

Article

# Adaptive Degenerate Space-Based Method for Pollutant Source Term Estimation Using a Backward Lagrangian Stochastic Model

Omri Buchman \* and Eyal Fattal \*

Applied Mathematics Department, Israel Institute for Biological Research (IIBR), POB 19, Ness Ziona 7410001, Israel

\* Correspondence: omrib@iibr.gov.il (O.B.); eyalf@iibr.gov.il (E.F.)

**Abstract:** A major challenge in accidental or unregulated releases is the ability to identify the pollutant source, especially if the location is in a large industrial area. Usually in such cases, only a few sensors provide non-zero signal. A crucial issue is therefore the ability to use a small number of sensors in order to identify the source location and rate of emission. The general problem of characterizing source parameters based on real-time sensors is known to be a difficult task. As with many inverse problems, one of the main obstacles for an accurate estimation is the non-uniqueness of the solution, induced by the lack of sufficient information. In this study, an efficient method is proposed that aims to provide a quantitative estimation of the source of hazardous gases or breathable aerosols. The proposed solution is composed of two parts. First, the physics of the atmospheric dispersion is utilized by a well-established Lagrangian stochastic model propagated backward in time. Then, a new algorithm is formulated for the prediction of the spatial expected uncertainty reduction gained by the optimal placement of an additional sensor. These two parts together are used to construct an adaptive decision support system for the dynamical deployment of detectors, allowing for an efficient characterization of the emitting source. This method has been tested for several scenarios and is shown to significantly reduce the uncertainty that stems from the insufficient information.



Academic Editors: William A. Anderson and Ching-Yuan Chang

Received: 3 December 2024

Revised: 2 January 2025

Accepted: 5 January 2025

Published: 10 January 2025

**Citation:** Buchman, O.; Fattal, E. Adaptive Degenerate Space-Based Method for Pollutant Source Term Estimation Using a Backward Lagrangian Stochastic Model. *Environments* **2025**, *12*, 18. <https://doi.org/10.3390/environments12010018>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** source term estimation; Lagrangian stochastic model; optimal sensor placement

## 1. Introduction

Air pollution modeling is the basic framework for better understanding, investigating, and regulating atmospheric air quality, as well as assessing the distribution of toxic environmental pollutants. The focus of many studies is on environmental sources, such as traffic and industrial sources containing inhalable aerosols up to 10  $\mu\text{m}$  in diameter (PM10), which are usually modeled as air-borne tracers at the city scale [1,2]. Such exposure was found to be related to cardiovascular disease [3], risk of asthma among children [4] and incidence rates of lung cancer among males [5]. Several field campaigns were held in different urban regions around the world in which particulate matter concentrations were measured at different heights, ranging from ground level up to several hundreds of meters above ground. These have exhibited large spatial variability both horizontally and vertically [6–9].

The realistic modeling of the transport and dispersion of pollutants at the city scale is a very challenging problem. This is due to the non-uniform three-dimensional complex nature of the turbulent flow field in the canopy layer and just above the canopy (known as the roughness sub-layer, RSL) [10,11]. The RSL consists of the lowest part of the atmospheric

surface layer from the ground up to two to five times the average height of the canopy elements [10]. The main flow feature of the RSL is the inflected wind profile induced by strong shear layers just above the rooftop level, separating high-speed to low-speed regions, leading to mixing layer instability and canopy-scale turbulent eddies [11–14]. This flow structure dominates the scalar, momentum and energy transfer in urban canopies and, in particular, the transport of air flow and dispersion of the pollutant [15–17]. Moreover, turbulence in canopies is non-local in nature since a large fraction of turbulent kinetic energy is produced at the top of the obstacles and is then transported into the canopy layer itself [11,18–20]. Recent direct Lagrangian measurement shows a fast Lagrangian velocity de-correlation timescale due to the highly inhomogeneous nature of the flow within the canopy [21–23]. This fast de-correlation was further shown to result in a quasihomogeneous flow regime, recovered well by a Lagrangian stochastic pollutant dispersion model (LSM) formulated using a pre-assumed Gaussian shape of velocity fluctuations probability distribution [21,22]. The LSM modeling approach has been found to provide a quantitative modeling of pollutant dispersion in realistic heterogeneous urban environments [1,2].

The inertial sub-layer lies above the RSL and is the surface layer's upper part, where the flow can be approximated as horizontally homogeneous. The defining characteristic of the inertial sub-layer is its turbulent kinetic energy (TKE) budget, which can be considered close to "local equilibrium" between TKE production and dissipation [24]. In the inertial sub-layer, since the atmosphere is both stationary and horizontally homogeneous, the turbulent fluxes are approximately independent of height, allowing for the description of this layer by the Monin–Obukhov similarity theory [10,25,26]. In addition, using the horizontal homogeneity assumption, estimates for rooftop wind essential for pollutant dispersion modeling can be derived with improved accuracy [27,28].

The ability to identify the location and emission rate of hazardous gasses or breathable aerosols has great importance for the ongoing effort of reducing the effect of accidental or unregulated releases. For example, the evaluation of the emission rate (which integrates into the total gas released in some time interval) can be used for the verification of compliance with regulatory standards in chemical industrial parks, intensive farming areas [29,30] or odorous pollutants emitted from landfills [31], sparing the need for placing detectors at specific suspected sites. In addition, fast and accurate identification of the source location can improve the efficiency of the emergency response and the accuracy of the risk assessment process in cases of accidental chemical releases.

The characterization of the source parameters based on a comparison between real-time measurements at static pre-deployed concentration detectors and a reliable dispersion model is known as the source term estimation (STE) problem. A detailed review on possible methodologies and approaches for solving this problem is beyond the scope of this paper, and can be found in other places [32]. Briefly, the two most common approaches for solving the STE problem are optimization or probabilistic approaches. In the optimization approach, one seeks the optimal combination of the source parameters that minimizes some regularized objective function that measures the discrepancy between the calculated and observed concentrations [33]. On the other hand, in the probabilistic approach, the probability density function of the source parameters is constructed based on Bayesian inference theory [34,35].

One of the main difficulties in solving the STE problem in realistic scenarios is the lack of sufficient measurements leading to non-uniqueness of the solutions. Clearly, adding more information by increasing the number of detectors will gradually solve this problem, but the deployment of a dense array of detectors for covering a wide area may not be feasible. For example, about 40 detectors were used for covering an area of less than 200 m × 200 m in the the well-known Mock Urban Setting Trial (MUST campaign; [36]),

which is frequently used for testing STE methods. Trying to maintain such a density of detectors for a wider area may be impractical for many applications.

A possible solution for this difficulty can be found by using a network of mobile detectors rather than that of static detectors, which has become possible in recent years due to the accelerated development of drones and other mobile platforms. The use of mobile network has two advantages: first, it allows one to spare the need for a pre-deployed dense array of detectors. More than that, the possibility of an adaptive update of the detector's position enables the network to optimize the gained information as time progresses. Moreover, since the general optimization problem of designing a network of detectors by individually placing all of them simultaneously over a finite grid of points is NP-hard [37], the extension of an existing detector network is a more approachable problem. Therefore, one of the key ingredients for a successful implementation of such an approach is the decision step, i.e., a mechanism for choosing the position of the next measurement based on a given set of already measured concentrations.

Several strategies for such a dynamic deployment mechanism have been proposed. For example, in the the 'Infotaxis' approach [38], the searcher chooses the move that maximizes the expected reduction in entropy of the posterior probability field, which amounts to having less uncertainty on the source parameters. In the 'Entrotaxis' approach [39], based on the maximum entropy sampling principles [40], the entropy of the predictive measurement distribution is taken as the reward function, which guides the searcher to where there is the most uncertainty in the next measurement outcome. Ristic [41] compared a number of search strategies based on different information rewards functions for determining the location of a diffusive source in turbulent flows. Keats [42] applied the Bayesian adaptive exploration (BAE) methodology [43,44], which provides a general methodology for choosing how future experiments should be performed so that information about the phenomenon of interest is maximized, thus also addressing the problem of adding a new detector to an existing array of detectors.

In this study, we propose a new approach for the adaptive deployment of mobile detectors. The general outline of this paper is as follows. First, the Lagrangian stochastic dispersion model (LSM), a well-established, realistic and consistent modeling approach, is utilized as the solution for the "Forward problem" (Section 2). As further discussed in Section 2, the computational effort that is required for solving the inverse problem can be reduced dramatically by introducing the backward time propagating LSM (BLSM) rather than the standard LSM. Then, in Section 2.2, a new mechanism for choosing the next measurement is presented, based on two sequential steps: first, a simple and efficient mechanism is proposed for identifying the suspected points in the parameter space that characterizes the source based on a given set of measurements. The method for the construction of this sub-space, which will be referred to as a "degenerate space", is given. Section 2.2 also presents a decision mechanism for choosing the location for a new detector based on a statistical evaluation of the expected degeneracy reduction as a function of the detector's position. These two steps can be combined iteratively to generate an adaptive algorithm that exploits the given information from a set of "old" measurements to plan the next measurement. The performance of this adaptive decision-making method was examined on a quasi-steady-state release for several scenarios differing by the initial locations of the detectors. The algorithm is shown to converge fast, sparing the need for a large number of pre-deployed detectors (Section 3).

## 2. Materials and Methods

### 2.1. The Lagrangian Stochastic Model

In order to solve the source estimation problem, one has to specify the underlying turbulent pollutant dispersion model. In this work, a Lagrangian stochastic model (LSM) developed at IIBR [1,2,21,45] has been used. This modeling approach is known to be able to consistently describe tracer dispersion phenomena for different temporal and spatial scales (near-field as well as far-field) and in rather complicated atmospheric scenarios, such as non-homogeneous turbulent regimes, complex terrain and canopies (urban and vegetation) [2,46,47], and is known to be superior to advection–diffusion-based approaches [48]. The LSM has been described in detail in many articles (see, e.g., [2,16,49]). Briefly, the basic idea is to propagate the position and velocity of the Lagrangian fluid particles according to the Langevine equations:

$$dr_i = u_i dt \quad (1)$$

$$du_i = a_i(r, u, t)dt + b_{ij}(r, u, dt)dW_j(t)$$

where  $r$  and  $u$  are the position and velocity of the particles and  $dW_i$  is a random increment selected from a normal distribution with variance  $dt$ . The deterministic coefficients  $a_i$  can be determined via the Fokker–Planck equation by satisfying the well-mixed (necessary) condition of Thomson [50]. According to the Kolmogorov–Obukhov similarity theory [51], the stochastic term takes the form of  $b_{ij} = \sqrt{(C_0\epsilon)}\delta_{ij}$ , where  $\epsilon$  is the average dissipation rate of the kinetic energy [52]. An additional step needed to formulate the drift term  $a_i$  is based on a pre-assumed shape of the probability density function for the velocities. This was found to be of a Gaussian quadratic shape in the highly inhomogeneous canopy flows case [21,22].

In order to perform backward-in-time (BLSM) simulations, a slight modification of the above forward Langevin model (Equation (1)) is needed [53,54]: the time increment  $dt$  will be negative, and there is a sign reversal in the the “damping” part within the drift terms  $a_i$  (i.e., only in the part containing  $b_{ij}^2$ ).

By counting all ‘touchdowns’ of the stochastic particles in some volume that represents the detector, one can calculate the probability density  $P^f(r, t|r', t')$  of whether a particle spreading from  $r'$  at time  $t'$  will reach the detector located at point  $r$  after time  $t$ . Fleck and Wilson [53] showed that, by modifying the Langevine equation (as stated above), one can propagate the particles backward in time (BLSM), i.e., from the future to the past, and that, for an incompressible fluid, the following relation holds:

$$P^f(r, t|r', t') = P^b(r', t'|r, t) \quad (2)$$

where  $P^b(r', t'|r, t)$  is the probability density function of whether a particle evolving from the detector located at  $r$  and propagated backward in time will reach the ‘source’ located at point  $r'$  at time  $t'$ . The transition probability for such a process can be related to the ensemble averaged concentration [53,55] by

$$d(r, t) = \int \int_{-\infty}^t S(r', t')P^f(r, t|r', t')dt'dr' \quad (3)$$

where  $S(r', t')$  is the mass source density ( $\text{kg m}^{-3} \text{ s}^{-1}$ ). The calculated concentration, averaged over the detector’s volume  $v_d$ , which is centered at position  $R$ , at stationary turbulence and a sustained uniformly distributed source can be written as

$$d(R) = \frac{q}{V_d} \int_{v_d} \left( \int_{v_s} \int_0^{\infty} P^f(r, t|r', 0)dt dr' \right) dr \quad (4)$$

where  $q$  is the emission rate and  $v_s$  is the volume of the source.

### 2.2. The Degenerate Space

Assume that the hypotheses space  $\Theta$  about the source is discretized by setting a grid of  $N$  cells. Each hypothesis will be labeled as  $\theta_j = (r_j, q_j)$ , where  $r_j$  is the coordinates of the center of the  $j$ 'th cell and  $q_j$  is the corresponding emission rate. We also assume that the source is fully characterized by one and only one hypothesis  $\theta_{ex} \in \Theta$ .

The concentration at the  $i$ 'th detector, located at  $R_i$  due to a source characterized by the  $j$ 'th hypothesis, can be calculated by Equation (4):

$$d(R_i|\theta_j) = \frac{q_j}{V} \int_{v_i} \left( \int_{v_j} \int_0^\infty P^f(r, t|r', 0) dt dr' \right) dr = \frac{q_j}{V} \int_{v_j} \left( \int_{v_i} \int_0^\infty P^b(r', 0|r, t) dt dr \right) dr' \quad (5)$$

where  $v_j$  is the  $j$ 'th cell in the parameter space and  $v_i$  is a small volume centered on  $R_i$ . The last passage in Equation (5), due to the forward-backward relation, can be very useful as will be seen later.

The model-measurement deviation  $\sigma$  represents the maximal discrepancy between the measured concentration  $d$  and the calculated concentration if the source is characterized by the exact hypothesis  $\theta_{ex}$ , i.e.,

$$d \in [d(R|\theta_{ex}) - \sigma, d(R|\theta_{ex}) + \sigma] \quad (6)$$

In this work, we shall assume that the error has a linear dependency on the concentration in order to retain the relative error as fixed. In order to avoid unrealistic small errors associated with low concentrations, a constant term  $\sigma_1$  will also be added, so the error will be

$$\sigma(\theta) = \sqrt{(\kappa d(R|\theta))^2 + \sigma_1^2} \quad (7)$$

where  $\kappa$  is some proportional constant.

Now, assume that we have measured concentration  $d$ . Which hypotheses are important in this case? Consider an arbitrary hypothesis  $\theta_k$ . If it was the exact hypothesis, the measurement outcome should be in the range described by Equation (6). This claim can be reversed—if the R.H.S of Equation (6) does not hold, then  $\theta_k$  is not the correct hypothesis. Formally,

$$d \notin [d(R|\theta_k) - \sigma(\theta_k), d(R|\theta_k) + \sigma(\theta_k)] \Rightarrow \theta_k \neq \theta_{ex} \quad (8)$$

This is a good way to select the good hypotheses from the bad ones—for a given measurement, eliminate all the hypotheses for which the constraint in Equation (6) does not hold because they cannot be considered as the exact hypothesis. This is the basis for the “Degenerate space”, defined as the set of all equally probable hypotheses for the source. If there are  $M$  measurements,  $d_1 \dots d_M$  located at positions  $R_1 \dots R_M$ , define the  $M$ -detectors degenerate space  $S(R_1 \dots R_M)$  as

$$S(R_1 \dots R_M) = \{\theta_k \in \Theta \mid |d(R_i|\theta_k) - d_i| \leq \sigma(\theta_k), i = 1 \dots M\} \quad (9)$$

where  $d(R_i|\theta_k)$  designates the calculated concentration due to the  $k$ 'th hypothesis at the  $i$ 'th detector's position, calculated by Equation (5). The last equation can be written recursively as

$$S(R_1 \dots R_M) = \{\theta_k \in S(R_1 \dots R_{M-1}) \mid |d(R_M|\theta_k) - d_M| \leq \sigma(\theta_k)\} \quad (10)$$

which will be useful later.

The construction of  $S(R_1 \dots R_M)$  by applying Equation (9) imposes a severe computational obstacle since, for every hypothesis in the parameters space, the expected concentration at every detector must be calculated, and then only the hypotheses for which the criterion holds are included. The number of required LSM operations for this task is

the number of different cells used to discretize the  $xy$  plane, which can be very large in practical situations (for example, in the following analyzed cases, the number of  $xy$  cells is  $2.5 \times 10^5$ ). Since the time taken for a typical LSM simulation is typically about 30 min, the total computational time could take weeks, even by operating on 100 parallel processors.

A significant reduction in the computational time can be achieved by introducing a backward Lagrangian stochastic model (BLSM) [33,56–58]. This is since a single BLSM simulation enables the calculations of the concentration at a specific detector from all the cells in the  $xy$  plane of the parameter space (see Equation (5)). Hence, the number of BLSM simulations is the number of detectors rather than the number of  $xy$  cells, which is much smaller, leading to a dramatic reduction in the total computational effort. To conclude, the degenerate space can be calculated efficiently by  $M$  BLSM simulations by combining its definition and the right part of Equation (5) as  $d(R_i|\theta_k)$ .

### 2.3. Adaptive Degeneracy Reduction

Now, let us take another step forward. After the construction of  $S(R_1 \dots R_M)$ , we are allowed to perform another measurement in any place we wish. Where should we place the new measurement? Clearly a “good” measurement will dramatically reduce the degenerate space, while a “bad” one leaves the degenerate space as almost the same. This intuition can be quantified by a small or large (up to 1) value of the reduction factor, defined as  $\Gamma(R_{M+1}) = \frac{|S(R_1 \dots R_{M+1})|}{|S(R_1 \dots R_M)|}$ . Note that the reduction factor can be calculated only after the new detector has been deployed at  $R_{M+1}$  since the measurement  $d_{M+1}$  is a crucial part of the degenerate space definition.

How can we predict the expected reduction as a function  $R_{M+1}$  of before the deployment of the new detector? Although it cannot be calculated directly, in Appendix A, we show that the probability that the reduction factor will be smaller than some positive number  $\alpha$ , designated by  $Pr(\Gamma(R_{M+1}) \leq \alpha)$ , can be bound from below by an accessible quantity. Formally, for every position  $R_{M+1}$  and for every  $0 < \alpha < 1$ , the following inequality holds:

$$Pr(\Gamma(R_{M+1}) \leq \alpha) \geq Pr(\gamma(R_{M+1}) \leq \alpha) \tag{11}$$

where  $Pr(\gamma(R_{M+1}) \leq \alpha)$  can be calculated directly from the  $M$ -detectors degenerate space and the BLSM as explained in the Appendix A. One possible strategy that can be taken from here is to search for the candidate  $R_{M+1}^*$  that maximizes  $Pr(\gamma(R_{M+1}) \leq \alpha)$ . Although this choice does not necessarily maximize  $Pr(\Gamma(R_{M+1}) \leq \alpha)$ , we are guaranteed that

$$Pr(\Gamma(R_{M+1}^*) \leq \alpha) \geq \max_{R_{M+1}} Pr(\gamma(R_{M+1}) \leq \alpha) \tag{12}$$

What is the meaning of this criterion? Equation (12) ensures that the probability that the reduction factor is smaller than a chosen  $\alpha$  is larger than the optimized value. For the rest of this paper, we shall use  $\alpha = 0.2$ , meaning that the reduction in degenerate space is at least by a factor of 5. As we shall see later, in many cases, the probability of such a reduction will be larger than 0.6.

Equation (12) can be used iteratively, where, in every cycle, the degeneracy of the previous step is used to deploy the new detector. A schematic description for such an algorithm is the following:

While  $N_d > toll$  :

1. Prediction: based on  $S(R_1 \dots R_M)$ , use Equation (12) to estimate  $R_{M+1}^*$ ;
2. Measurement: deploy the new detector at  $R_{M+1}^*$  and measure  $d_{M+1}$ ;
3. Reduction: use Equation (10) to construct the new degenerate space  $S(R_1 \dots R_M, R_{M+1}^*)$ ;
4. Calculate  $N_d = |S(R_1 \dots R_M, R_{M+1}^*)|$  and return to step 1.

where  $N_d$  is the number of degenerate points in every cycle,  $N_{toll}$  is the tolerance level, that is, the minimal number of points required for the termination of the algorithm, and  $\alpha$  should be specified by the user.

### 3. Results

#### 3.1. Numerical Details

Since our main interest in this section is the study of the proposed decision mechanism for the deployment of a new detector based on previous measurements, we shall focus on the relatively simple meteorological scenario and use the same dispersion model used for the source estimation to generate the input concentrations ('synthetic data'). In all examined cases, the external wind is 2 m/s in the east direction in neutral stability. Unless mentioned otherwise, the source is located at the center of a 10 km  $\times$  10 km grid composed by cells of 20 m  $\times$  20 m and the emission rate of the source is 1 kg/min. The Lagrangian stochastic model is used to calculate the concentration based on 400,000 Lagrangian particles. It is important to note that, in the context of these test cases, which focus on pollutant dispersion analysis for a period not exceeding one hour, we assume that the prevailing averaged wind and turbulent fluxes remain stationary ([59] [Chap. 2]). The LSM, however, is designed to model realistic environmental conditions, including non-homogeneous turbulence (see Section 2).

As mentioned in Section 2.2, the construction of the degenerate space at each cycle requires the specification of the error dependency on the concentration. In order to estimate this relation, a series of 30 identical LSM operations was performed, differing by their random seeds. Based on these calculations, the ratio between the standard deviation and the average concentration was calculated at seven typical points located downstream of the source position. This ratio, averaged over these points, is  $0.15 \pm 0.14$ , which implies that the proportional constant should be  $\kappa = 0.3$ . In addition, the fixed term in the error expression (see Equation (6)) was taken as  $\sigma_1 = 10^{-8}$ .

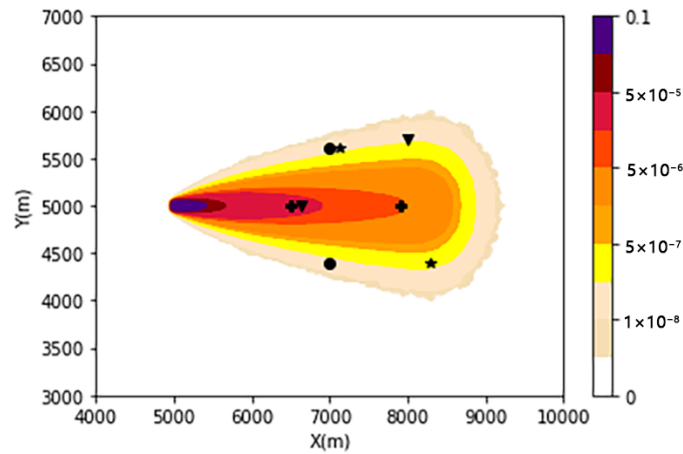
#### 3.2. Testing the Adaptive Source Algorithm

We shall examine the adaptive source term estimation algorithm, specified in Section 2.3, in six cases differing by the initial arrangement of the first two pre-deployed detectors and the source parameters. The concentration map generated by operating the Lagrangian stochastic model for the source located at (5000, 5000) is shown in Figure 1. A different marker is used to designate the pair of detectors of the first four cases. In the first case (plus marker), the first two detectors are pre-deployed along the same line parallel to the wind direction. In the second case (circle), the detector is aligned with the same line perpendicular to the external line. In cases 3 (star) and 4 (triangle), the symmetry is removed. The fifth case is similar to the fourth case but the source location is shifted to (4000, 5200). In the sixth case, the setup is identical to the third case, but the emission rate is 2 kg/min.

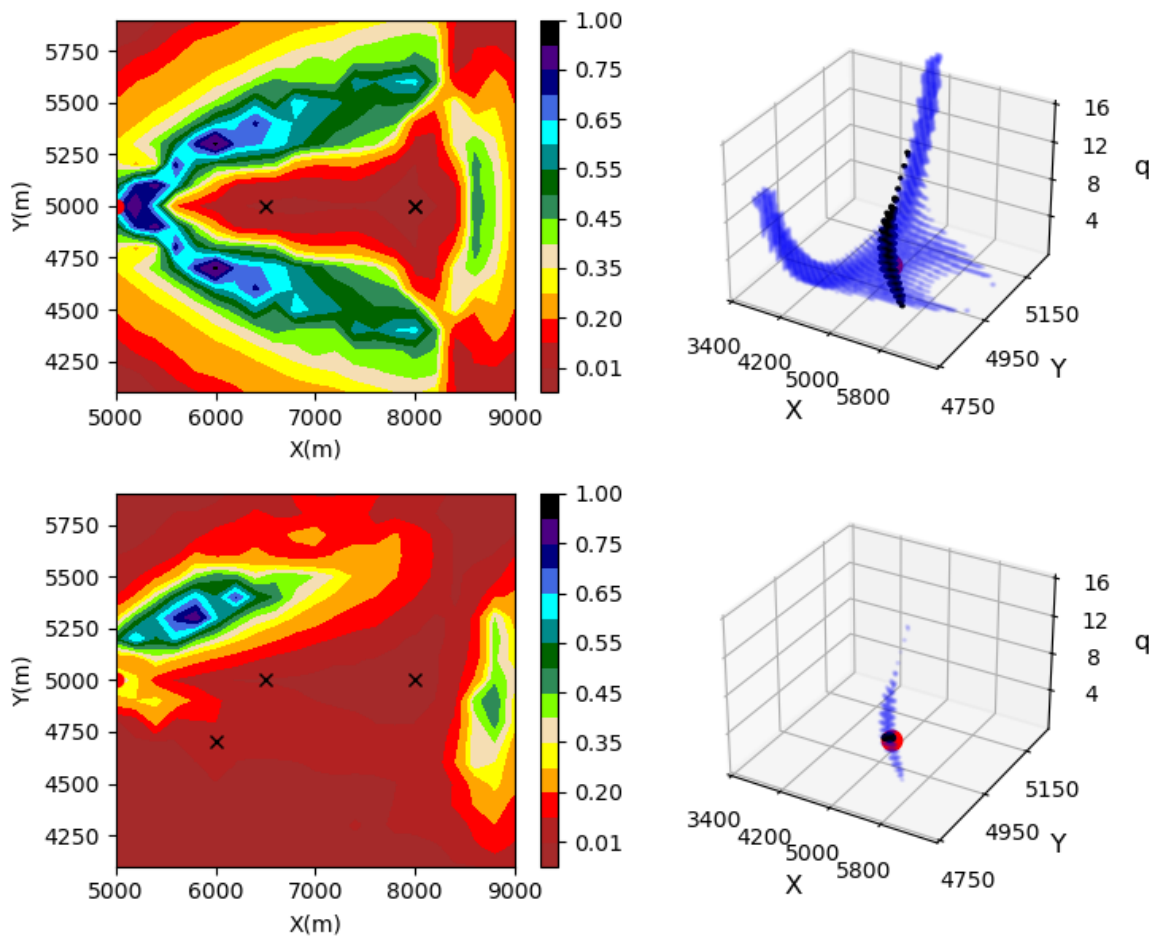
In all cases, the input concentration for the source estimation process is taken as the value of the map at the detector's positions. Full analysis for cases 1 and 4 can be seen in Figures 2 and 3 and shall now be described.

**Case 1:** In the upper-right panel of Figure 2, a map of  $Pr(\gamma(R_3) \leq 0.2)$  is shown based on the available measurements at the first two detectors located at  $R_1 = (6500, 5000)$ ,  $R_2 = (8000, 5000)$  (also shown as black crosses on the map). One can easily see two stretches of preferred points and a wide non-preferred regime enclosed by these two stretches. The position of the third detector is chosen at (6000, 4700), for which the criterion is maximized. The effect of this choice on the degeneracy reduction can be seen in the upper-right panel of Figure 2. The degenerate space before and after the third measurement can be seen in blue and black points, respectively. Note that the wide and spread-out

original space that contained 8704 points has been reduced to a narrow diagonal stretch containing only 460 points.



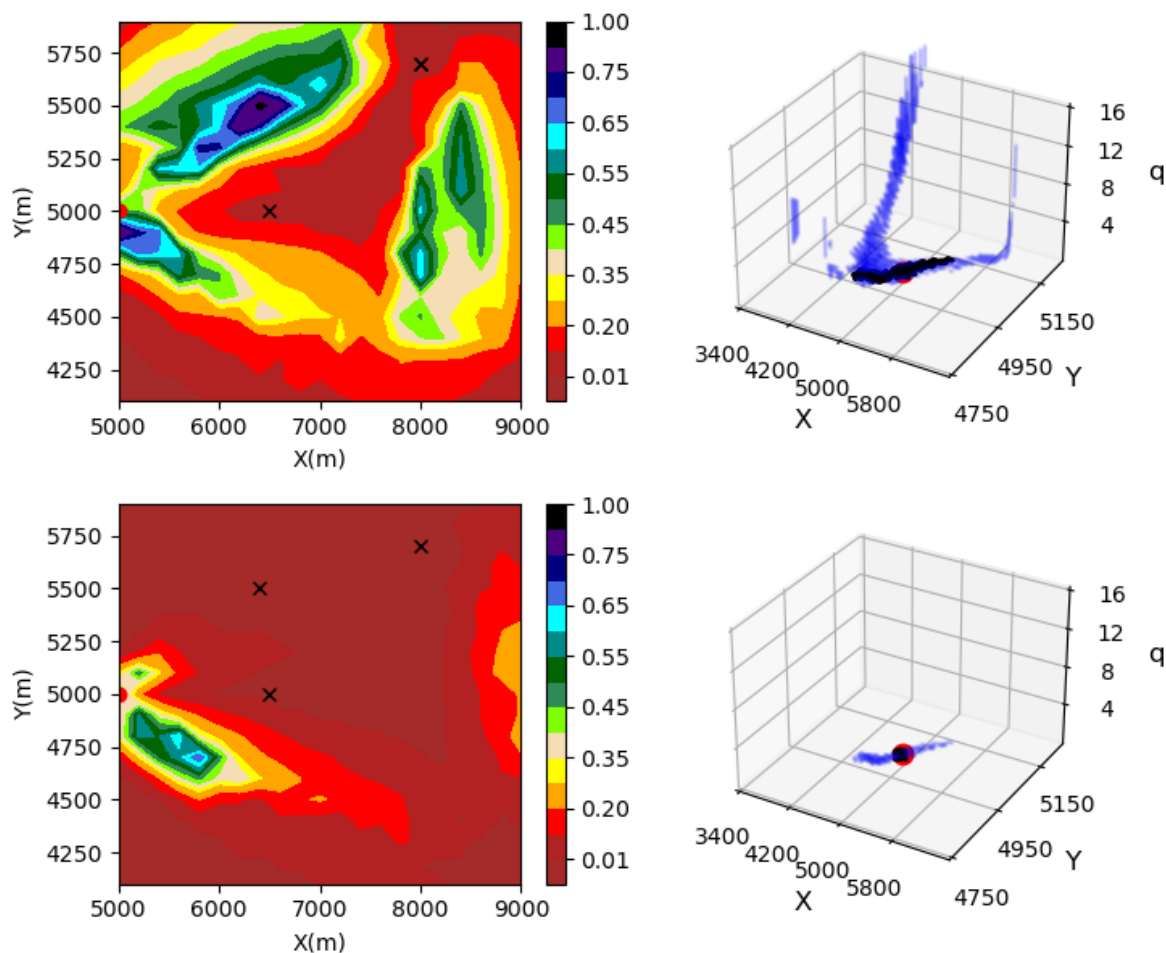
**Figure 1.** The concentration map (normalized by the source total emission) generated by a source located at (5000 m × 5000 m) is shown. The positions of the two detectors for the first four cases are also shown, each case represented by different markers (plus marker for case 1, circle for case 2, star for case 3 and triangle for case 4).



**Figure 2.** Case 1 analysis: each row describes a cycle of the algorithm, starting from two detectors located at (6500, 5000) and (8000, 5000). In the left side of each row, the expected reduction according to the criterion described in Equation (12) is shown for all possible locations of the new detector. The already deployed detectors, used for the construction of the degenerate space in the beginning of



the cycle, are designated by black crosses. The degenerate space before (blue) and after (black) the deployment of the new detector can be seen in the right panel of every row. The actual reduction gained by the procedure can be seen by the difference between the two colors. After two cycles, the degenerate space is reduced almost entirely to the correct source parameters, represented by the red dot.



**Figure 3.** Case 4 analysis: the same analysis as described in the previous figure, starting from two detectors located at (6500, 5000) and (8000, 5700).

The first row in Figure 2 describes the first cycle in the operation of the adaptive algorithm for case 1. The next row can be understood in a similar way—in the left panel, a map of the predicted reduction is formed based on the degenerate space of the previous iteration (note that a new black cross was added at the optimized position of the previous cycle), and a point that optimizes this map is chosen as the position of the new (fourth) detector. The effect of the new measurement can be shown in the right panel in the second row. As before, blue points represent the degenerate space before the addition of the new detector, which is the outcome of the last cycle, and black points represent the remaining points after the addition of the new information.

Note that the initial arrangement of the detectors along a line parallel to the wind direction led to a very large and spread-out initial degenerate space. Nevertheless, every addition of the detector dramatically reduces the size and shape of the degenerate space of the last cycle, so, after the deployment of two new detectors, the degenerate space is condensed to the small proximity of the source, represented by the red dot in the lower-right panel.

The performance of the adaptive algorithm along two cycles for the first case is summarized in the upper part of Table 1. For each cycle, the number of degenerate points,

as well as the averaged source parameters and their corresponding standard deviations, is shown (the actual source position and emission rate were subtracted from the averages). One can clearly see fast convergence to the correct value in all these criteria.

**Case 4:** In Figure 3, a similar analysis for the fourth case (triangle markers in Figure 1) is shown. The asymmetric deployment of the first two detectors dictates an asymmetric shape of the probability map. In addition, the regime of high probability is far more narrow and difficult to predict intuitively without a detailed calculation. In this case, the initial averages and standard deviations of all parameters are large, as can be seen in the first row of case 4 in Table 1. Adding new detectors dramatically reduces the error of all parameters, as can be seen in the following rows of the table, as well as by comparing the blue and black points in the figure.

A summary of all six cases is shown in Table 1. The number of cycles is limited by the demand that the number of degenerate points is lower than some small number (taken as 100 points in this work). One can see that, in all cases, a rapid decrease in the number of suspected points occurs after the first iteration (i.e., by adding the third detector). Furthermore, both the averaged parameters (after subtraction of the exact values) and the standard deviation of the degenerate space reduce along the procedure (the errors in the y direction are typically much smaller than the x direction because the wind is aligned along the x axis). Note that the search was conducted in a parameter space of 10,000 m × 10,000 m, so the relative error of every spatial parameter, normalized by the corresponding length scale, at the end of the algorithm operation is very small. More than that, the resolution of the source estimation is limited by the size of the grid cells used to predict the concentration from the LSM operation, which was 20 m × 20 m for the spatial coordinates and 0.15 for the emission rate in these cases. Therefore, in all cases, the errors are within a deviation of three spacial cells at the most. Note that since the wind direction is aligned along the x axis, the errors in this direction (both the average and standard deviation) are much larger than along the y axis.

**Table 1.** The results of the adaptive source estimation algorithm for six cases are shown in the table. For every case, the first row describes the features of degenerate space before applying the adaptive algorithm based on two pre-deployed detectors. In every row, the number of degenerate points ( $N_d$ ), the averages of the source parameters (after subtraction of the exact value) and their corresponding standard deviations are shown.

Cycle	$N_{detectors}$	$N_d$	$\langle x \rangle - x_{sor}$	$\langle y \rangle - y_{sor}$	$\langle q \rangle - q_{sor}$	$\sigma_x$	$\sigma_y$	$\sigma_q$
CASE 1								
	2	8704	231	0	3.76	400	210	4.24
1	3	460	130	32	0.68	296	72	0.92
2	4	32	32	0	0.12	45	0	0.19
CASE 2								
	2	2583	896	0	3.53	1062	22	3
1	3	235	87	0.6	0.75	394	11	1.2
2	4	12	45	0	0.18	27	0.0	0.12
CASE 3								
	2	2183	30	30	3.9	630	139	3.1
1	3	73	39	11	0.21	74	12	0.25
2	4	31	30	12	0.06	40	12	0.15
CASE 4								
	2	3762	451	87	3.5	604	171	3.9
1	3	441	56	24	0.35	186	53	0.45
2	4	43	13	5	0.12	60	8	0.18

**Table 1.** *Cont.*

Cycle	$N_{detectors}$	$N_d$	$\langle x \rangle - x_{sor}$	$\langle y \rangle - y_{sor}$	$\langle q \rangle - q_{sor}$	$\sigma_x$	$\sigma_y$	$\sigma_q$
CASE 5								
1	2	2743	428	2	3.5	771	102	3.22
2	3	245	120	4	0.21	188	19	0.25
3	4	24	70	5	0.08	94	13	0.14
CASE 6								
1	2	3060	2	50	5.3	575	117	3.3
2	3	158	49	10	0.46	78	12	0.49
3	4	63	28.5	11	0.19	39	12	0.32

### 4. Discussion

In this paper, we have presented a new approach for an adaptive deployment of detectors that aims to reduce the existence of many hypotheses with approximately the same probability for describing the source parameters. This adaptive approach relies on two concepts: the identification of the suspected points based on current measurements and the prediction of the effect of the new detector’s position on the degenerate space reduction.

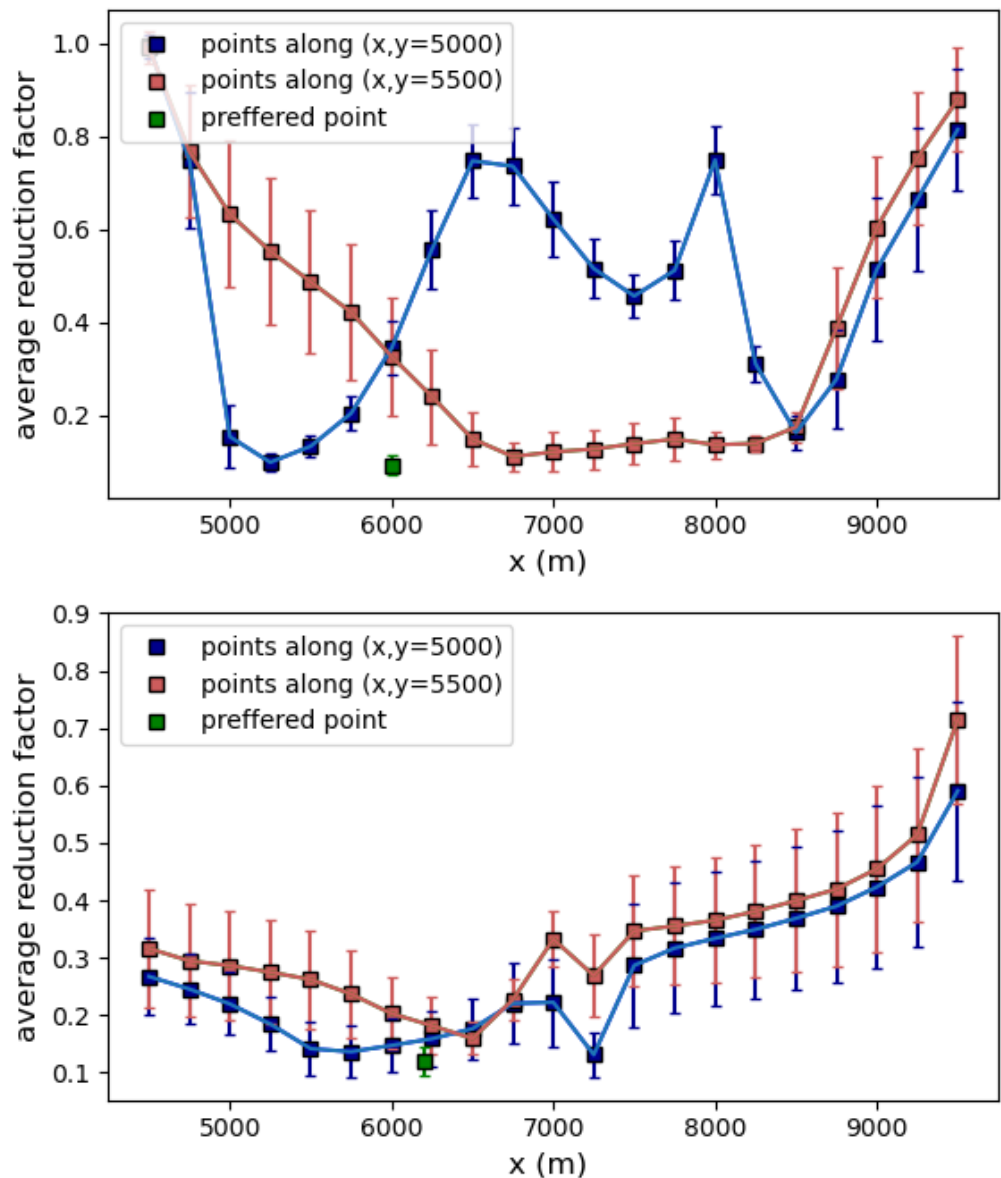
The construction of the degenerate space defined in Section 2.2 can be intuitively thought of as the intersection of the detectors’ “iso-surfaces”, which are the sets of all points in the parameter space that will retrieve the actually measured concentration in every detector. This idea is somewhat similar in its character to the approach presented by Keats [34], who demonstrated how the ‘regimes of influences’, generated by solving the adjoint diffusion–advection equation for every detector, can be used to select the possible locations of the source. There is, however, a significant difference between the two approaches, since the construction of these regimes does not rely on the concentrations that were actually measured, in contrast to the definition of the degenerate space given before. Furthermore, to the best of our knowledge, this idea was not implemented quantitatively for the BLSM in the form of Equation (10). The estimation of the effect of the new detector’s position on the degenerate space reduction, which is the heart of the adaptive algorithm, is novel and does not rely on Bayesian design techniques.

How should the adaptive method developed in this study be tested? The straightforward answer may be to calculate the actual reduction by deploying the new detector at the proposed location, measure the concentration and calculate the new degenerate space formed by this deployment. Then, this outcome could be compared with the actual reduction formed by the deployment of the new detector at other places. Note, however, that all the points in the degenerate space can serve as the source and each one of them will reproduce a similar degenerate space. If the starting point would be a different point taken from the degenerate space, applying the previously mentioned test may lead to a completely different reduction. Therefore, a better test for the performance of the proposed mechanism is to compare the actual obtained reduction as a function of the detector’s position, averaged over all equally probable hypotheses.

In the upper panel of Figure 4, the averaged actual reduction is shown for different possible positions for the third detector in the first case along two lines,  $(x, 5000)$  and  $(x, 5500)$  (an error bar was added to indicate the standard deviation associated with the average at every point). The average reduction for the preferred point, chosen by the criterion mentioned in Section 2.3, is also shown as a green rectangle. In the lower panel, the same analysis is shown for the second case. One can see that the best outcome is obtained for the proposed position in both cases. In the first case, the difference between the preferred and other points can be dramatic for most of the examined points (note that the averaged reduction for the preferred point is less than 0.1, i.e., on average, less the 10% of the degenerate points survived the elimination induced by the third measurement). Note that, in this case, because the first two detectors are positioned along the line parallel to the

wind direction, there is a preference for breaking the symmetry in the deployment of the third detector as can be seen by the sequence of red squares with low average reduction.

In the second case, the difference is less dramatic, although the reduction for the preferred point is still significantly lower than all other points (0.12). In addition, a similar trend between the two lines can be seen, induced by the symmetric deployment of the first two detectors along the axis perpendicular to the wind direction (see the circles in Figure 1, which represents the position of the detectors in the second case). Hence, a different behavior can be seen between the two cases: while, in the first case, choosing the preferred point is crucial, the second case is less sensitive to the position of the third detector (although the preferred point is still optimal). We should emphasize that the actual averaged reduction can be used only for analysis of the results, not as a design criterion.



**Figure 4.** The actual reduction, averaged over all possible hypotheses in the first cycle of the first and second cases, is shown for different points along the line  $(x, 5000)$  and  $(x, 5500)$  and for the preferred point (located at  $(6000, 5300)$  in the first case and at  $(6200, 4600)$  in the second case). Note that the best performance is achieved for the preferred point (green point).

In addition, we compared the reduction obtained by deploying the third detector at the proposed position to the reduction obtained by randomly choosing its position. The reduction averaged over 200 random choices taken from a box of  $[3000, 10,000] \times [4000, 6000]$  is 0.6, much larger than the reduction obtained for the proposal position (0.05).

The effect of the detector's number and its arrangement on source estimation accuracy was addressed by [60], who used the results from a wind tunnel experiment as input for gradient-based source estimation using a Gaussian dispersion model. In their work, 11 different sets of four detector configurations were used to estimate the source parameters in a similar setup as taken here (homogeneous external wind is 1.3 m/s). As can be seen in the first row of Table 1, in Rudd's paper, the error in the source position divided by a typical length scale, taken as the averaged distance between the source and detectors, is about 0.19 and 0.24 for four and five detectors, respectively. In order to compare our results, we normalized the error of the second cycle (after four detectors had been deployed) with the averaged distance ( $\sim 2000$  m), yielding an almost an order of magnitude smaller error in all six cases.

As mentioned in Section 2.2, the use of a BLSM instead of a standard (forward) LSM enables a linear dependency of the computational effort required for the construction of the degenerate space with the number of detectors. Furthermore, replacing the underlying dispersion model with a more complex and costlier model will not dramatically increase the computation effort required for the task (as long as an 'adjoint' dispersion model, analog to the BLSM, is available). The computational bottleneck of the proposed adaptive algorithm is the double loop (over all hypotheses in the parameter space) required for the calculation of Equation (12). Increasing the dimensionality of the parameter space will increase the number of hypotheses. Typically, every degree of freedom was discretized to  $N_i$  cells (500 cells for the special parameters and 100 for the emission rate in this work). So, the number of hypotheses is expected to grow as  $\prod_i^d N_i$ . Therefore, the scalability of the algorithm is expected to grow exponentially with the number of parameters  $d$ , limiting the applicability of the algorithm to a relatively small number of parameters. It should be noticed, however, that, in this work, a dense grid was used, which can be easily diluted under mild prior assumptions. For example, the width of the dispersion in the  $y$  direction in this case is much smaller than in the  $x$  direction as can be seen in Figure 1. Therefore, the relevant range in the  $y$  direction can be significantly reduced. In addition, the spatial cell size ( $20 \text{ m} \times 20 \text{ m}$  in this study) can be increased at the expense of resolution. Furthermore, the computation of each case was performed on a standard off-the-shelf single i-7 processor and took a few hours to complete. It should be noticed that the calculations of Equation (11) for different candidates ( $R_{M+1}$ ) are independent, such that parallel computing can be applied linearly with the processor number. To conclude, for reasonable computational resources ( $\sim 200$  processors), the algorithm can be used in real time for problems characterized by up to five source parameters.

One mandatory requirement for operating this procedure is prior knowledge of the noise functional form. If this form is unknown, then a calibrating procedure of the detector is required as a preliminary step. In this work, an "almost linear" form of the noise (see Equation (7)) was taken, which corresponds to the work of Keats and Yee [42]. Naturally, as  $\kappa$  and  $\sigma_0$  increase, more hypotheses will satisfy the condition described in Equation (9), so the degenerate space will become larger. Besides increasing the computational burden, this might require more cycles in order to reach the tolerance level of the algorithm. Furthermore, all the environmental conditions, such as the wind, turbulence parameters and stability state, should be supplied as an input before or during the operation of the procedure. Hence, for practical use, it is important to incorporate this algorithm in a wider meteorological framework.

## 5. Conclusions

The need for an adaptive source term estimation method is critical in real-life scenarios where the available information is expected to be small, and thus the deployment of a sparse array of static detectors will be inefficient. In this study, a novel method for an adaptive STE is presented. The method is based on a simple and fast identification of the relevant hypotheses based on a given set of measurements, and then exploiting these data to choose the next measurement position. The method was tested on six cases differing by the initial arrangement of the first two pre-deployed detectors, and was shown to retrieve the source location and emission rate accurately and efficiently.

The focus of this study was on presenting the new methodology and testing it on a small initial number of pre-deployed detectors for a relatively simple meteorological scenario. We should emphasize that, since the underlying dispersion model is an LSM, which is known for its ability to accurately describe pollutant dispersion in many complicated scenarios, it can be implemented for realistic scenarios. One future extension of this work is testing the performance of the algorithm in cases where the wind is non-stationary and heterogeneous and the source has a time-dependent release profile. A further comparison of this algorithm’s numerical performance with that of the Bayesian adaptive exploration method is currently being investigated in scenarios analogous to those analyzed in this study.

**Author Contributions:** Conceptualization, O.B. and E.F.; methodology, O.B. and E.F.; software, O.B.; validation, O.B.; formal analysis, O.B. and E.F.; investigation, O.B. and E.F.; data curation, O.B.; funding acquisition, E.F.; writing—original draft preparation, O.B. and E.F.; writing—review and editing, O.B. and E.F.; All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The data used for this study can be found at <https://doi.org/10.5061/dryad.4mw6m90fb> (accessed on 1 January 2025). The computer code used during the study is available from the corresponding author, O. Buchman, on request.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Appendix A. Derivation of the Probability of Reduction Factor Estimation

How can we predict the expected reduction as a function  $R_{M+1}$  of before the deployment of the new detector? If  $\theta_j = \theta_{ex}$  and the measurement outcome is  $d_{M+1}$ , then the new degenerate space will be

$$S(R_{M+1}|\theta_j) = \{\theta_k \in S(R_1 \dots R_M) \mid |d(R_{M+1}|\theta_k) - d_{M+1}| \leq \sigma(\theta_k)\}$$

and the conditioned reduction factor will be defined as

$$\Gamma(R_{M+1}|\theta_j) = \frac{|S(R_1 \dots R_{M+1}|\theta_j)|}{|S(R_1 \dots R_M)|} \tag{A1}$$

(we omitted the dependency of the reduction factor on the already deployed detectors ( $R_1 \dots R_M$ ) since their positions cannot be changed in contrast to the position of the new detector). As mentioned before, this is not useful since  $d_{M+1}$  can be known only after the deployment and measurement of the new detector at  $R_{M+1}$ . Therefore, define the readily attainable set:

$$s(R_{M+1}|\theta_j) = \{\theta_k \in S(R_1 \dots R_M) \mid |d(R_{M+1}|\theta_k) - d(R_{M+1}|\theta_j)| \leq \sigma(\theta_k) + \sigma(\theta_j)\}$$

and its corresponding conditioned reduction factor:

$$\gamma(R_{M+1}|\theta_j) = \frac{|s(R_1 \dots R_{M+1}|\theta_j)|}{|S(R_1 \dots R_M)|}. \tag{A2}$$

Now, let us relate  $\gamma(R_{M+1}|\theta_j)$  to  $\Gamma(R_{M+1}|\theta_j)$ . By definition, for every  $\theta_k \in S(R_{M+1}|\theta_j)$ ,

$$|d(R_{M+1}|\theta_k) - d_{M+1}| \leq \sigma(\theta_k).$$

In addition, since we assumed that  $\theta_j = \theta_{ex}$ , then, from (6),

$$|d(R_{M+1}|\theta_j) - d_{M+1}| \leq \sigma(\theta_j).$$

The combination of the last two inequalities leads to

$$|d(R_{M+1}|\theta_k) - d(R_{M+1}|\theta_j)| \leq |d(R_{M+1}|\theta_k) - d_{M+1}| + |d(R_{M+1}|\theta_j) - d_{M+1}| \leq \sigma(\theta_j) + \sigma(\theta_k).$$

So, we have shown that  $\theta_k \in s(R_{M+1}|\theta_j)$  and therefore  $S(R_{M+1}|\theta_j) \subseteq s(R_{M+1}|\theta_j)$ , which implies that, for every  $\theta_j$ ,

$$\Gamma(R_{M+1}|\theta_j) \leq \gamma(R_{M+1}|\theta_j). \tag{A3}$$

Now, consider the probability that, if we randomly choose a hypothesis from  $S(R_1 \dots R_M)$ , the reduction factor and its estimator will be smaller than some small number  $\alpha$ . Since all points are equally probable, this can be calculated as

$$Pr(\Gamma(R_{M+1}) \leq \alpha) = \frac{|\{\theta_j|\Gamma(R_{M+1}|\theta_j) \leq \alpha\}|}{|S(R_1 \dots R_M)|}, \tag{A4}$$

and its “companion”:

$$Pr(\gamma(R_{M+1}) \leq \alpha) = \frac{|\{\theta_j|\gamma(R_{M+1}|\theta_j) \leq \alpha\}|}{|S(R_1 \dots R_M)|}. \tag{A5}$$

From Equation (A3), we know that, for every  $\theta_j$  for which  $\gamma(R_{M+1}|\theta_j) \leq \alpha$ , we obtain  $\Gamma(R_{M+1}|\theta_j) \leq \gamma(R_{M+1}|\theta_j) \leq \alpha$ , so  $\{\theta_j|\Gamma(R_{M+1}|\theta_j) \leq \alpha\} \subseteq \{\theta_j|\gamma(R_{M+1}|\theta_j) \leq \alpha\}$ , and we have shown that

$$Pr(\Gamma(R_{M+1}) \leq \alpha) \geq Pr(\gamma(R_{M+1}) \leq \alpha). \tag{A6}$$

Clarification of the last claim can be seen by the following example. Assume that we have 100 points and we obtain  $Pr(\gamma \leq 0.2) = 0.6$ . Therefore, there are 60 points for which  $\gamma(R_{M+1}|\theta_j) \leq 0.2$  and 40 points for which  $\gamma(R_{M+1}|\theta_j) \geq 0.2$ . For each point from the first class, we know that  $\Gamma(R_{M+1}|\theta_j) \leq \gamma(R_{M+1}|\theta_j) \leq 0.2$  and therefore  $Pr(|\Gamma(R_{M+1}) \leq 0.2|) \geq 0.6$ .

## References

1. Fattal, E.; David-Saroussi, H.; Klausner, Z.; Buchman, O. An urban lagrangian stochastic dispersion model for simulating traffic particulate-matter concentration fields. *Atmosphere* **2021**, *12*, 580. [CrossRef]
2. Fattal, E.; David-Saroussi, H.; Buchman, O.; Tas, E.; Klausner, Z. Heterogenous Canopy in a Lagrangian-Stochastic Dispersion Model for Particulate Matter from Multiple Sources over the Haifa Bay Area. *Atmosphere* **2023**, *14*, 144. [CrossRef]
3. Meier-Girard, D.; Delgado-Eckert, E.; Schaffner, E.; Schindler, C.; Künzli, N.; Adam, M.; Pichot, V.; Kronenberg, F.; Imboden, M.; Frey, U.; et al. Association of long-term exposure to traffic-related PM10 with heart rate variability and heart rate dynamics in healthy subjects. *Environ. Int.* **2019**, *125*, 107–116. [CrossRef] [PubMed]

4. Khreis, H.; Kelly, C.; Tate, J.; Parslow, R.; Lucas, K.; Nieuwenhuijsen, M. Exposure to traffic-related air pollution and risk of development of childhood asthma: A systematic review and meta-analysis. *Environ. Int.* **2017**, *100*, 1–31. [[CrossRef](#)] [[PubMed](#)]
5. Eitan, O.; Barchana, M.; Dubnov, J.; Linn, S.; Carmel, Y.; Broday, D.M. Spatial analysis of air pollution and cancer incidence rates in Haifa Bay, Israel. *Sci. Total Environ.* **2010**, *408*, 4429–4439. [[CrossRef](#)]
6. Sasaki, K.; Sakamoto, K. Vertical differences in the composition of PM10 and PM2.5 in the urban atmosphere of Osaka, Japan. *Atmos. Environ.* **2005**, *39*, 7240–7250. [[CrossRef](#)]
7. Zhang, Y.F.; Xu, H.; Tian, Y.Z.; Shi, G.L.; Zeng, F.; Wu, J.H.; Zhang, X.Y.; Li, X.; Zhu, T.; Feng, Y.C. The study on vertical variability of PM10 and the possible sources on a 220 m tower, in Tianjin, China. *Atmos. Environ.* **2011**, *45*, 6133–6140. [[CrossRef](#)]
8. Chan, C.; Xu, X.; Li, Y.S.; Wong, K.; Ding, G.; Chan, L.; Cheng, X. Characteristics of vertical profiles and sources of PM2.5, PM10 and carbonaceous species in Beijing. *Atmos. Environ.* **2005**, *39*, 5113–5124. [[CrossRef](#)]
9. Manousakas, M.; Bairachtari, K.; Kantarelou, V.; Eleftheriadis, K.; Vasilakos, C.; Assimakopoulos, V.; Maggos, T. The traffic signature on the vertical PM profile: Environmental and health risks within an urban roadside environment. *Sci. Total Environ.* **2019**, *646*, 448–459.
10. Roth, M. Review of atmospheric turbulence over cities. *Q. J. R. Meteorol. Soc.* **2000**, *126*, 941–990. [[CrossRef](#)]
11. Finnigan, J. Turbulence in plant canopies. *Annu. Rev. Fluid Mech.* **2000**, *32*, 519–571. [[CrossRef](#)]
12. Raupach, M.R.; Finnigan, J.J.; Brunet, Y. Coherent eddies and turbulence in vegetation canopies: The mixing-layer analogy. In *Boundary-Layer Meteorology*; In 25th Anniversary Volume, 1970–1995: Invited Reviews and Selected Contributions to Recognise Ted Munn’s Contribution as Editor over the Past 25 Years; Springer: Dordrecht, The Netherlands, 1996; pp. 351–382.
13. Castro, I.P.; Cheng, H.; Reynolds, R. Turbulence over urban-type roughness: Deductions from wind-tunnel measurements. *Bound.-Layer Meteorol.* **2006**, *118*, 109–131. [[CrossRef](#)]
14. Coceal, O.; Dobre, A.; Thomas, T.G.; Belcher, S.E. Structure of turbulent flow over regular arrays of cubical roughness. *J. Fluid Mech.* **2007**, *589*, 375–409. [[CrossRef](#)]
15. Hanna, S.R.; Brown, M.J.; Camelli, F.E.; Chan, S.T.; Coirier, W.J.; Hansen, O.R.; Huber, A.H.; Kim, S.; Reynolds, R.M. Detailed simulations of atmospheric flow and dispersion in downtown Manhattan: An application of five computational fluid dynamics models. *Bull. Am. Meteorol. Soc.* **2006**, *87*, 1713–1726. [[CrossRef](#)]
16. Wilson, J.; Yee, E.; Ek, N.; d’Amours, R. Lagrangian simulation of wind transport in the urban environment. *Q. J. R. Meteorol. Soc.* **2009**, *135*, 1586–1602. [[CrossRef](#)]
17. Wang, C.; Li, Q.; Wang, Z.H. Quantifying the impact of urban trees on passive pollutant dispersion using a coupled large-eddy simulation–Lagrangian stochastic model. *Build. Environ.* **2018**, *145*, 33–49. [[CrossRef](#)]
18. Finnigan, J.J.; Shaw, R.H.; Patton, E.G. Turbulence structure above a vegetation canopy. *J. Fluid Mech.* **2009**, *637*, 387–424. [[CrossRef](#)]
19. Christen, A.; Rotach, M.W.; Vogt, R. The budget of turbulent kinetic energy in the urban roughness sublayer. *Bound.-Layer Meteorol.* **2009**, *131*, 193–222. [[CrossRef](#)]
20. Shig, L.; Babin, V.; Shnapp, R.; Fattal, E.; Liberzon, A.; Raviv, Y. Quadrant Analysis of the Reynolds Shear Stress in a Two-Height Canopy. *Flow Turbul. Combust.* **2023**, *111*, 35–57. [[CrossRef](#)]
21. Shnapp, R.; Bohbot-Raviv, Y.; Liberzon, A.; Fattal, E. Turbulence-obstacle interactions in the Lagrangian framework: Applications for stochastic modeling in canopy flows. *Phys. Rev. Fluids* **2020**, *5*, 094601. [[CrossRef](#)]
22. Shnapp, R. On small-scale and large-scale intermittency of Lagrangian statistics in canopy flow. *J. Fluid Mech.* **2021**, *913*, R2. [[CrossRef](#)]
23. Shnapp, R.; Liberzon, A.; Raviv-Bohbot, Y.; Fattal, E. On local isotropy and scale dependence of pair dispersion in turbulent canopy flows. *J. Fluid Mech.* **2024**, *978*, A3. [[CrossRef](#)]
24. Charuchittipan, D.; Wilson, J. Turbulent kinetic energy dissipation in the surface layer. *Bound.-Layer Meteorol.* **2009**, *132*, 193–204. [[CrossRef](#)]
25. Wilson, J. Monin-Obukhov functions for standard deviations of velocity. *Bound.-Layer Meteorol.* **2008**, *129*, 353–369. [[CrossRef](#)]
26. Klausner, Z.; Ben-Efraim, M.; Arav, Y.; Tas, E.; Fattal, E. The Micrometeorology of the Haifa Bay Area and Mount Carmel during the summer. *Atmosphere* **2021**, *12*, 354. [[CrossRef](#)]
27. Klausner, Z.; Kaplan, H.; Fattal, E. The Similar Days Method for Predicting Near Surface Wind Vectors. *Meteorol. Appl.* **2009**, *16*, 569–579. [[CrossRef](#)]
28. Klausner, Z.; Fattal, E. An objective and automatic method for identification of pattern changes in wind direction time series. *Int. J. Climatol.* **2011**, *31*, 783–790. [[CrossRef](#)]
29. Wilson, J.; Thurtell, G.; Kidd, G.; Beauchamp, E. Estimation of the rate of gaseous mass transfer from a surface source plot to the atmosphere. *Atmos. Environ.* **1982**, *16*, 1861–1867. [[CrossRef](#)]
30. Oxoli, D.; Gianquintieri, L.; Borghi, F.; Fanti, G.; Spinazzè, A. Non-Conventional Data for Farming-Related Air Pollution: Contributions to Modelling and Risk Assessment in the Lombardy Region, Italy. *Environments* **2024**, *11*, 229. [[CrossRef](#)]



31. Gallego, E.; Perales, J.F.; Aguasca, N.; Domínguez, R. Determination of emission factors from a landfill through an inverse methodology: Experimental determination of ambient air concentrations and use of numerical modelling. *Environ. Pollut.* **2024**, *351*, 124047 [[CrossRef](#)] [[PubMed](#)]
32. Hutchinson, M.; Oh, H.; Chen, W.H. A review of source term estimation methods for atmospheric dispersion events using static or mobile sensors. *Inf. Fusion* **2017**, *36*, 130–148. [[CrossRef](#)]
33. Issartel, J. Rebuilding sources of linear tracers after atmospheric concentration measurements. *Atmos. Chem. Phys.* **2003**, *3*, 2111–2125. [[CrossRef](#)]
34. Keats, A.; Yee, E.; Lien, F.S. Bayesian inference for source determination with applications to a complex urban environment. *Atmos. Environ.* **2007**, *41*, 465–479. [[CrossRef](#)]
35. Johannesson, G.; Hanley, B.; Nitao, J. *Dynamic Bayesian Models via Monte Carlo-An Introduction with Examples*; Technical Report; Lawrence Livermore National Lab.(LLNL): Livermore, CA, USA, 2004.
36. Yee, E.; Biloft, C.A. Concentration fluctuation measurements in a plume dispersing through a regular array of obstacles. *Bound.-Layer Meteorol.* **2004**, *111*, 363–415. [[CrossRef](#)]
37. Ko, C.W.; Lee, J.; Queyranne, M. An Exact Algorithm for Maximum Entropy Sampling. *Oper. Res.* **1995**, *43*, 684–691. [[CrossRef](#)]
38. Vergassola, M.; Villermaux, E.; Shraiman, B.I. 'Infotaxis' as a strategy for searching without gradients. *Nature* **2007**, *445*, 406–409. [[CrossRef](#)]
39. Hutchinson, M.; Oh, H.; Chen, W.H. Entotaxis as a strategy for autonomous search and source reconstruction in turbulent conditions. *Inf. Fusion* **2018**, *42*, 179–189. [[CrossRef](#)]
40. Sebastiani, P.; Wynn, H.P. Maximum entropy sampling and optimal Bayesian experimental design. *J. R. Stat. Society. Ser. B Stat. Methodol.* **2000**, *62*, 145–157. [[CrossRef](#)]
41. Ristic, B.; Skvortsov, A.; Gunatilaka, A. A study of cognitive strategies for an autonomous search. *Inf. Fusion* **2016**, *28*, 1–9. [[CrossRef](#)]
42. Keats, A.; Yee, E.; Lien, F.S. Information-driven receptor placement for contaminant source determination. *Environ. Model. Softw.* **2010**, *25*, 1000–1013. [[CrossRef](#)]
43. Lindley, D.V. On a Measure of the Information Provided by an Experiment. *Ann. Math. Stat.* **1956**, *27*, 986–1005. [[CrossRef](#)]
44. Loredo, T.J. Bayesian Adaptive Exploration. In *AIP Conference Proceedings 707*; Cornell University: Ithaca, NY, USA, 2004; pp. 330–346. [[CrossRef](#)]
45. Fattal, E. A non-homogeneous non-Gaussian Lagrangian-stochastic model for pollutant dispersion in complex terrain, and its comparison to Haifa 2009 tracer campaign. *IIBR Sci. Rep.* **2014**, *56*, 5614. [[CrossRef](#)]
46. Wilson, J.D.; Sawford, B.L. Review of Lagrangian stochastic models for trajectories in the turbulent atmosphere. *Bound.-Layer Meteorol.* **1996**, *78*, 191–210. [[CrossRef](#)]
47. Cassiani, M.; Stohl, A.; Brioude, J. Lagrangian stochastic modelling of dispersion in the convective boundary layer with skewed turbulence conditions and a vertical density gradient: Formulation and implementation in the FLEXPART model. *Bound.-Layer Meteorol.* **2015**, *154*, 367–390. [[CrossRef](#)]
48. Mooney, C.J.; Wilson, J.D. Disagreements between gradient-diffusion and Lagrangian stochastic dispersion models, even for sources near the ground. *Bound.-Layer Meteorol.* **1993**, *64*, 291–296. [[CrossRef](#)]
49. Weil, J.C. Linking a Lagrangian particle dispersion model with three-dimensional Eulerian wind field models. *J. Appl. Meteorol. Climatol.* **2008**, *47*, 2463–2467. [[CrossRef](#)]
50. Thomson, D.J. Criteria for the selection of stochastic models of particle trajectories in turbulent flows. *J. Fluid Mech.* **1987**, *180*, 529–556. [[CrossRef](#)]
51. Pope, S.B. *Turbulent Flows*; Cambridge University Press: Cambridge, UK, 2000. [[CrossRef](#)]
52. Obukhov, A.M. Description of Turbulence in Terms of Lagrangian Variables. *Adv. Geophys.* **1959**, *6*, 113–116.
53. Flesch, T.K.; Wilson, J.D.; Yee, E. Backward-time Lagrangian stochastic dispersion models and their application to estimate gaseous emissions. *J. Appl. Meteorol. Climatol.* **1995**, *34*, 1320–1332. [[CrossRef](#)]
54. Flesch, T.K.; Wilson, J.D.; Harper, L.A.; Crenna, B.P.; Sharpe, R.R. Deducing ground-to-air emissions from observed trace gas concentrations: A field trial. *J. Appl. Meteorol.* **2004**, *43*, 487–502. [[CrossRef](#)]
55. Sawford, B.L. Lagrangian statistical simulation of concentration mean and fluctuation fields. *J. Clim. Appl. Meteorol.* **1985**, *24*, 1152–1166. [[CrossRef](#)]
56. Issartel, J.P. Emergence of a tracer source from air concentration measurements, a new strategy for linear assimilation. *Atmos. Chem. Phys.* **2005**, *5*, 249–273. [[CrossRef](#)]
57. Keats, W.A. *Bayesian Inference for Source Determination in the Atmospheric Environment*; University of Waterloo: Waterloo, ON, Canada, 2009.
58. Kumar, P.; Feiz, A.A.; Singh, S.K.; Ngae, P.; Turbelin, G. Reconstruction of an atmospheric tracer source in an urban-like environment. *J. Geophys. Res.* **2015**, *120*, 12589–12604. [[CrossRef](#)]

- 
59. Stull, R.B. *An Introduction to Boundary Layer Meteorology*; Springer: Berlin/Heidelberg, Germany, 1988; p. 684.
  60. Rudd, A.C.; Robins, A.G.; Lepley, J.J.; Belcher, S.E. An Inverse Method for Determining Source Characteristics for Emergency Response Applications. *Bound.-Layer Meteorol.* **2012**, *144*, 1–20. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.