next step is to use the framework to prove rigorous results on the first order inverse method of footprints, e.g. [5, 6]. The framework is also easily augmented to incorporate the forward dispersion problem as well. Our preliminary investigations into uncertainty analysis of the forward dispersion problem indicates that this is a fruitful approach.

**References**

[1] Gudiksen P H, Harvey T F and Lange R 1989 Chernobyl source term, atmospheric dispersion, and dose estimation Health Phys. 57 697–706

[2] Stohl A, Selbert P, Wotawa G, Arnold D, Burkhart J F, Eckhardt S, Tapia C, Vargas A and Yasunari T J 2012 Xenon-133 and Caesium-137 releases into the atmosphere from the Fukushima Dai-ichi nuclear power plant: determination of the source term, atmospheric dispersion, and deposition Atmos. Chem. Phys. 12 2313–43
[3] Ringbom A et al 2014 Radioxenon detections in the VTBT international monitoring system likely related to the announced nuclear test in North Korea on 12 February 2013 J. Environ. Radioact. 128 47–63
[4] Schmid H P 2002 Footprint modeling for vegetation atmosphere exchange studies: a review and perspective Agric. Forest Meteorol. 113 159–83
[5] Pudykiewicz J A 1998 Application of adjoint tracer transport equations for evaluating source parameters Atmos. Environ. 32 3039–50
[6] Robertson L 2004 Extended Back-Trajectories by Means of Adjoint Equations (Norrköping, Sweden: Swedish Meteorological and Hydrological Institute) pp 109–33 RMK No.105
[7] Yee E and Flesch T K 2010 Inference of emission rates from multiple sources using Bayesian probability theory J. Environ. Monit. 12 622–34
[8] Stuart A M 2010 Inverse problems: a Bayesian perspective Acta Numer. 19 451–559
[9] Franklin J N 1970 Well-posed stochastic extensions of ill-posed linear problems J. Math. Anal. Appl. 31 682–716
[10] Keats A, Yee E and Lien F-S 2007 Bayesian inference for source determination with applications to a complex urban environment Atmos. Environ. 41 1128–34
[11] Yee E 2007 Bayesian inversion of concentration data for an unknown number of contaminant sources Technical Report DRDC Suffield TR 2007-085
[12] Yee E 2012 Probability theory as logic: data assimilation for multiple source reconstruction Pure Appl. Geophys. 169 499–517
[13] Green P 1995 Reversible jump Markov chain Monte Carlo computation and Bayesian model determination Biometrika 82 711–32
[14] Yee E 2012 Inverse dispersion for an unknown number of sources: model selection and uncertainty analysis ISRN Appl. Math. 2012 465320
[15] Robertson L and Langner J 1998 Source function estimate by means of variational data assimilation applied to the ETEX-I tracer experiment Atmos. Environ. 32 4219–25
[16] Thomson L C, Hirst B, Gibson G, Gillespie S, Jonathan P, Skeldon K D and Padget M J 2007 An improved algorithm for locating a gas source using inverse methods Atmos. Environ. 41 1128–34
[17] Allen C T, Young G S and Haupt S E 2007 Improving pollutant source characterization by better estimating wind direction with a genetic algorithm Atmos. Environ. 41 2283–9
[18] Issartel J P, Sharan M and Singh S K 2012 Indentification of a point source by use of optimal weighted least squares Pure Appl. Geophys. 169 467–82
[19] Sharan M, Singh S K and Issartel J P 2012 Least square data assimilation for identification of the point source emissions Pure Appl. Geophys. 169 483–97
[20] Butler T, Estep D, Tavener S, Dawson C and Westerink J 2014 A measure-theoretic computational method for inverse sensitivity problems: III. Multiple quantities of interest J. Uncertain. Quantification 2 174–202
[21] Bocquet M 2005 Reconstruction of an atmospheric tracer source using the principle of maximum entropy I: theory Q. J. R. Meteorol. Soc. 131 2191–208
[22] Rudin W 1966 Real and Complex Analysis (New York: McGraw-Hill)
[23] Marchuk G I 1986 Mathematical models in environmental problems Stud. Math. Appl. 16 40–1
[24] Blaschke W and Pick G 1916 Distanzschätzungen im Funktionenraum: II Math. Ann. 77 277–300
[25] Nocedal J and Wright S J 2006 Numerical Optimization 2nd edn (Berlin: Springer)
[26] Wolfe P 1959 The simplex method for quadratic programming Econometrica 27 383–98
[27] Franklin J N 1980 Methods of Mathematical Economics (Berlin: Springer)
[28] Daley D J and Vere-Jones D 2003 An Introduction to the Theory of Point Processes vol 1 2nd edn (Berlin: Springer)