A minimum curvature algorithm for tomographic reconstruction of atmospheric chemicals based on optical remote sensing

Sheng Li, Ke Du

Department of Mechanical and Manufacturing Engineering, University of Calgary, Calgary, AB T2N 1N4, Canada

Correspondence to: Ke Du (kddu@ucalgary.ca)

Abstract. Optical remote sensing (ORS) combined with the computerized tomography (CT) technique is a powerful tool to retrieve ~~a~~ two-dimensional concentration map over ~~the~~ an area under investigation. ~~But unlike the~~Whereas medical CT usually uses ~~, the~~ beam number ~~used in ORS CT is usually dozens comparing to up to~~ of hundreds of thousands ~~in the former~~, ORS-CT usually uses a beam number of dozens, ~~which~~ thus severely ~~limit~~limitings the spatial resolution and the quality of the reconstructed map. T~~his situation makes t~~he 'smoothness' *a priori* information is therefore crucial ~~especially necessary~~ for ORS-CT. Algorithms that~~which~~ produce smooth reconstructions include smooth basis function minimization ~~(SBFM)~~, grid translation and multiple grid (GT-MG), and low third derivative (LTD), among which the LTD algorithm is ~~a~~ promising ~~one with~~because of fast speed ~~and simple realization~~. ~~But~~ However, its ~~characteristics and the~~ theoretical~~y~~ basis ~~are not clear~~must be clarified to better understand the characteristics of its smoothness constraints. Moreover, the computational efficiency and ~~the~~ reconstruction quality need to be improved for practical applications. This paper first treated the LTD algorithm as a special case of Tikhonov regularization that uses the approximation of the third-order derivative as the regularization factor. Then, to seek more flexible smoothness constraints, we successfully incorporated the smoothness seminorm used in variational interpolation theory into the reconstruction problem. Thus, the smoothing effects can be well understood according to the close relationship between the variational approach and the spline functions. Furthermore, ~~employs two theories, i.e., Tikhonov regularization and spatial interpolation, to produce a smooth reconstruction by ORS CT. Within the two theories' frameworks, new~~other algorithms can be ~~explored in order to improve the performance~~formulated by using different seminorms. ~~For example~~On the basis of this idea, we propose a new minimum curvature (MC) algorithm by ~~based on the variational approach in the theory of the spatial interpolation~~using a seminorm approximating the sum of the squares of the curvature, which reduces the number of linear equations ~~by~~ to half ~~comparing to~~ that in the LTD algorithm ~~using the biharmonic equation instead of the smoothness seminorm~~. ~~We compared our~~ The MC algorithm was compared with the non-negative least square (NNLS), GT-MG, and LTD algorithms by using multiple test maps. ~~The MC and the LTD algorithms have~~The MC algorithm, compared with the LTD algorithm, shows similar performance as ~~on~~ in terms of ~~the~~ reconstruction quality~~. But the MC algorithm~~ but ~~needs~~ requires only ~~about~~ approximately 65% the computation time ~~of the LTD algorithm~~. It is ~~much~~ also simpler ~~in realization than~~ to perform than the GT-MG algorithm ~~by using~~because it directly uses high-resolution grids ~~directly~~ during the reconstruction process ~~to generate a high-resolution map immediately after one reconstruction process is done~~. Compar~~ed~~ing

~~to~~ with the traditional NNLS algorithm, it shows better performance in three aspects: (1) the nearness of reconstructed maps is improved by more than 50%; (2) the peak location accuracy is improved by 1- 2 m; and (3) the exposure error is improved by more than ~~10~~ ten times. T~~he t~~esting results ~~show~~ indicated the effectiveness of the new algorithm ~~based on~~according to the ~~spatial~~ variational ~~interpolation theory~~approach. More specific algorithms could be similarly further formulated and evaluated. ~~Similarly, other algorithms may also be formulated to address problems such as the over-smooth issue in order to further improve the reconstruction equality.~~ Thise study~~ies~~ ~~will~~ promotes the practical application of ~~the~~ ORS-CT mapping of atmospheric chemicals.

1 Introduction

Measuring the concentration distribution of atmospheric chemicals over ~~a~~ large areas is required in many environmental applications, ~~including~~ such as locating hot-spots or emission sources of air pollutants (Wu et al., 1999), understanding air pollutant dispersion and airflow patterns, and quantifying emission rates or ventilation efficiency (Samanta and Todd, 2000; Belotti et al., 2003; Arghand et al., 2015). The traditional network method~~s~~ uses multiple point samplers placed at various locations in the region under investigation~~, which~~. This method is ~~is~~ intrusive, time-consuming, and limited in temporal and spatial resolution (Cehlin, 2019). The advanced method is based on the combination of optical remote sensing ~~(ORS)~~ and computerized tomography ~~(CT)~~ techniques (ORS-CT). ORS-CT is ~~is~~ a powerful technique for sensitive mapping of air contaminants ~~measurement, which can detect a large~~throughout kilometer-size areas in ~~situ and provide near~~ real-time ~~information~~ (Du et al., 2011). ~~The path-integrated concentration (PIC) is measured along each path using techniques like~~Two commonly used ORS techniques use an open-path tunable diode laser (~~OP~~-TDL) ~~or~~ and open-path Fourier transform infrared spectrometer ~~(OP-FTIR)~~. The ORS analyzer emits a light beam targeted at multiple mirrors, which reflect the beam back to the analyzer. For each beam path, the path-integrated concentration (PIC) is obtained. After multiple PICs are collected, ~~the~~ a two-dimensional concentration map can be ~~reconstructed~~ generated ~~by~~ through ~~the CT technique~~tomographic reconstruction algorithms (Hashmonay et al., 2001). The ORS-CT method provides better spatial and temporal resolution than the network approach, and it is more sensitive than the range-resolved optical techniques. ~~Comparing to other techniques, ORS-CT~~It is also non-intrusive and ~~fast. It also has good spatial resolution and can works~~suitable for continuous~~ly~~ long-term ~~and automatically~~monitoring. ~~These advantages make ORS-CT have the potential to be an excellent tool for investigating air dispersion problems under various conditions.~~

In ORS-CT mapping of atmospheric chemicals, owing to factors including system cost, response time, beam configuration, ~~In environmental applications,~~ the number of beams ~~in ORS-CT~~ is ~~usually less than 50~~only tens-, whereas the number of beams in medical CT is hundreds of thousands~~comparing to hundreds of thousands of beams used in the medical CT due to factors including cost, response, configuration, and other practical considerations.~~ The very small beam number poses several

challenges in tomographic reconstruction algorithms. ~~The reconstruction techniques have been well studied in medical CT. Theoretically, the true distribution can be determined exactly by Radon transformation approach, which requires infinite beams (Radon, 1986).~~ In practice, ~~the~~ transform~~ation~~ methods~~approaches~~ based on the theory of Radon transformation using a ~~include back projection and~~ filtered back projection ~~(FBP)~~ formula ~~methods~~are not feasible because of noise and artifacts in the reconstructions (Radon, 1986; Herman, 2009). ~~Series-expansion-based methods are another type of approaches which estimate~~ Series expansion methods, which discretize the reconstruction problem before any mathematical analysis, are usually used in ORS-CT. The underlying distribution is represented by a linear combination of a finite set of basis functions (Censor, 1983). ~~the true distribution by finding a finite set of numbers or superposition of a set of simple functions. In environmental applications, the number of beams in ORS-CT is usually less than 50 comparing to hundreds of thousands of beams used in the medical CT due to factors including cost, response, configuration, and other practical considerations. Therefore, only the series-expansion-based methods are applied in ORS-CT, which include~~The simplest type is the ~~two main types of approaches: pixel-based approach and basis function based approach. P~~pixel-based approach, which ~~-~~ divides an area into multiple grid pixels ~~(grids)~~ and ~~assigns a concentration value to each pixel~~assigns a unit value inside each pixel. The ~~PIC~~ path integral is approximated by ~~calculated by summarising~~the summation of the product of the pixel ~~concentration~~ value and the length of the path in ~~each~~ that pixel. A system of linear equations can be set up for multiple beams. The inverse ~~question~~ problem ~~is to~~involves find~~ing~~ the ~~best~~ optimal set of pixel concentrations~~. Basis function approach assumes that the true distribution is determined by the superposition of a set of simple basis functions with unknown parameters. Possible basis functions include bilinear function, bivariate Gaussian function, etc (Giuli et al., 1999; Hashmonay et al., 1999). A system of non-linear equations is set up for multiple beams. The question is to find the best set of parameters. For both approaches, the inverse questions are solved by minimizing the error function constructed based on~~according to ~~some~~ criteria ~~(Price et al., 2001) ,~~ including the least square criterion to minimize the summation of the squared errors between the observed and model-predicted PICs; the maximum likelihood (ML) criterion to maximize the probability of the PIC observations given the distribution of the random variables of the concentrations and observation errors; and the maximum entropy criterion to maximize the entropy of the reconstructed maps, given that the average concentration of the map is known~~including minimizing the L_2 norm of error (finding a least squares (LS) solution), maximum likelihood (ML), maximum entropy , etc~~(Herman, 2009). ~~In pixel-based reconstruction, the inverse problem is linear but usually ill-posed. The number of equations is very large.~~ Commonly used~~Numerical iterative techniques~~ pixel-based algorithms are ~~used to estimate the solution, including~~ algebraic reconstruction techniques (ART) ~~(ART)~~, non-negative least square (NNLS)~~steepest descent, conjugate gradient (CG)~~, and expectation-maximization (EM) (Tsui et al., 1991; Lawson and Janson, 1995; Todd and Ramachandran, 1994~~; Drescher et al., 1996~~). The NNLS algorithm has similar performance to the ART algorithm but shorter computation time (Hashmonay et al., 1999). It has been used in US EPA OTM-10 for horizontal radial plume mapping of air contaminants (EPA, 2005). The EM algorithm is mainly used for ML-based minimization. These traditional pixel-based algorithms are suitable for rapid CT, but they produce maps with poor spatial resolution, owing to the requirement that the pixel number must not exceed the beam number, or they

may have problem of indeterminacy associated with substantially underdetermined systems (Hashmonay, 2012). ~~In the basis-function-based approach, the inverse problem is usually non-linear. Stochastic optimization algorithms need to be used to find the global minimum (e.g., simulated annealing (SA)).~~

To mitigate the problem of indeterminacy and improve the spatial resolution of reconstructions without substantially increasing the system cost, the smooth basis function minimization (SBFM) algorithm has been proposed. This algorithm represents the distribution map by a linear combination of several bivariate Gaussian functions (Drescher et al., 1996; Giuli et al., 1999). Each bivariate Gaussian has six unknown parameters (normalizing coefficient, correlation coefficient, peak locations and standard deviations) to be determined. The problem requires fitting these parameters to the observed PIC data. ~~For ORS-CT mapping of atmospheric chemicals, l~~This method ~~iteratures show that the smooth basis function minimization (SBFM) algorithm performs better~~ performs better than ~~other~~the traditional pixel-based algorithms ~~(Drescher et al., 1996). This is b~~ for ORS-CT applications because the patterns of air dispersion are physically smooth in shape (Wu and Chang, 2011). ~~Therefore, an algorithm converging toward smooth concentration distribution consistent with the path-integrated data is a rational choice (Cehlin, 2019).~~ However, the resultant equations defined by the PICs are non-linear because of the unknown parameters. The search for the best-fit set of parameters minimizing the mean-squared difference between predicted and measured path integrals can be performed through an iterative minimization procedure, such as the simplex method or simulated annealing.

The reported methods using simulated annealing to find a global minimization are ~~SBFM is~~ highly computationally ~~-~~intensive ~~due to the non-linear optimization, which was reported to be about 100 times slower than pixel-based algorithms (Price et al., 2001). This makes it unfavorable~~, thereby limiting the SBFM algorithm's practical applications, such as ~~for rapid reconstruction, which is usually required~~ rapid reconstruction in industrial ~~applications such as~~ monitoring of chemical plants. However, an algorithm converging toward a smooth concentration distribution consistent with the path-integrated data has been demonstrated to be a rational choice. To improve the computational speed and append the smoothness *a priori* information to the inverse problem~~A representative pixel based algorithm is the non negative least square (NNLS), which is also the algorithm used in the USEPA OTM-10 for horizontal radial plume mapping (HRPM) measurements (EPA, 2005). It achieves similar results as the multiplicative algebraic reconstruction technique (MART), but with shorter computing time (Hashmonay et al., 1999). The traditional pixel based algorithms, although suitable for rapid CT, have poor performance comparing to SBFM algorithm (Wu and Chang, 2011). This is mainly due to the coarse grid resolution limited by the beam number. Because the algorithm requires the number of grids to be less than or equal to the number of beams (Hashmonay, 2012).~~

~~To improve the performance of traditional pixel based algorithms, there are two reported approaches used in ORS-CT. One method is t~~, the pixel-based "low third derivative" (LTD) algorithm has been proposed~~, which~~. This algorithm ~~adds~~ sets the third derivative at each pixel to zero~~as a smoothness restriction (Price et al., 2001).~~, thus resulting in a new system of linear equations that is overdetermined. The LTD algorithm ~~was~~ has been reported to work ~~nearly~~ as well as the SBFM algorithm, but ~~was~~ is ~~about~~ approximately 100 times faster (Price et al., 2001). Another method to produce the smoothness effect is the

'grid translation' (GT) algorithm, which shifts the basis grid by different distances (e.g., 1/3, or 2/3 the width of the basis grid) horizontally and vertically while keeping the basis grid fixed (Verkruysse and Todd, 2004). The performance is improvedSmoothness is achieved by averaging the reconstruction results after each shifting. There is also aAn improved version called 'grid-translation and multi-grid' (GT-MG), which applies the GT algorithm at different basis grid resolutions (Verkruysse and Todd, 2005). This method has been successfully used with the maximum likelihood expectation maximization (ML-EM) algorithm to improve the reconstruction accuracy, especially particularly in determining the peak location and value (Cehlin, 2019).

The successes of these algorithms show demonstrates the necessary need of to applying the smoothness restriction to the ORS-CT gas mapping. Using With the LTD algorithm, we can easily achieve a smooth reconstruction is achieved by simply adding the third-order derivative constraints. The generated solutions are locally quadratic. However, like the Gaussian model used in the SBFM algorithm, the smoothness in the LTD algorithm is also achieved based on ideal model (quadratic functions)To. It is necessary to understand the characteristics of this these model constraints andin order to apply the algorithm method to specific application, the . But the theoreticaly basis of the LTD algorithm must be understood. However, this basis is was not clearly givendefined in the literature. When we tried toWith the purpose of introducinge the smoothness as *a priori* informationconstraints to the reconstruction process, we found that the LTD algorithm can be treated as a special case of a smooth reconstruction could be achieved based on two well-established theories. The first one is thethe Tikhonov regularization, a well-known technique to theory originated from Tikhonov for solving solve the ill-posed inverse problem (Tikhonov and Arsenin, 1977; Rudin et al., 1992). The most well-known approach is the Tikhonov L_2 regularization , which uses a penalty term defined by the squared norm of the *ith*-order derivative of the function as a regularization factor and produces a smoothing effect on the resulting solution (Gholami and Hosseini, 2013). From The third-order derivative is used in the view of regularization theory, the LTD algorithm is a regularized algorithm which uses the third-order derivative. In fact, although the first, second, and higher order derivatives can also produce smooth results. A more flexible method of regularization uses the smoothness seminorm The second one is the interpolation theory based on which almost all algorithms introduce the smooth effect to the output. Specifically, theaccording to the variational interpolation approach theory, given its similar formula (Mitasova et al., 1995). The variational method is another way of achieving spline interpolation, given that the interpolation polynomial splines can be derived as the solution of certain variational problems of minimizing an integral whose integrant can produce a regularized inverse problem similar to Tikhonov regularization by using a smoothness seminorm. The seminorm consists of different order derivatives or their combinations. The solution to this inverse problem is a set of spline functions. For example, a bivariate smoothness seminorm with squares of second derivative leads to a thin plate spline (TPS) function. Based on the variational interpolation theory, the LTD algorithm is a special case of the regularized spline with tension (RST) (Mitasova et al., 1995).

The interpolation techniques are based on the given sample points, in contrast to tomographic reconstruction, in which only the line integrals are known. However, we have found that the interpolation can be adopted in the reconstruction process to produce a smooth solution by using the smoothness seminorm for interpolation as a smoothness regularization factor for the tomographic reconstruction problem. In ~~the~~ view of variational spline~~framework of the regularized inversion and spatial~~ interpolation~~ theories~~, the characteristics of ~~the derivative-based algorithms~~algorithms using different seminorms ~~–~~have ~~can be~~ been well explored in the literature. The LTD algorithm can be considered as one case that minimizes the seminorm consisting of the third-order derivatives (Bini and Capovani, 1986). Other algorithms can also be formulated by using different seminorms. ~~In this paper~~On the basis of this idea, we propose a new minimum curvature (MC) algorithm ~~based on the variational interpolation~~using a seminorm approximating the integral of the squares of the curvature. This algorithm generates a smooth reconstruction approximating the application of cubic spline interpolation. We compared the algorithm with the NNLS, LTD, and GT-MG algorithms by using multiple test maps. We demonstrated its effectiveness and two main aspects of this method. First, smooth effect similar to spline interpolation is achieved during the reconstruction process by using high-resolution grid division, and second, ~~It improves~~ the computational efficiency is markedly better than that of~~comparing to~~ the LTD algorithm ~~by~~ through ~~using the corresponding biharmonic equation~~halving the number of linear equations ~~instead of the~~according to the new smoothness seminorm~~ to construct the additional linear equation at each pixel of the map, the new MC algorithm reduces the number of equations to half of the LTD algorithm and eventually reduces the computation time~~. This approach achieves the same performance but is easier to perform ~~Another innovation of the new algorithm is to use much larger number of grids than the number of beams instead of the comparable numbers of grids and beams in the traditional methods. C~~than~~omparing to~~ the GT-MG algorithm which has ~~the~~ complicated operations involving multiple grids and grid translation~~ operations in GT-MG algorithm~~. More specific algorithms applied for the ORS-CT method for mapping atmospheric chemicals could be further formulated and evaluated similarly~~, this approach offers the same performance but is much easier to realize. We compared the new MC algorithm with the NNLS, LTD, and GT-MG algorithms using multiple test maps. These tests showed the effectiveness of the new MC algorithm. This study also demonstrates the feasibility of introducing techniques from the Tikhonov regularization and spatial interpolation to the ORS-CT method for mapping atmospheric chemicals~~.

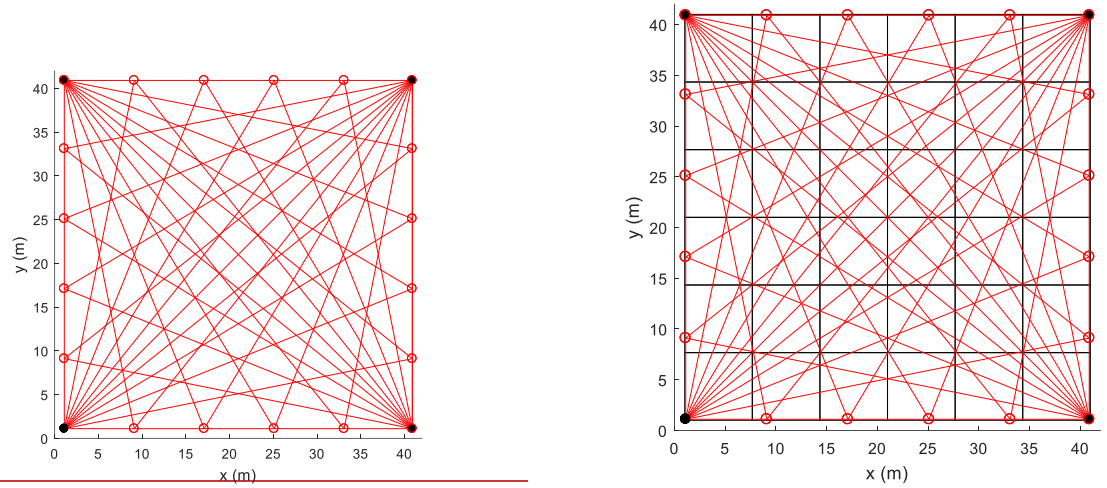
2 Materials and method~~ologie~~s

2.1 ORS-CT and beam geometry

The area of the test field ~~is~~ was 40 m ~~×~~ 40 m. ~~We use the open-path tunable diode laser (OP~~ Open-path TDL~~)~~ was used as the ORS analyzer, which ~~is~~ was installed on a scanner and ~~aims~~ aimed at multiple retroreflectors by scanning periodically and continuously. ~~The beam geometry can be categorized into overlapped and non-overlapped beam geometry based on the way~~

~~the beams are deployed.~~ To compare ~~with~~ the results with those of the GT-MG algorithm, we used an overlapping beam configuration similar to ~~the one~~that used ~~in~~ by Verkruysse and Todd (2005).

As shown in Fig. 1, four TDL analyzers ~~are~~ were located at the four corners of the test field. ~~Th~~The~~e~~ retroreflectors were ~~are~~ evenly distributed along the ~~on the~~ edges of the field. The total number of retroreflectors was 20. Each retroreflector reflected the laser beams coming from two different directions. Excluding the overlapped beams along the diagonals, the ~~The~~ total beam number ~~is~~ was 38. For the traditional pixel-based algorithm, the pixel number ~~of grids~~ should ~~be less than~~not exceed ~~or equal to~~ the beam number. Therefore, we divided the test field into 6~~-~~×~~-~~6=36 pixels~~grids~~. The concentration within each ~~grid~~ pixel ~~is~~ was assumed to be uniform.



~~(a)~~ ~~(b)~~

Figure 1. The ~~test field and~~ beam configuration and grid division. ~~(a) beam geometry; (b) beam geometry with~~The field was divided into 6 ×~~-~~6 grid pixels. Four open-path TDL analyzers were located at the four corners. A total of 20 retroreflectors were distributed on the edges of the field.

~~At~~ For each ~~retroreflector~~laser beam, the ~~the path-integrated concentration (~~PIC was measured by the analyzer~~) is measured~~. The predicted PIC for one beam was equal to the sum of the multiplication of the pixel concentration and the length of the beam inside the pixel. In general, let us assume that the site is divided into $N_c=m\times n$ pixels, which are arranged as a vector according to the left-to-right and top-to-bottom sequence and indexed by j. The average concentration for the j-th pixel is c_j. The total number of laser beams is N_b, which are indexed by i. The length of the i-th laser beam passing the j-th pixel is L_{ij}. Then, for the i-th beam, the measured PIC b_i is contributed by all pixels. We have the following linear equation

~~The PIC of the i-th beam and the system of linear equations for all the paths are~~

$$~~p~~b_i = \sum_{j=1}^{N_c~~N_n~~} L_{ij}c_j ~~,\ c_j \geq 0~~ \quad (1)$$

A system of linear equations can be set up for all beams

$$~~\boldsymbol{p}~~\boldsymbol{b} = \boldsymbol{Lc} ~~,\ c_j \geq 0~~ \quad (2)$$

where $\boldsymbol{L}$ is the kernel matrix that incorporates the specific beam geometry with the pixel dimensions, $\boldsymbol{c}$ is the unknown concentration vector of the pixels, and $\boldsymbol{b}$ is a vector of the measured PIC data. With the least squares approach, the reconstruction can be solved by minimizing the following problem~~where p is the PIC, $N=m\times n$ is the total number of grids, m, n the row and column number of the grids, i, j is the index of path number and cell number respectively, L_{ij} is the beam length of the ith beam in the jth cell, c_j is the concentration of the jth cell.~~

$$\min_{\boldsymbol{c}} \ ||\boldsymbol{Lc}-\boldsymbol{b}||_2^2, \text{ subject to } \boldsymbol{c} \geq 0 \quad (3)$$

where $\|\cdot\|_2$ denotes the Euclidean norm. This non-negative constrained linear least squares ~~inverse~~ problem ~~is~~ can be solved by the widely used ~~iterative~~ NNLS optimization ~~algorithms~~algorithm ~~of NNLS~~ (Lawson and Janson, 1995), which is an active-set optimization method using an iterative procedure to converge on the best fit of positive values. The routine "lsqnonneg" in MATLAB software was used in this study. The optimal least squares solution is not smooth because the minimizing process does not introduce smoothness *a priori* information. Herein, the NNLS algorithm in the tomographic reconstruction refers to solving the original problem by using the NNLS optimization algorithm without adding additional *a priori* information. When the system of linear equations is underdetermined, the solution is not unique. Additional information must be introduced to choose the appropriate solution.~~.~~

2.2 LTD algorithm and Tikhonov regularization ~~and LTD algorithm~~

The LTD algorithm introduces the smoothness ~~*feature*~~information ~~by setting~~through setting the ~~the~~ third-order derivatives~~s~~ of the concentration to ~~be~~ zero at each ~~grid~~ pixel in both ~~each~~x and y directions, ~~which will~~thus generating~~e~~ solutions that are locally quadratic (Price et al., 2001). We define c_j as an element of a one-dimensional (1-D) concentration vector of the pixels, but the pixels also have two-dimensional (2-D) structure according to the grid division of the site area and can be indexed by the row number k and column number l, where $j=(k-1)n+l$. We use $C_{k,l}$ to denote the pixel concentration at the pixel located at the k-th row and l-th column of the grids. The third-derivative prior equations at the (k, l) pixel are defined as ~~Assuming the grid indices in x, y directions are i, j, the new equations at one grid are~~

$$\frac{d^3Ce}{dx^3} = (C_{k+2,l} - 3C_{k+1,l} + 3C_{k,l} - C_{k-1,l})\frac{1}{\Delta x}\frac{1}{\Delta d}(\cancel{c_{i+2,j} - 3c_{i+1,j} + 3c_{i,j} - c_{i-1,j}}) = 0$$

$$\frac{d^3Ce}{dy^3} = (C_{k,l+2}\frac{1}{\Delta d}(\cancel{c_{i,j+2} -} -3C_{k,l+1}\cancel{c_{i,j+1} +} +3C_{k,l}\cancel{c_{i,j}} - C_{k,l-1})\cancel{- c_{i,j-1}})\frac{1}{\Delta y} = 0$$

(~~3~~4)

where ~~$\Delta d=$~~$\Delta x=\Delta y=\Delta d$ is the grid length in the x, y direction ~~divided in Eq. 3 to convert the grid-based derivatives into physical units~~. Therefore, two additional linear equations are introduced at each ~~grid~~pixel defined by Eq. (~~3~~4). There will be $2N_c$ linear equations appended to the original linear equations (defined by Eq. (2), thus results~~s~~ing in a new over-determined system of linear equations with ($2N_c+N_b$) equations and N_c unknowns. ~~that is over-determined.~~

A weight ~~needs to be~~needs to be assigned to each equation depending on the uncertainty of the observation. Under the assumption that the analyzers have the same performance, the uncertainty is mainly associated with the path length. Therefore,

equations are assigned ~~For the laser paths, w~~weights ~~are assigned~~ inversely proportional to the path length to ~~make sure~~ensure that different paths have equal ~~weights~~influences. ~~In this application~~Herein, the lengths of the laser paths are ~~comparable (less than two times difference)~~approximately equal to each other. ~~Thus~~Therefore, ~~,~~ their weights ~~are~~are ~~all~~ set to the same value and scaled to be ~~one~~1. The weights for the third-derivative prior equations ~~are~~are assigned as the same value ~~(~~of w~~)~~, because they ~~are~~are all based on the ~~single~~same grid length. The determination of w follows the scheme for determining the regularization parameter described below. With the least squares approach, the reconstruction is intended to minimize the following problem

$$\min_{\boldsymbol{c}} \left\| \begin{bmatrix} \boldsymbol{L} \\ w\boldsymbol{T} \end{bmatrix} \boldsymbol{c} - \begin{bmatrix} \boldsymbol{b} \\ 0 \end{bmatrix} \right\|_2^2, \text{ subject to } \boldsymbol{c} \geq 0 \quad (5)$$

where $\boldsymbol{T}$ is the kernel matrix for the third-derivative prior equations. Assuming that the new augmented kernel matrix is $\boldsymbol{A}$ and the observation vector is $\boldsymbol{p}$, the new system of linear equations will be $\boldsymbol{Ac}=\boldsymbol{p}$. The non-negative least squares solution ~~The approximation solution can be~~was also found ~~by~~by ~~traditional methods like~~the NNLS optimization algorithm. If the non-negative constraints are ignored, the least squares solution can be found analytically as $(\boldsymbol{A}^T\boldsymbol{W}\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{W}\boldsymbol{p}$, where $\boldsymbol{W}$ is a diagonal matrix whose diagonal elements are the weights (Price et al., 2001).

The -LTD algorithm actually constructs a regularized inverse problem. It can be viewed as ~~one~~a special case of ~~the Tikhonov L_2 regularization, which is~~ the ~~most~~ well-known Tikhonov regularization technique. ~~Its form~~The Tikhonov L_2 regularization can be written as the following minimization problem (Gholami and Hosseini, 2013)[21]

~~$\arg\min_{\boldsymbol{c}\in\mathbb{R}^n} min \quad \|\boldsymbol{Lc}-\boldsymbol{b}\|_2^2+\mu\|\boldsymbol{D}_k\boldsymbol{c}\|_2^2 \{\|\boldsymbol{Lc}-\boldsymbol{p}\|_2^2+\mu\|\boldsymbol{R}_t\boldsymbol{c}\|_2^2\}$~~

(~~3~~46)

~~Where~~where the first term represents the discrepancy between the measured and predicted values, the second term is the regularization term adding a smoothness penalty to the solution, μ is the regularization parameter controlling the conditioning of the problem, and matrix ~~$\boldsymbol{R}_t$~~ $\boldsymbol{D}_k$ is the regularization operator, which is typically ~~an approximation of the~~a ~~ith~~kth-order difference ~~derivative~~ operator. The -first- and ~~the~~ second-order difference ~~derivative~~ operators are commonly used.

We can see that the~~For~~ LTD algorithm~~, the~~ uses the third-order

~~The LTD algorithm uses the third-order derivative operator. Assuming the grid indices in x, y directions are i, j, the third-order derivative at y-direction is~~

~~$\frac{d^3c}{dj^3} = c_{i,j+2} - 3c_{i,j+1} + 3c_{i,j} - c_{i,j-1}$~~ ~~(4)~~

~~The third-order derivative~~forward difference operator~~in matrix form is~~

$$\boldsymbol{RD}_3 = \begin{bmatrix} -1 & 3 & -3 & 1 & & & \\ & -1 & 3 & -3 & 1 & & \\ & & & \ddots & & & \\ & & & 1 & 3 & -3 & 1 \\ & & & & 1 & 3 & -3 & 1 \end{bmatrix} \frac{1}{\Delta d} \in \mathbb{R}^{(m-3)\times n} \quad (~~5~~7)$$

For pixelgrids onat the edges, we use the second-order derivative and first-order difference operators derivative can be used. The regularization parameter is set to be inversely propositional to the grid length. By setting the derivative to zero, the algorithm introduces two linear equations at each grid. The resulted system of linear equations is over-determined, which can be solved by NNLS algorithm. The regularization parameter is the same as theanalogous to the weight assignedparameter for the prior equations to the third-derivative equations in the LTD algorithm. Therefore, $\mu = w/\Delta d$, which is inversely proportional to the grid length.

The regularization parameterIt determines the balance between data fidelity and regularization terms. Determination of the optimum regularization parameter is an important step ofin the regularization method. It determines the balance between data fidelity and regularization terms. There are some parameter-choice methods existed with or without *a priori* information of the noise (Hamarik et al., 2012). However, the regularization parameter is problem and data dependent. There is no general-purpose parameter-choice algorithm whichthat will always produce a good parameter. For simplicity, we use the commonly used method based on the discrepancy principle (Hamarik et al., 2012). The regularization parameter μ (μ_i) is chosen from a finite section of a monotonic sequence. For each value of μ_i, an optimal solution is derived by solving Eq. 4the inverse problem. Then the discrepancy (the first term in Eq. 4) can then be calculated. The regularization parameter is determined to be the highest value that makes the discrepancy the $\|\boldsymbol{Lc} - \boldsymbol{bp}\|_2^2 = \|\boldsymbol{e}\|^2$, equal to *where* $\|\boldsymbol{e}\|^2$ *is the residual norm* $N_b\sigma^2$, where σ is the standard deviation of the noisewhich can be determined by examining the variance and uncertainty. In this applicationstudy, the reconstructions variedy only slowly with the regularization parameters,. Therefore, ppprecise selection of the weightsparameter wasis not necessary. For computational efficiency, the regularization parameter can bewas selected from four widely varying values. The one producesing the smallest discrepancy wasis used.

2.3 Variational approach interpolation and minimum curvature algorithm

Splines are special types of piecewise polynomials, which have been demonstrated to be very useful in numerical analysis and in many applications in science and engineering problems. They match given values at some points (called knots) and have continuous derivatives up to some order at these points (Champion et al., 2000). Spline interpolation is preferred over polynomial interpolation by fitting low-degree polynomials between each of the pairs of the data points instead of fitting a single high-degree polynomial. Normally, the spline functions can be found by solving a system of linear equations with unknown coefficients of the low-degree polynomials defined by the given boundary conditions.

The variational approach for interpolation provides a new way to find the interpolating splines and opens up directions in theoretical developments and new applications (Champion et al., 2000). Variational interpolation was motivated by the minimum curvature property of natural cubic splines, i.e., the interpolated surface minimizes an energy functional that corresponds to a physical bending energy. This principle provides flexibility in controlling the behavior of the generated spline. Given an observation z_k (k=1, …, N) measured at the k-th point whose position vector is $\boldsymbol{r}_k$, a spline function $F(\boldsymbol{r})$ for

interpolating the data points can be found through the variational approach by ~~is to~~ minimizinge the sum of the deviation from the measured points and the smoothness seminorm of the spline function~~the following problem~~

$$\min_{c \in \mathbb{R}^{Nn}} \sum_{k=1}^{N} |F(\boldsymbol{r}_k) - z_k|^2 + \omega I(F) \quad \{\|\boldsymbol{Lc - p}\|_2^2 + \omega |I(c)|^2\}$$

(~~6~~8)

where ω~~ω~~ is a positive weight, and ~~and~~ $|I(cF)|^2$ denotes the ~~square of~~ smoothness seminorm. The seminorm can be defined in various forms~~.~~ , ~~The~~ commonly ~~used ones are~~ the first, second, third derivatives, or their combinations. The solution~~s~~ ~~of~~ to the minimizing problem~~s~~ ~~are~~ is spline functions, which can also be found by solving a Euler-Lagrange differential equation corresponding to the given seminorm (Briggs, 1974).~~.~~

We can see that the minimizing problem in Eq. (8) has a similar form to the Tikhonov regularization but with a more flexible regularization term. The problem is that the variational interpolation is based on given data points, whereas the tomographic reconstruction is based on measured line integrals. However, we show herein that the variational approach for interpolation can also be applied to the latter problem to produce a smoothness solution with an effect similar to spline interpolation. In addition, on the basis of different seminorms, we can formulate many different reconstruction algorithms. In this way, we propose a new minimum curvature (MC) algorithm.

Under the assumption that the unknown concentration distribution is described by a function $f(x, y)$, (x_k, y_l) are the smallest coordinates of the j-th pixel at row k and column l of the 2-D grids, then the concentration c_j equals the average concentration of the pixel

$$c_j = \frac{1}{(\Delta d)^2} \int_{x_k}^{x_{k+1}} \int_{y_l}^{y_{l+1}} f(x, y) dx dy \quad (9)$$

The minimization problem according to the variational approach is formulated as

$$\min_f \sum_{i=1}^{N_b} \sum_{j=1}^{N_c} || L_{ij} c_j - b_i ||_2^2 + \omega I(f) \quad (10)$$

For the MC algorithm, we define the seminorm ~~Based on different seminorms, we can formulate many different reconstruction algorithms. As a demonstration, we propose the minimum curvature (MC) algorithm~~according to the minimum curvature principle, which is used in the ~~,~~geographic data interpolation to seek a 2-D surface with continuous second derivatives and minimal total squared curvature (Briggs, 1974). The minimum-curvature surface is analogous to elastic plate flexure, and it approximates the shape adopted by a thin plate flexed to pass through the observation data points with a minimum amount of bending. This method generates the smoothest possible surface while attempting to follow the observation data as closely as possible. ~~which uses t~~The seminorm in the MC algorithm is ~~equal to~~defined to be equal to the total squares curvature~~:~~

$$|I(f\boldsymbol{\epsilon})|^2 = \int \int \left(\frac{\partial^2 \epsilon f}{\partial x^2} + \frac{\partial^2 \epsilon f}{\partial y^2} \right)^2 dx dy \quad (~~7~~11)$$

This integral must be discretized according to the grid division. ~~According to the minimum curvature principle, the minimization can be carried out by solving the biharmonic equation. (Briggs, 1974)~~

~~$$\frac{\partial^4 \epsilon}{\partial x^4} + 2 \frac{\partial^4 \epsilon}{\partial x^2 \partial y^2} + \frac{\partial^4 \epsilon}{\partial y^4} = 0 \quad (8)$$~~

~~T deduce the normal difference equations, Eq. 7 is constructed directly in terms of grid values $c_{i,j}$.~~The discrete total squares curvature is

$$I = \sum_{k=1}^{n} \sum_{l=1}^{m} (I_{k,l}^2)^2$$

(~~9~~12)

where $I_{i,j}$ is the curvature at the (k,l) pixel~~x_i, y_j~~, which is a function of ~~c~~$C_{i,j}$ and its neighboring ~~grid~~pixel values. In two dimensions the approximation to the curvature is~~. In two dimensions, it is~~

$$I_{k,l} = (cC_{k+1,l} + cC_{k-1,l} + cC_{k,l+1} + cC_{k,l-1} - 4cC_{i,j})/(\Delta d)^2$$

(1~~0~~3)

To minimize the total squared curvature, we need

$$\frac{\partial I}{\partial cC_{k,l}} = 0$$

(1~~1~~4)

Combining Eq. (~~9~~11), (1~~0~~2), and (1~~1~~3), we ~~get~~obtain the ~~common~~following difference equation~~ for the biharmonic equation~~

$$\begin{aligned}&[cC_{i+2,j} + cC_{i,j+2} + cC_{i-2,j} + cC_{i,j-2} \\ &+2(cC_{i+1,j+1} + cC_{i-1,j+1} + cC_{i+1,j-1} + cC_{i-1,j-1}) \\ &-8(cC_{k+1,j} + cC_{k-1,j} + cC_{k,j-1} + Cc_{k,j+1}) + 20cC_{k,j}]/(\Delta d)^2 = 0\end{aligned}$$

(~~9~~125)

This equation is appended at each ~~grid~~ pixel as a smoothness regularization. Therefore, there is only one ~~additional~~prior equation at each grid~~The seminorm can be calculated at each pixel using the finite difference approach. Each item in the summation is set to be zero to generate an equation at that pixel. Multiple items will lead to multiple equations at one pixel. This is how the LTD algorithm does to add additional equations. In this paper, however, we will only add one equation at each pixel to reduce the number of equations. This is done by using the corresponding Euler Lagrange differential equation to the minimizing problem (Briggs, 1974). According to the minimum curvature principle, the minimization can be carried out by solving the biharmonic equation.~~

~~$$\frac{\partial^4 c}{\partial x^4} + 2\frac{\partial^4 c}{\partial x^2 \partial y^2} + \frac{\partial^4 c}{\partial y^4} = 0 \quad (8)$$~~

~~The corresponding finite difference equation is~~

~~$$\begin{aligned}&c_{i+2,j} + c_{i,j+2} + c_{i-2,j} + c_{i,j-2} \\ &+2(c_{i+1,j+1} + c_{i-1,j+1} + c_{i+1,j-1} + c_{i-1,j-1}) \\ &-8(c_{i+1,j} + c_{i-1,j} + c_{i,j-1} + c_{i,j+1}) + 20c_{i,j} = 0\end{aligned} \quad (9)$$~~

~~Therefore, one equation is set up for each grid~~ instead of two equations in the LTD algorithm. For ~~grids~~ pixels on the edges, we set the approximation of the first and second derivatives to ~~be~~ zeros. Under the assumption that $\boldsymbol{M}$ is the kernel matrix of the prior equations, the reconstruction aims to minimize the following problem ~~The weight ω is set to be inversely propositional to the grid length~~determined in the same way as determining the regularization parameter in the Tikhonov regularization ~~method~~.

$$\min_{\boldsymbol{c}} \ \|\boldsymbol{Lc}-\boldsymbol{b}\|_2^2+\lambda\|\boldsymbol{Mc}\|_2^2, \text{ subject to } \boldsymbol{c}\geq 0 \quad (16)$$

where the parameter λ is determined in the same manner as the regularization parameter in Tikhonov regularization method. Similar to the LTD approach, Tthe resultinged constrained system of linear equations is over-determined and iscan be solved by the NNLS optimization algorithm.

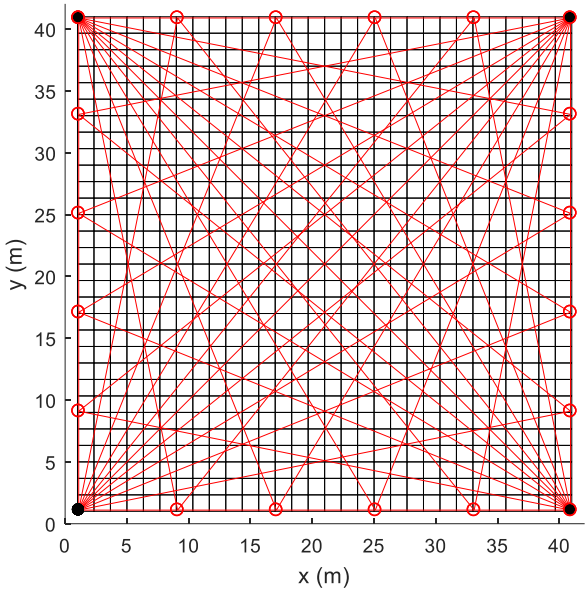


Figure 2. Beam geometry and a with 30×30 grid division of the sitegrids.

For conventional pixel-based reconstruction algorithms, the number of gridpixels (unknowns) should be no more thannot exceed the number of beams (equations) in order to get obtain a well-posted system of linear equationsproblem. Because there are only dozens tens of beamsbeams are usually used in an ORS-CT applications, the resultant grid spatial resolution is very coarse. The GT algorithm is one way to increase the resolution. B, but it needs requires several steps to complete the whole entire translation because . each translation uses a different grid division, and Athend the high-resolution map is generated after each reconstruction process must be conducted for each grid division. In the MC algorithm, we use only one division of very high-resolution grids directly before during the reconstruction. T he resultant system of linear equations remains determined Bbecause there is one of the derivative smoothness restriction at each pixel, the resulted system of linear equations is determined. As shown in Fig. 2, 30×30 grids pixels are used in the MC algorithm instead of the 6×6 grids pixels in the traditional NNLS algorithmapproach. Under this configuration, the number of linear equations for the LTD algorithm is approximately 38+30×30×2=1838, whereas the number for the MC algorithm is approximately 38+30×30=938. Thus, the MC approach decreases the number of linear equations to approximately half that of the LTD algorithm. The smoothness feature seminorm of the MC algorithm will guaranteeensures the a smooth effect between all the pixelssolution. This smoothing smooth effect is similar to the spline interpolation applied after the reconstruction process, except that the interpolationit is

achieved ~~automatically when the~~during the inverse ~~problem is solved~~process. This aspect is important because an interpolation after the reconstruction cannot correct the error resulting from the reconstruction in terms of coarse spatial resolution. The MC approach evaluates the discrepancy based on the high-resolution values that are the same as the reconstruction outcomes. Errors due to coarse spatial resolution are corrected during the process.

2.4 Test concentration data

The NNLS, LTD, and MC algorithms were compared by using multiple test maps. The results were also compared with those of the GT-MG algorithm. ~~In order t~~To ~~do this, w~~e set up test conditions similar to those~~at~~ used in Verkruysse and Todd (2005). The ~~source~~ concentration distribution from one source is defined by a bivariate Gaussian distribution

$$~~c~~g(x,y) = Q\exp[-\left(\frac{(x-x_0)^2}{\sigma_x^2}+\frac{(y-y_0)^2}{\sigma_y^2}\right)] \quad (~~10~~137)$$

where Q (0 to 40 ~~ppm~~mg/m^3) is the source strength, x_0, y_0 (0 to 40 m) is the peak location, and σ_x, σ_y are~~is~~ the width of the peaks ~~(~~with possible values of 2.8, 4.2, 5.7, and~~or~~ 7.1~~.)~~.

The source number varies from 1 to 5. For multiple sources, the resultant concentration distribution is the superposition value due to each source. For each source number, 100 maps were generated by randomly setting the source strength, location, and peak width from the defined ranges or set above.

2.5 Evaluation of reconstruction quality

A conventional image quality measure called nearness is used to describe the discrepancy between the original maps and the reconstructed maps. Nearness evaluates errors over all ~~the~~ grid cells ~~in~~ on the map (Verkruysse and Todd, 2005)

$$\text{Nearness}=\sqrt{\frac{\sum_i^{m\times n}(c_i^*-c_i)^2}{\sum_i^{m\times n}(c_i^*-c_{avg}^*)^2}} \quad (~~11~~148)$$

where m, n are the grid divisions on the x, y direction of the map, c_i^* is the synthetic value of concentration in the ith grid ~~on the map~~ generated by the Gaussian distribution model, c_i is the estimated value for the ith grid ~~in the map~~, and c_{avg}^* is the mean concentration of all ~~the~~ grids ~~in the map~~. A nearness value of zero implies a perfect match.

The effectiveness of locating the emission source is evaluated by the peak location error, which calculates the distance between the true and reconstructed peak locations.

$$\text{Peak location error} =\sqrt{(x_r-x_0)^2+(y_r-y_0)^2} \quad (~~12~~159)$$

where x_r, y_r are the peak locations ~~in~~ on the reconstruction map. For multiple peaks, only the location of the highest peak is~~was~~ calculated. The peak is located by searching for the largest concentration on the map. When multiple locations have the same values, the centroid of these locations is used.

Exposure error percentage is used to evaluate how well average concentrations in the whole field are reconstructed. It can reflect the accuracy of measuring chemical air emissions and emission rates from fugitive sources, such as agricultural sources and landfills (Verkruysse and Todd, 2004)

$$\text{Exposure error \%} = \left|\frac{\sum_i^{m\times n} c_i^* - \sum_i^{m\times n} c_i}{\sum_i^{m\times n} c_i^*}\right| \times 100\% \quad (13\,16\,20)$$

Herein, a measure using the resolution matrix is also applied to predict the reconstruction error due to different regularization approaches. Resolution matrices are commonly used to determine whether model parameters can be independently predicted or resolved, and how regularization limits reconstruction accuracy (Twynstra and Daun, 2012; von Clarmann et al., 2009). Ignoring the non-negative constraints, the generalized inverse matrices for the NNLS, LTD, and MC algorithms can be found by

$$\boldsymbol{G}_{NNLS} = (\boldsymbol{L}^T\boldsymbol{L})^{-1}\boldsymbol{L}^T$$

$$\boldsymbol{G}_{LTD} = (\boldsymbol{L}^T\boldsymbol{L} - \mu^2\boldsymbol{D}_3^T\boldsymbol{D}_3)^{-1}\boldsymbol{L}^T$$

$$\boldsymbol{G}_{MC} = (\boldsymbol{L}^T\boldsymbol{L} - \lambda^2\boldsymbol{M}^T\boldsymbol{M})^{-1}\boldsymbol{L}^T \quad (21)$$

The resolution matrix is defined as $\boldsymbol{R}=\boldsymbol{GL}$. The reconstruction error is given by

$$\delta\boldsymbol{c} = \boldsymbol{c}_{model} - \boldsymbol{c}_{exact} = (\boldsymbol{R} - \boldsymbol{I})\boldsymbol{c}_{exact} - \boldsymbol{G}\delta\boldsymbol{b} \quad (22)$$

where $\boldsymbol{c}_{model}$ and $\boldsymbol{c}_{exact}$ are the model-predicted and the exact concentrations, respectively, $\delta\boldsymbol{b}$ is the perturbation of the observation, $\boldsymbol{I}$ is the identity matrix, $(\boldsymbol{R} - \boldsymbol{I})\boldsymbol{c}_{exact}$ is the regularization error caused by the inconsistency between the measurement data equations and the prior information equations, and $\boldsymbol{G}\delta\boldsymbol{b}$ is the perturbation error.

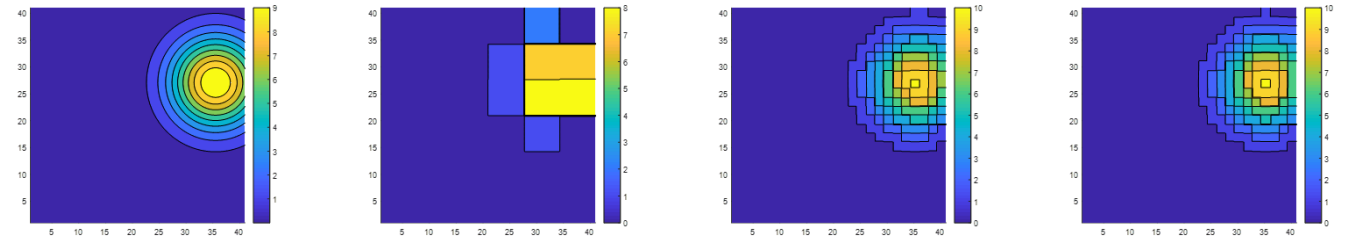
For the LTD and MC approaches using high-resolution grids, the kernel matrix $\boldsymbol{L}$ is rank-deficient, and the regularized solution is robust to perturbation error over a wide range of regularization parameters. Thus, the perturbation error is negligible, and the reconstruction error is dominated by regularization error (Twynstra and Daun, 2012). Because the resolution matrix is determined only by the beam configuration and the regularization approach, it is independent of the actual concentration distribution. Therefore, it is best used to evaluate different beam configurations that considerably influence the reconstruction accuracy. However, in this study the beam configurations are fixed. We can therefore use the resolution matrix to measure different regularization approaches. In an ideal experiment, $\boldsymbol{R}=\boldsymbol{I}$, thus implying that each unknown pixel value can be independently resolved from the measurement data. The regularization term forces the off-diagonal terms in $\boldsymbol{R}$ to be nonzero, thereby making the estimated concentration of each pixel a weighted average of the concentration of the surrounding pixels. We can use the Frobenius distance between $\boldsymbol{R}$ and $\boldsymbol{I}$ defining a measure of fitness to predict the reconstruction error (Twynstra and Daun, 2012).

$$\varepsilon = \frac{1}{N_c}\|\boldsymbol{R} - \boldsymbol{I}\|_F^2 \quad (23)$$

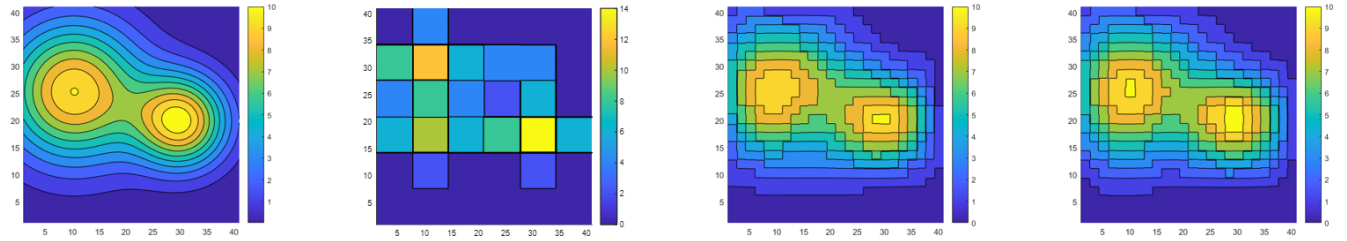
3 Results and discussions

In these tests, the traditional NNLS algorithm uses 6-×-6 grids, whereasile the LTD and MC algorithms both use 30-×-30 grids. The results of the GT-MG algorithm are from Verkruysse and Todd (2005), in which. They used a maximum basis grid resolution of 10-×-10 with 1/4 grid size as translation distance was used. Of nNote, that the test conditions wereare not exactly the same as thosee ones used by the GT-MG algorithm, which. did not measure the peak location error and used a different method to calculate the exposure error by limiting the calculation domain to a small area near the peak instead of the entire map. Therefore, the results of the GT-MG algorithm are listed provided as a reference. O and only the measure of nearness results of the GT-MG algorithm arewas compared because It it did not give the peak location error, and the exposure error was calculated only in a small area near the peak instead of the whole map in this paper. The original resolution of the reconstruction map by the NNLS algorithm is too coarse (6.7 m). In order toTo determine the peak locations more accurately, all the reconstructed concentration maps reconstructed by the NNLS algorithm are were spline interpolated by spline function using with a resolution of 0.5 m. Fig. 3 depicts some examples of the test maps and reconstructed maps generated by different algorithms with different source numbers.

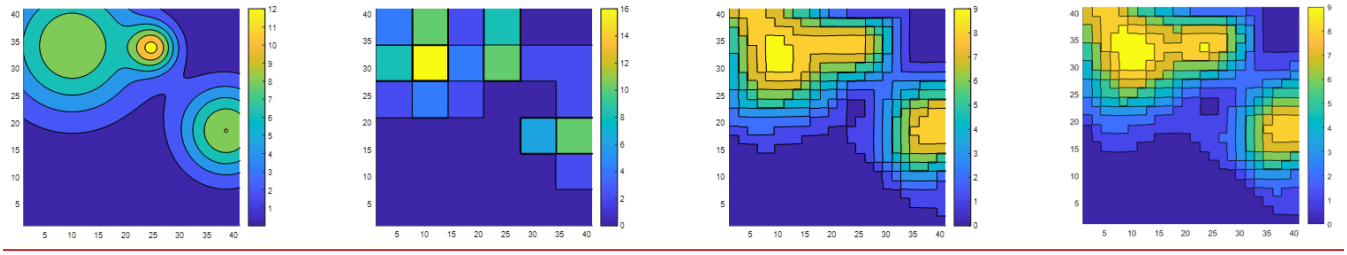
(a)



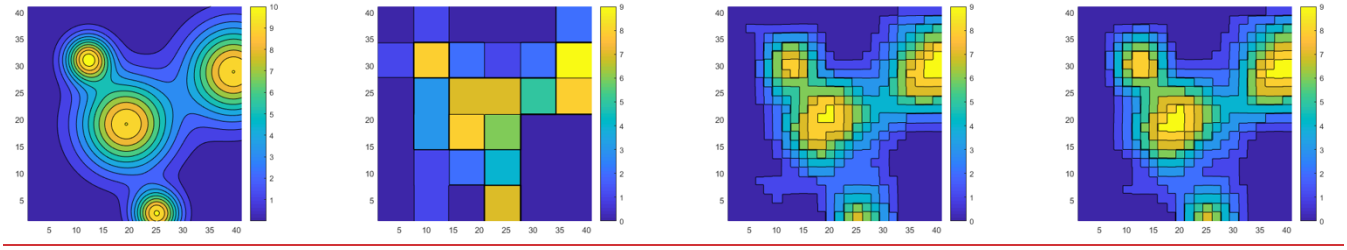
(b)



(c)



(d)



(e)

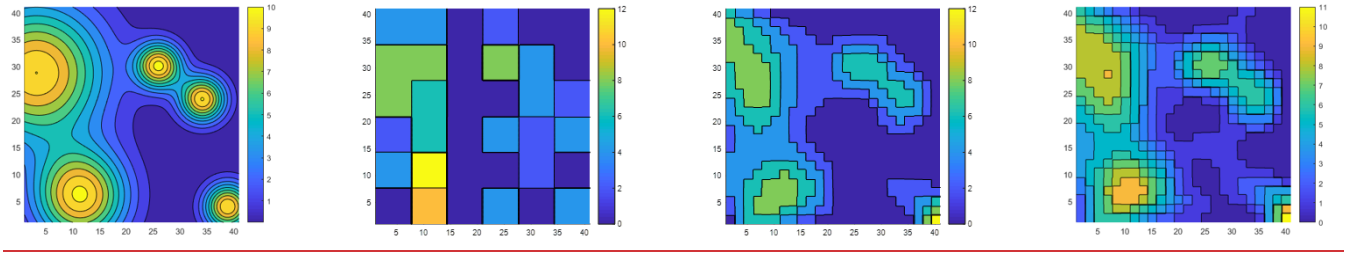


Figure ~~3~~3. Original test maps (first column) and corresponding ~~reconstructed~~ maps reconstructed ~~using~~ with the NNLS (second column), LTD (third column), and MC (forth column) algorithms. (a) ~~one~~ One source; (b) two sources; (c) three sources; (d) four sources; (e) five sources.

~~Fig. 3 depicts some examples of the test maps and reconstructed maps generated by different algorithms under different source numbers. As indicated by the nearness results, LTD and MC algorithms show better accuracy than the NNLS algorithm in the shapes and source strengths. Because the differences in accuracy (i.e., nearness, peak distance) between LTD and MC algorithms are very small, more specific evaluations may be needed to compare them using more complicated and realistic conditions.~~

3.1 Nearness

Table 1. Mean and standard deviation of nearness.

Source number	NNLS	LTD	MC	GT-MG*
1	0.40 (0.21)	0.13 (0.08)	0.11 (0.07)	0.09 (0.05)
2	0.38 (0.16)	0.15 (0.07)	0.13 (0.06)	0.16 (0.07)

3	0.40 (0.14)	0.18 (0.08)	0.17 (0.08)	0.19 (0.06)
4	0.40 (0.12)	0.20 (0.08)	0.19 (0.08)	0.25 (0.08)
5	0.43 (0.13)	0.22 (0.09)	0.21 (0.08)	0.27 (0.09)

*: The results of the GT-MG algorithm are from Verkruysse and Todd (2005), whose test conditions are not exactly the same as the conditions used ~~in this paper~~herein.

N~~The n~~earness is the most important measure of accuracy of the reconstructed map. It ~~, which~~ represents the reconstruction of peak heights, shapes, and the production of artifacts. The smaller the ~~number of sources~~nearness value, the better the reconstruction quality. In Table 1, the LTD, MC~~,~~ and GT-MG algorithms generally reduce the nearness values by more than 50% with respect to~~of~~ the values obtained by the NNLS algorithm. Under~~In~~ the condition of one source, they ~~can~~ reduce the nearness by approximately~~about~~ 70% with respect to the NNLS. The LTD, MC and~~,~~ GT-MG algorithms show increasing trends as the source number increases, thus~~which~~ imply~~ies~~ that the performance of the algorithm is affected by the ~~complicity~~ complexity of the underlying distribution. The nearness results of NNLS for different numbers of sources are almost the same because they are the results after spline interpolation. In fact, the original un-interpolated results also show increasing trends. The interpolation improves the results of the NNLS algorithm more than those of the LTD and MC algorithms, which already use high-resolution grids. The overall performance~~s~~ of the LTD, MC, and GT-MG algorithms is~~are~~ very similar, whereas~~ile~~ the new MC algorithm's performance~~s~~ is slightly better.

3.2 Peak location error

Table 2. Mean and standard deviation of peak location error.

Source number	NNLS (m)	LTD (m)	MC (m)
1	1.78 (0.93)	0.41 (0.45)	0.40 (0.56)
2	4.88 (8.21)	1.97 (5.98)	1.62 (4.81)
3	5.17 (8.35)	2.58 (6.77)	2.34 (6.17)
4	8.40 (11.53)	5.22 (10.28)	5.58 (10.76)
5	8.95 (11.32)	5.51 (10.15)	5.77 (10.41)

As shown ~~I~~in Table 2, the LTD and MC algorithms show better performance in peak location error than the NNLS algorithm. They generally improve the accuracy of peak location by 1 to 2 m. The errors of all ~~the~~ algorithms increase with~~as~~ the source number ~~increase~~.

One reason for this finding is that ~~Because only the highest peak was calculated, a large error may happen w~~when ~~there are~~ two or more peaks with ~~that are~~ comparable peak magnitudes on the map ~~to each other in magnitude~~ (~~see~~ Fig. 3), the algorithm may not identify the correct location of the highest peak. Therefore, a large error may occur ~~Therefore~~when the highest value on the reconstructed map is located on the wrong peak~~, this measure is more meaningful when the source number is small.~~.

3.3 Exposure error

Table 3. Mean and standard deviation of exposure error.

Source number	NNLS (%)	LTD (%)	MC (%)
1	29.26 (19.29)	1.51 (2.11)	1.61 (2.15)
2	19.91 (11.99)	1.05 (1.26)	1.09 (1.19)
3	16.05 (9.78)	1.16 (0.97)	1.31 (0.92)
4	12.99 (6.92)	1.15 (1.04)	1.24 (1.00)
5	12.13 (8.62)	1.30 (1.14)	1.45 (1.16)

As shown ~~I~~in Table 3, ~~we can see~~ a significant improvement in~~of~~ exposure error is achieved by the MC and LTD algorithms compared~~ing~~ with~~to~~ the NNLS algorithm. The exposure error reflects the accuracy of the overall emissions measurement other than the concentration distribution. The performance~~s~~ of the LTD and MC algorithms is~~are~~ very similar ~~to each other~~. Unlike the trends shown by ~~that~~ the NNLS ~~shows~~ in the nearness and peak location error, its performance in exposure error improves ~~becomes better~~ with ~~the~~ increasing~~e of~~ source number. A plausible cause of this phenomenon may~~could~~ be that the distribution becomes more uniform with larger numbers of sources. Because ~~, implying that t~~the NNLS algorithm ~~uses coarse grid division,~~ it produces concentrations with very low spatial resolution and fits the true distribution better when the distribution becomes more uniform~~is more suitable for relatively uniform distributions~~.

3.4 Computation time

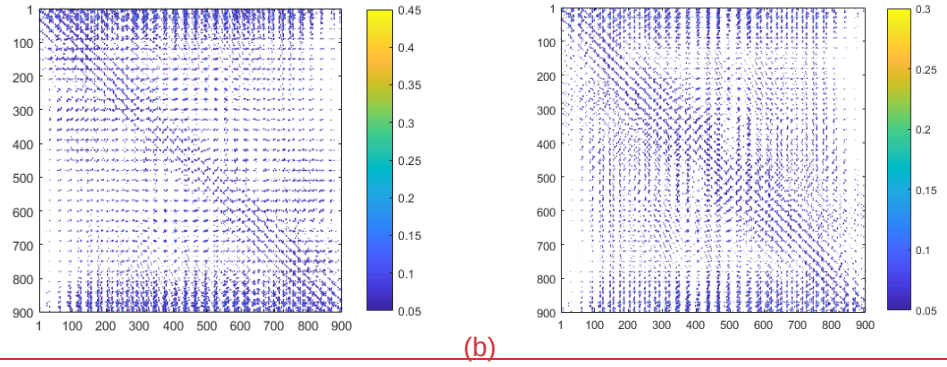
Table 4. Mean and standard derivation of computation time.

Source number	LTD (s)	MC (s)	Ratio (MC/LTD)
1	11.08 (14.27)	8.06 (10.03)	0.73
2	21.02 (17.89)	14.17 (11.46)	0.67
3	34.44 (19.93)	22.76 (13.61)	0.66
4	43.58 (20.45)	29.13 (12.92)	0.67
5	59.48 (21.98)	38.74 (15.05)	0.65

In Table 4, the computation times for the LTD and MC algorithms are compared. The computation time generally increases with ~~the~~ increasing~~e of~~ source number. The MC algorithm is faster than the LTD algorithm because it has approximately~~about~~ half the number of the linear equations ~~than that in~~as the LTD algorithm. The ratio results show~~s~~ that the MC algorithm's computation time is approximately~~about~~ 65% ~~of~~ that of the LTD algorithm when the source number is five. The trend of the ratio implies that the advantage of the MC algorithm becomes more clearer~~obvious~~ with ~~the~~ increasing complexity of the underlying distribution.

3.5 Fitness

Contour plots of the resolution matrices for the LTD and MC algorithms are shown in Fig. 4. (a) and (b). The fitness values for the LTD and MC algorithm are 1.4411 and 1.3878, respectively. The MC algorithm shows slightly better performance. The off-diagonal elements are not zeros. The reconstructed concentration at each pixel is a weighted average of the concentrations of the surrounding pixels according to the smoothness regularization. Each row of the resolution matrix can be regarded as smoothing weights. Because the pixels have a 2-D arrangement, we show the 2-D display of the row of the 106th pixel (row and column indices are 4 and 16) in the resolution matrix for the LTD and MC algorithms in Fig. 4. (c) and (d) as an example. The dependence on the beam geometry can be seen on both pictures. Because the beam configuration is fixed, the difference between the fitness values is mainly caused by the use of different regularization approaches. The fitness difference between the LTD and MC algorithms is very small, which may indicate that both algorithms have similar smoothness effects. This result coincides with the results from other measures discussed above. The 2-D display of the diagonal elements of the resolution matrix are shown in Fig. 4. (e) and (f), which are not much useful in this case.



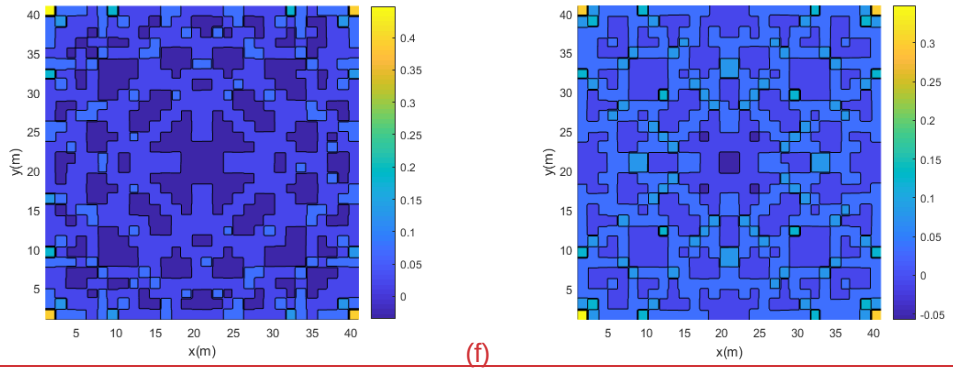


Figure 4. Contour plot of the resolution matrix for (a) the LTD algorithm (b) the MC algorithm. 2-D display of the row vector of the 106th pixel in the resolution matrix for (c) the LTD algorithm (d) the MC algorithm. 2-D display of the diagonal elements of the resolution matrix for (e) the LTD algorithm (f) the MC algorithm.

3.6 Influence of the grid size

The derivatives are approximated by the finite differences during the discretization process. The finite grid length causes discretization error and affects the reconstruction results. We studied the influences of different grid divisions by investigating the changes of the nearness, peak location error, exposure error, and computation time with respect to the pixel number. Five different grid divisions were used: 6×6, 12×12, 18×18, 24×24, and 30×30. The peak number was five. A total of 100 maps were tested for each grid division. The results of the averaged values are shown in Fig. 5.

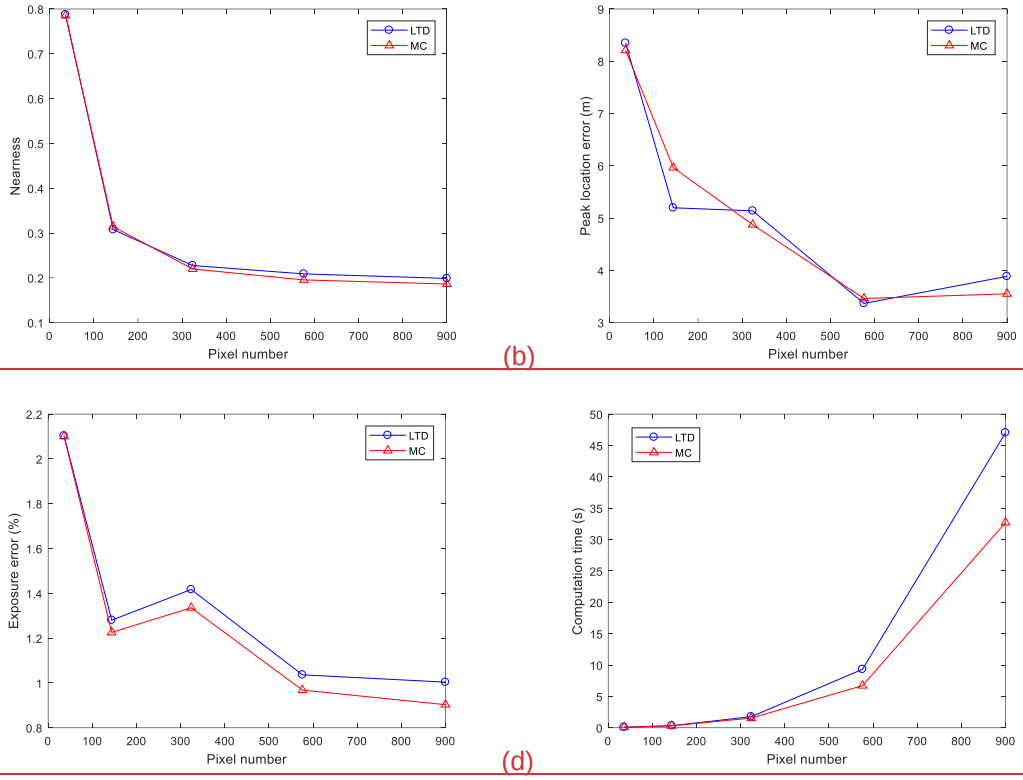


Figure 5. The change of (a) nearness (b) peak location error (c) exposure error percentage (d) computation time with respect to the pixel number.

The nearness, peak location error, and exposure error generally illustrate decreasing trends with increasing pixel number. The MC algorithm shows slightly better performance than the LTD algorithm with increasing pixel number. The performance improvement becomes slow for both algorithms when the division is finer than 24×24. The computation time shows approximately exponential growth trend with increasing pixel number. The LTD algorithm has a faster increasing rate than the MC algorithm. To conclude, the reconstruction performance is improved for both LTD and MC algorithms with increasing pixel numbers, but at the cost of exponential growth of the computation time. And the improvement becomes small when the resolution is higher than certain threshold value (24×24 herein). Therefore, there should be a balance between the performance and the computation time.

4 Conclusion

~~With the purpose of~~To understand the characteristics of the smoothness constraints and to seek more flexible ~~achieving a~~ smooth reconstruction, we first ~~found~~ identified the LTD algorithm as a special case of~~that the~~ Tikhonov regularization. Then,

more flexible smoothness constraints were found through the smoothness seminorms ~~and the~~according to variational ~~spatial~~ interpolation ~~theories~~ theory. The smoothness seminors were successfully adopted ~~can be applied to~~in ~~the~~ ORS-CT inverse problems. ~~Based on~~On the basis of ~~the~~ variational approach ~~approach in the interpolation theory~~, we proposed a new MC algorithm by using a seminorm approximating the sum of the squares of the curvature~~the corresponding biharmonic equation instead of the smoothness seminorm to construct an additional linear equation at each pixel of the map~~. The new algorithm improves ~~the~~ computational efficiency ~~comparing to the LTD algorithm~~ through~~by~~ reducing the number of linear equations ~~by~~ to half that of the LTD algorithm. It is simpler to perform than ~~. Comparing to~~ the GT-MG algorithm~~, which achieves high-resolution reconstruction by shifting the low-resolution base grids and averaging the resulted maps, the new MC algorithm uses~~ by directly using high-resolution grids ~~directly~~ during the reconstruction ~~process to generate a high-resolution map immediately after the reconstruction is done. Therefore, it is much simpler in realization~~.

The MC, LTD, and NNLS algorithms were compared by using multiple test maps ~~with the GT-MG algorithm as a reference~~. The new MC algorithm shows ~~almost the same~~similar performance as the LTD algorithm, but only ~~needs~~ requires ~~about~~ approximately 65% ~~of its~~the computation time. The smoothness-related algorithms of LTD, MC, and GT-MG all show better performance than the traditional NNLS algorithm: the nearness of reconstructed maps is improved by more than 50%, the peak distance accuracy is improved by 1- 2 m, and the exposure error is improved by more than ten~~10~~ times. Because ~~the~~ differences in accuracy ~~(i.e., nearness, peak distance)~~ between the LTD and MC algorithms are very small, more specific evaluations may be needed ~~to compare them~~by using more complicated and realistic conditions.

These comparisons ~~prove~~ demonstrate the feasibility of introducing ~~theories of Tikhonov regularization and~~the theory of ~~spatial~~ variational interpolation ~~techniques to mapping distribution of atmospheric chemicals using the ORS-CT techniques~~. ~~With~~ On the basis of the seminorms~~these theories~~, it is easier to understand the advantages and the drawbacks of ~~the current~~different algorithms. Common problems such as the over-smooth issue may be ~~solved~~ improved by ~~testing~~ formulating more algorithms suitable for ~~techniques in these theories and using them in the~~ ORS-CT applications. ~~We need to n~~Note that although the smoothness is very good *a priori* information for the reconstruction problem, beam configuration and underlying concentration distribution are also important factors affecting the reconstruction equality. To further improve the reconstruction quality, extra *a priori* information ~~based on~~according to ~~the~~ specific application may be added to the inverse problem. For example, ~~the~~ statistic information of the underlying distribution ~~,~~ or ~~the~~ information resulting from the fluid mechanics.

Code and data availability. Data and code are available on request by contacting the authors.

Author contributions. KD was responsible for acquiring funding for this research. SL designed the algorithm and conducted the tests. SL and KD were both involved in data analysis. Both authors contributed to writing and editing the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors are grateful to the supports by the following grants: Discovery Grant from Natural Sciences and Engineering Research Council (NSERC) of Canada (RGPIN-2020-05223), John R. Evans Leaders Fund (JELF) and Infrastructure Operating Fund (IOF) from Canada Foundation for Innovation (CFI) (35468), University Research Grant Committee (URGC) seed grant from University of Calgary (1050666).

References

Arghand, T., Karimipanah, T., Awbi, H.B., Cehlin, M., Larsson, U., Linden, E.: An experimental investigation of the flow and comfort parameters for under-floor, confluent jets and mixing ventilation systems in an open-plan office. Building and Environment, 92, 48-60. DOI: 10.1016/j.buildenv.2015.04.019, 2015.

Belotti, C., Cuccoli, F., Facheris, L., Vaselli, O.: An application of tomographic reconstruction of atmospheric CO2 over a volcanic site based on open-path IR laser measurements. IEEE Transactions on Geoscience and Remote Sensing, 41:11, 2629-2637. DOI: 10.1109/TGRS.2003.815400, 2003.

Briggs, I.C.:, 1974. Machine contouring using minimum curvature. Geophysics, 39:1, 39. DOI: 10.1190/1.1440410, 1974.

Cehlin, M.: Mapping tracer gas concentrations using a modified Low Third Derivative method: numerical study. International Journal of Ventilation, 18:2, 136-151. DOI: 10.1080/14733315.2018.1462935, 2019.

Drescher, A.C., Gadgil, A.J., Price, P.N., Nazaroff, W.W.: Novel approach for tomographic reconstruction of gas concentration distributions in air: Use of smooth basis functions and simulated annealing. Atmospheric Environment, 30:6, 929–940. DOI: 10.1016/1352-2310(95)00295-2, 1996.

Censor, Y.: Finite series-expansion reconstruction methods. Proceedings of the IEEE, 71:3, 409-419. DOI: 10.1109/PROC.1983.12598, 1983.

Champion, R., Lenard, C. T., Mills, T. M.: A variational approach to splines. The ANZIAM Journal, 42:1, 119-135. DOI: 10.1017/S1446181100011652, 2000.

von Clarmann, T., De Clercq, C., Ridolfi, M., Hoepfner, M., and Lambert, J.-C.: The horizontal resolution of MIPAS, Atmos. Meas. Tech., 2, 47–54, https://doi.org/10.5194/amt-2-47-2009, 2009.

Du, K., Rood, M. J., Welton, E. J., Varma, R. M., Hashmonay, R. A., Kim, B. J., Kemme, M. R.: Optical Remote Sensing to Quantify Fugitive Particulate Mass Emissions from Stationary Short-Term and Mobile Continuous Sources: Part I. Method and Examples. Environ. Sci. & Technol., 45, 658-665. DOI: 10.1021/es101904q-, 2011.

EPA: Measurement of Fugitive Emissions at a Landfill Practicing Leachate Recirculation and Air Injection. EPA-600/R-05/088, 2005.

Gholami, A., Hosseini, M.: A balanced combination of Tikhonov and total variation regularizations for reconstruction of piecewise-smooth signals. Signal Processing, 93:7, 1945-1960. DOI: 10.1016/j.sigpro.2012.12.008, 2013.

Giuli, D., Facheris, L., Tanelli, S.: Microwave tomographic inversion technique based on stochastic approach for rainfall fields monitoring. IEEE Transactions on Geoscience and Remote Sensing, 37, 5, 2536-2555. DOI: 10.1109/36.789649, 1999.

Hamarik, U., Palm, R., Raus, T. A family of rules for parameter choice in Tikhonov regularization of ill-posed problems with inexact noise level. Journal of Computational and Applied Mathematics, 236:8, 2146–2157. DOI: 10.1016/j.cam.2011.09.037, 2012.

Hashmonay, R.A., Natschke, D.F., Wagoner, K., Harris, D.B., Thompson, E.L, Yost, M.G.: Field Evaluation of a Method for Estimating Gaseous Fluxes from Area Sources Using Open-Path Fourier Transform Infrared. Environ. Sci. Technol., 35:11, 2309–2313. DOI: 10.1021/es0017108, 2001.

Hashmonay, R.A., Yost, M.G., Wu, C.F.: Computed tomography of air pollutants using radial scanning path-integrated optical remote sensing. Atmospheric Environment, 33, 267-274. DOI: 10.1016/S1352-2310(98)00158-7, 1999.

Hashmonay, R.A.: Theoretical evaluation of a method for locating gaseous emission hot spots. Journal of the Air and Waste Management Association, 58:8, 1100-1106. DOI: 10.3155/1047-3289.58.8.1100, 2012.

Herman, G. T.: Fundamentals of computerized tomography: Image reconstruction from projection, 2nd edition, Springer, 2009.

Lawson, C.L., Janson, R.J.: Solving least squares problems. Society for Industrial and Applied Mathematics: Philadelphia, 23, 158-165. DOI: 10.1137/1.9781611971217, 1995.

Twynstra, M. G., Daun, K. J.: Laser-absorption tomography beam arrangement optimization using resolution matrices. Applied Optics 51:29,7059-7068. DOI: 10.1364/AO.51.007059, 2012.

Mitasova, H., Mitas, L., Brown, W.M., Gerdes, D.P., Kosinovsky, I., Baker, T.: Modelling spatially and temporally distributed phenomena: new methods and tools for GRASS GIS, International Journal of Geographical Information Systems, 9:4, 433-446. DOI: 10.1080/02693799508902048, 1995.

Price, P.N., Fischer, M.L., Gadgil, A.J., Sextro, R.G.: An algorithm for real-time tomography of gas concentrations, using prior information about spatial derivatives. Atmos. Environ., 35, 2827. DOI: 10.1016/S1352-2310(01)00082-6, 2001.

Radon, J., On the determination of functions from their integral values along certain manifolds. IEEE transactions on medical imaging, **5**:4: 170-176, 1986.

Rudin, L., Osher, S.J., Fatemi, E.: Non-linear total variation based noise removal algorithms. Physica D, 60, 259–268.

Samanta, A., Todd, L.A., 2000. Mapping chemicals in air using an environmental CAT scanning system: evaluation of algorithms. Atmospheric Environment, 34, 699-709. DOI: 10.1016/S1352-2310(99)00331-3, 1992.

Tikhonov, A.N., Arsenin, V.Y.: Solutions of Ill-posed Problems, Winston and Sons, Washington, 1977.

Todd L., Ramachandran, G.: Evaluation of Optical Source-Detector Configurations for Tomographic Reconstruction of Chemical Concentrations in Indoor Air, American Industrial Hygiene Association Journal, 55:12, 1133-1143, DOI: 10.1080/15428119491018204, 1994.

Tsui, B.M.W., Zhao, X., Frey, E.C., Gullberg, G.T.: Comparison between ML-EM and WLS-CG algorithms for SPECT image reconstruction. IEEE Transactions on Nuclear Science, 38:6, 1766-1772. DOI: 10.1109/23.124174, 1991.

Verkruysse, W., Todd L.A.: Improved method “grid translation” for mapping environmental pollutants using a two-dimensional CAT scanning system. Atmospheric Environment, 38, 1801–1809. DOI: 10.1016/j.sigpro.2012.12.008, 2004.

Verkruysse, W., Todd, L.A.: Novel algorithm for tomographic reconstruction of atmospheric chemicals with sparse sampling. Environ Sci Technol., 39:7, 2247-2254. DOI: 10.1021/es035231v, 2005.

Wu, C.F., Chang, S.Y.: Comparisons of radial plume mapping algorithms for locating gaseous emission sources. Atmospheric Environment, 45, 1476-1482. DOI: 10.1016/j.atmosenv.2010.12.016, 2011.

Wu, C.F., Yost, M.G., Hashmonay, R.A., Park, D.Y.: Experimental evaluation of a radial beam geometry for mapping air pollutants using optical remote sensing and computed tomography. Atmospheric Environment, 33, 4709-4716. DOI: 10.1016/S1352-2310(99)00218-6, 1999.