Development of methods for adaptive and optimal deployment of heterogeneous sensor array for accurate representation of air pollution and its sources

פיתוח שיטות לפריסה אופטימלית ואדפטיבית של מערך חיישנים הטרוגני לייצוג מדויק של זיהום אוויר ומקורותיו

Submitted by:

Idit Belachsen

Faculty of Civil and Environmental Engineering Technion-Israel Institute of Technology

31.07.2019

Advisors: Associate Professor Barak Fishbain, Associate Professor Shai Kendler

Abstract

Adequate monitoring of ambient air pollution is needed in order to prevent population exposure to hazardous concentrations of harmful materials, for regulation on industrial activity and for improved urban planning and development. The major challenge in these tasks is producing high resolution pollution concentration maps, using the current monitoring methods – standard air quality monitoring (AQM) and short-term measurements campaigns. The two methods have complementary benefits - AQM stations provide consistent but sparse high-quality data while short-term measurement campaigns achieve higher spatial resolution, but for limited time only. An alternative or complementary solution is to use Wireless Distributed Environmental Sensor Network (WDESN), usually comprised of portable and relatively low-cost Micro Sensing Units (MSUs), that can measure, process and transmit data to some base station. The use of a WDESN enables covering larger areas and obtaining finer spatial and temporal resolution of measurements than the standard methodology, but certainly poses a new challenge – finding the optimal way to deploy it while keeping operational costs low. The optimal deployment problem becomes even more challenging when the time varying weather and emission conditions are considered, and the optimal layout varies accordingly. The following proposal offers to investigate the theoretical and practical aspects of the deployment of WDESN in a site with complex sources' configuration and changing weather conditions, through a multi-objective optimization model, coupled with an air pollution dispersion model. The basic model aims at finding an algorithm for placing MSUs, under a set of fiscal and geographical constraints, that will allow decision-makers to discover tradeoffs between performance criteria and to consider alternative modes of action according to the prevailing circumstances. It considers the time varying meteorological conditions (wind velocity and atmospheric stability) and uses as the main objective a quantitative measure of the complexity of the given set of sources/sensors geometries: the pairwise Euclidean distance (PED) between the sensor network potential readings of different number of active sources. The task in hand is very complex and in order to accomplish it successfully, three modules are suggested: i) a *spatial* optimization model that computes sensors' deployment, which minimizes the cost of the deployment while maximizing the sensitivity of the network to changes in emissions, ii) a *temporal* optimization model that computes sensors' redeployment, which minimizes the transfer effort from the previous deployment, considering a change in weather or emission conditions, and iii) a *spatial-temporal* optimization model comprised of the findings of the two previous models, considering the probability of change in weather conditions and emission patterns. An initial simulation setup was constructed to test the first suggested model. Optimal solutions for placing up to 300 homogeneous MSUs in a 1000x1000 m flat area were retrieved for various PED-related objective functions, showing promising results. More simulations will be conducted for each of the modules, and the potential improvement achieved by the deployment of a heterogenous network, comprised of different types of sensors with different characteristics (sensitivity, dynamic range, cost), will be examined. Other extensions will be considered as well, including the use of other more sophisticated dispersion models, other forms of objective functions related to the maximization of the PED values, and real-world test cases.

Table of Contents

1 Introduction

1.1 Air pollution

Degraded air quality is a consequence of heightened emissions from a wide range of pollution sources, both anthropogenic, such as transportation or industry, and natural, such as dust plumes or wildfires. Dense urban centers and industrialized regions serve as "hotspots" in a continuum of concentrations, dictated by the rate by which all emitted substances undergo transport, diffusion, chemical transformation and deposition to the ground [1]. Fossil fuel combustion processes that govern todays' industrial and transportation activities are major emitters of gaseous pollutants to the troposphere, mainly nitrogen oxides, NoX, i.e., nitrogen monoxide, NO, and dioxide, NO_2; carbon monoxide (CO) and dioxide (CO_2); volatile organic compounds (VOCs) and others. Particulate pollutants, known as particulate matter (PM), are emitted as well in combustion processes and can be divided to primary particles (such as black carbon (BC)) and secondary particles which are formed in the atmosphere by oxidation of primary gaseous pollutants. Another secondary pollutant is ozone (O_3), produced naturally in the troposphere by photochemical oxidation of primary pollutants [2]. Many other anthropogenic pollutants are emitted from various industrial processes, among them are ammonia, methane and fluorinated gases, emitted for example from fertilizers' industry, agricultural practices or waste decay, and aluminum or semiconductor manufacturing, respectively [3]. Emissions are commonly separated to two types: point source emissions - channeled through a pipe, such as a chimney stack or a vent, and non-point source emissions – caused by direct contact of volatile substances or PM with the environment, where tracing pollution back to a single source is difficult [4].

The main incentive to reduce and control emission rates from anthropogenic sources is of course assuring population health. World health organization (WHO) estimates that 4.2 million premature deaths every year globally are linked to ambient air pollution, mainly from heart disease, stroke, chronic obstructive pulmonary disease, lung cancer, and acute respiratory infections in children [5]. Pollutants with the strongest evidence for public health concern include fine and ultrafine particulate matter (less than 2.5 μm and 0.1 μm in diameter, respectively), ozone, nitrogen dioxide and sulfur dioxide. Another not less important aspect is the affect air pollution has on climate. Climate change is driven by air pollution, as many pollutants disturb the steady state condition in earth's energy balance when interacting with solar and terrestrial radiation, leading to changes in earth's surface temperature and in climate [6], [7].

1.2 Monitoring air pollution

Monitoring air pollution is therefore necessary and may serve for preventing population exposure by detecting pollution peaks, for urban planning and development, for climate research or for public health studies that try to assess past and present population exposure to air pollution and correlate the level of exposure to observed health effects [8]. Regulatory authorities monitor air pollution to enforce plants that temporarily exceed their emission permit, reveal leaks or new

unknown sources. For all these purposes, the major challenge is in producing accurate pollution concentration maps of high spatial and temporal resolution. In epidemiological studies for example, it can enable finding correlations at a personal level (e.g. estimating concentrations in the exact place of residence of a subject) or help in finding health impacts triggered by short-term exceedances of pollution concentrations.

Ambient air pollution concentrations are usually obtained by two methods: i) routine measurements reported by standard air quality monitoring (AQM) stations [9] and ii) short-term measurement campaigns which usually utilize large number of sensors [10]–[12]. Data obtained by the first method are considered very accurate, since AQM stations are equipped with pollutant-designated measuring devices of high quality, that are maintained and calibrated on an ordinary basis by regulatory authorities. However, these tend to suffer from a few apparent flaws: i) their deployment is usually sparse, and cannot represent well the spatial and temporal variability of a typical pollutant; ii) samples of air are taken a few meters above street level and hence cannot represent well the extent of exposure of a passerby, if required, and iii) they are costly to maintain. Short-term measurement campaigns on the other hand can provide higher spatial resolution of the desired region of research, but for a limited time only. Seasonality or any phenomenon having longer cycles than the time of the campaign, will not be accounted for when using the measured data.

An alternative or complementary solution is to use Wireless Distributed Environmental Sensor Network (WDESN), usually comprised of portable and relatively low-cost Micro Sensing Units (MSUs), that can measure, process and transmit data to some base station. Operational costs of WDESNs are much lower than of AQM stations and they can be deployed in various sites, in large numbers and with far less operational requirements. These properties enable covering larger areas and obtaining finer spatial and temporal resolution of measurements than the standard methodology. Nevertheless, the reliability of MSUs' measurements is still questionable. Recent laboratory and field evaluations of MSUs show that these units are less accurate than standard laboratory equipment or AQM stations. However, it also has been shown that they are able to effectively capture the spatial and temporal variability of the pollution [10], [11], [13]–[15].

1.3 Air pollution representation in time and in space

Spatial interpolation methods (Kriging, inverse distance weighting (IDW) or spline) may help overcome sparse representation of concentration and can be used to produce pollution dense maps of certain locations and times [16]. Land Use Regression (LUR) models can link environmental variables (e.g., road type, traffic volume, topography, land cover) with measurements in monitoring locations and then use these variables as model predictors at unmonitored locations, or in time [17].

Unlike the above-mentioned models, atmospheric transport and dispersion (ATD) models quantify the deterministic relationships between sources' emissions and concentrations in space

and in time. ATD modeling refers to the mathematical description of pollutant transport in the atmosphere. The term dispersion is comprised of diffusion (due to turbulent eddy motion) and advection (due to wind) that occurs within the air near the Earth's surface [18]. ATD models can forecast the spread of the pollutants, when available results from monitoring data are used for calibration and evaluation of model performance. Oftentimes, these models serve as a useful tool for regulatory authorities to assess baseline ambient concentrations, analyze the relative importance of various emission sources, or test emission reduction strategies [19]. Regardless of the method used for generating the dense pollution maps, they all rely on the input they receive from the sensors. Thus, sub-optimal deployment on the sensors hinders air-pollution accurate description.

1.4   Source term estimation

For an accurate forecast of pollutants' spread, several variables are needed as input to ATD models, including, among others: meteorological data, the quantities of the pollutants emitted and the emissions' locations. While meteorological data of reasonable resolution (of a few square kilometers and 15-30 minutes) are usually available from local weather stations or weather models, the quantity, time or even location of pollutants' releases are frequently unknown. Pollutant release and transfer register (PRTR), an inventory of pollutants released to the environment, requires reporting of emissions from various industrial sectors to the regulator, on a yearly basis. However, their reliability is often questionable, not necessarily because of any bias, but due to inherent difficulties in the application of measurement techniques [20]–[22].

As a consequence, sensor measurements are often being used to determine source parameters in an inverse modelling problem of source-term estimation (STE) [16], [23]. In these methods, source parameters serve as input to an ATD model and are modified until the difference between the calculated and observed measurements is minimal. Usually, an optimization technique is applied in order to search the solution space and reach an accurate evaluation. The estimation of source parameters can also serve as an objective of its own, rather than generating high resolution concentration maps; in the case of a chemical attack or an industrial leak, locating the position and the leak rate of the cause is the primarily step before taking any action. Similarly to the sparse representation problem, the WDESN's input has a dramatical impact on the accuracy of STE methodologies, and thus, sensors' location allocation [24] problem is a major factor that must be addressed.

1.5   Network deployment

The above given background describes some general aspects related to air pollution, while emphasizing the major issues found in the field of monitoring, including - the sparse representation in space of pollution concentrations achieved by todays' commonly used monitoring methods, the low availability of data on pollutant release rates from point sources which requires source term estimation, or the high cost of maintaining a monitoring network. As

shortly specified, the use of a WDESN may help overcome some of these challenges, but evidently poses a new one – finding the optimal way to deploy it.

When resources are unlimited, the challenge of representing the pollution field well or sufficiently estimating source parameters, becomes an easier task through a deployment of an extensive number of sensors to reach a full cover of the area. However, as financial and human resources are usually limited, a smart and rigorous deployment of a network of sensors is needed, one which can provide the best performance for its designated application with the lowest possible cost.

The cost of a network is usually determined by the number of sensors and their attributes, such as sensitivity and dynamic range. Placement locations need to be considered as well, as the deployment of sensors close to a leak may be costly due to additional operational costs to cope with the expected severe environment. On the other hand, it can enable using less-sensitive sensors, which are usually cheaper. It should also be noted that in many industrial sites, sensor locations may be limited due to economic and practical considerations.

A static deployment approach, that satisfies the steady-state average conditions, might work well when no changes take place in the sources, emissions and meteorological conditions. Nonetheless, as meteorological conditions and sources' characteristics might change over different time-scales, redeployment strategies of a WDESN should be considered. Redeployment techniques allow reorganizing the WDESN by relocating some or all of the sensors, in order to achieve better performance at a minimal effort, according to predefined objectives.

The following proposal offers to explore and investigate the theoretical and practical aspects of multi-objective optimization models for the deployment of WDESN in time and space, aimed at finding a protocol for placing MSUs, under a set of fiscal and geographical constraints, so the best balance between various objectives is obtained. The offered tool can serve decision-makers when either establishing a new network, when performing short-term routine samplings for emission increase detection, or when alternative modes of action should be considered due to sudden changes in plans. It considers the time varying meteorological conditions (wind velocity and atmospheric stability) and uses as an objective a quantitative measure of the complexity of the given set of sources/sensors geometries, making it possible to evaluate "scene complexity", i.e. the potential to separate overlapping pollution plumes, once sensors are deployed at specific locations.

2   Research hypothesis and objectives

2.1   Research hypothesis

An optimal deployment of a WDESN is required for the purpose of obtaining early warnings when an increase in industrial gas emission occurs, as well as for generating pollution dense maps. A set of optimal deployment solutions can be retrieved using multi-objective optimization techniques that balance between detection capabilities of the network and operational requirements. The suggested solutions will allow decision-makers to discover tradeoffs between performance criteria and to consider alternative modes of action that may not have been apparent prior to the analysis.

2.2   Research objectives

The general objective of the proposed study is to design and develop a decision-making tool for an optimal deployment in time and in space of heterogenous sensor array, so the best balance between various objectives is obtained.

The specific objectives are:

i) Design an optimization model, comprised of the following modules:

1. A *spatial* optimization model that computes sensors' deployment, which minimizes the cost of the deployment while maximizing the sensitivity of the network to changes in the source term.

2. A *temporal* optimization model that computes sensors' redeployment, which minimizes the transfer effort from the previous deployment, considering a change in weather or emission conditions

3. A *spatial-temporal* optimization model comprised of the findings of the two previous models, considering probability of change in weather and emission conditions.

ii) Conduct a set of simulations for each of the modules to derive engineering insights for effective deployments of air pollution sensors at various emission rates, sources/sensors configurations and weather conditions.

iii) Examine the potential improvement achieved by the deployment of a heterogenous network, comprised of different types of sensors with different characteristics (sensitivity, dynamic range or specificity), compared to a homogeneous network of sensors.

iv) Examine at least one case-study of real-world network deployment and analyze its performance, compared to a suggested optimal solution.

3   Research contribution

Technological developments in recent years have made WDESNs feasible to deploy, in a relatively low operational cost, using portable MSUs. These units can easily transmit data directly to cloud-based platforms, making it possible to immediately obtain the latest mapping of the pollution level in the environment. These advancements have produced new challenges in the field of environmental air pollution monitoring and modelling, such as selecting optimal sensors' locations of placement.

This proposed work will provide mathematical and engineering tools which facilitate effective sensor deployment schemes that successfully balance between detection and operational requirements. The work will allow decision makers to examine tradeoffs between performance criteria of the multi-objective optimization model and to consider alternative modes of action according to the prevailing circumstances.

This proposed work's innovativeness stems from the implementation of a multi-objective optimization model, allowing a different perspective on the way operational problems may be approached. In addition, it firstly offers an adaptive mechanism, which considers the time-varying pollution and meteorological conditions that characterize the region of study, and thus produces robust and reliable solutions.

4   Literature review

4.1   The problem of sensors deployment

The problem of optimal deployment of any type of sensors consists of determining sensors' positions, while ensuring the coverage and the connectivity of the network [25]. By optimizing coverage, the deployment strategy would guarantee that the area of the sensing field is optimally covered by sensors, as required by the underlying application. By ensuring that the network is connected, it is also ensured that the sensed information is transmitted to other nodes and possibly to a centralized base-station that can make some valuable decision [26]. Zhang and Liu [27] added to these challenges prolonging the network lifetime, balancing the load to save energy and improving the accuracy of the transmitted data. However, some argue that when dealing with a WDESN, challenges can be reduced to sensor's type suitability (see Section 4.3), coverage area and data accuracy only, since most MSUs in use today are independent units which transmit data directly to a centralized computer [28], [29]. Hence, the following review regards works that dealt with optimizing the property of coverage as the main objective.

4.2   The importance of network's application

An air quality network of sensors is usually deployed for several designated applications. The two main ones include i) the management of a chemical leak due to an industrial accident, a disaster or an attack, that requires the detection of the resulted plume and the mapping the pollutant level in the environment [23] and ii) monitoring concentrations of pollutants emitted from routine activities. The different applications may form various goals that guide the optimization process aimed at finding the optimal sensors' placement. For example, Kanaroglou et al. [24] used population exposure assessment as a goal. In their method, they first generated a "demand surface", which represented the spatial variability in pollution concentrations. Then, their "demand surface" was modified to also consider the density of a population of interest, such as children or elderly. Eventually, the "demand surface" was used as input to an algorithm that solved a constrained single optimization problem (see Section 4.5) and identified the optimal locations for a predefined number of air pollution sensors. Their methodology however was limited to already monitored areas, since it relied on existing measurements from monitoring stations. Carter and Ragade [30] proposed a probabilistic model which took into account the detection probabilities of sensors of different types, assuming a decay with distance from a "target" or an "event" (i.e., pollution source). Their optimization procedure utilized a genetic algorithm to ensure a certain level of detection of an event is reached by the distributed sensors, while minimizing costs. In air pollution monitoring, unlike in many other fields, pollutants detected by air quality sensors are governed by a time varying wind field. Simply assuming an exponential decay of the detection probability with distance therefore lacks a physical base. Boubrima et al. [31] accounted for the various weather conditions characterizing their region of study, and designed a model for the deployment of sensors with the objective of detecting threshold crossings in order to trigger an adequate alert. They used a Gaussian dispersion model

and formulated an integer linear programming problem to minimize operational cost while assuring the detection of a threshold crossing. In their model however, sensors were eventually placed at the highest mean pollution zones, as calculated by assuming constant emission rates from sources. In another work, Boubrima et al. [32] simulated the deployment of a network with the goal of achieving the most effective data assimilation of air pollution measurements for the correction of physical model simulations. Similarly, Berman et al. [33] aimed to improve the performance of an interpolation-based model, using a geostatistical simulation. The researchers used an associated measure of the kriging interpolation method, which considered the uncertainty of the prediction. In the first step, they assessed how many sensors to add to a given network, using a Monte Carlo approach which evaluated how additional monitors change prediction precision through minimized uncertainty. In the second step, they assessed where to place the new monitors, using a similar Monte Carlo scheme which considered locations that improved prediction uncertainty and provided high prediction accuracy. Their method however required a robust preliminary deployed network that already captured the spatial variability to a certain extent, in order to be applied. In addition, their implementation did not provide a simultaneous assessment of number of sensors and their placement, due to the great affect it had on computation time.

4.3   Network heterogeneity

The simplest approach for sensors' deployment is to assume all MSUs are identical, thus creating a homogeneous WDESN. In some complex environments however, the existence of gas mixtures, that were either emitted from the same sources or from different ones, produce the need to account for multiple pollutants [34]. This can be attained through the use of a heterogeneous WDESN, comprised of different types of monitoring units, each measuring different target pollutants. Another option is to use the cross-sensitivity property of the device, and to take advantage of the interference effect generated by other gases reacting with the sensor's electrode [13], [35].

Not many studies have dealt with this aspect with relations to air pollution. Chakrabarty et al. [36] presented a grid coverage strategy for target detection purposes using different types of sensors. Altinel et al. [37] considered the uncertainty associated with sensor readings and assumed that a sensor is able to cover a given point with a predefined probability. In these studies, however, sensors differ from each other mainly in their monitoring range, a property not relevant to the in-situ character of air pollution monitoring. Lerner et al. [29] developed a method to optimally deploy a network of low-cost sensors of two types, measuring O_3, NO and NO_2 emissions. Their optimization process, constrained by available locations and a given budget, maximized the overall utility of the sensor network. The utility was comprised, in part, of the suitability of the type of sensor to the location of deployment, so for example, a sensor that measured O_3 with high accuracy and NO_2 with low accuracy, was less suited to monitor traffic-

related pollution. Lerner et al. mainly focused on the land use of the region of interest and ignored any knowledge on the pollutant field or weather conditions.

4.4 Redeployment of sensors

Another aspect of network deployment dealt with in some works, is redeployment strategies that relocate some or all of the sensors in order to enhance network performance. Some focused on route optimization, where, for example, a network of sensors is mounted on mobile platforms [23], and real-time adaptive deployment of sensors is performed according to the spreading of the pollutant. Kuroki et al. [38] developed an expert system for navigating unmanned aerial vehicles for optimal contamination mapping that enables estimation of source parameters. Other works focused on static redeployment strategies, where a new distribution of sensors is suggested so relocation of sensors is optimized for lowest energetic cost. Belkhiri et al. [39] proposed several deployment models, that simulated air pollution concentrations as ground truth and estimated concentrations where no sensors were deployed, using IDW interpolation. In their model, sensors were redeployed so estimation error (i.e., the absolute difference between ground truth and estimation) was minimized while regarding execution time.

4.5 Optimization aspect

As the basic problem of sensors' optimal deployment aims at maximizing the utility of the network, while minimizing its cost, the problem is equivalent to the "0-1 knapsack" problem [29], [40], where a subset of items out of n items, possessing each some value v_i and some cost c_i, are selected so values' sum is maximized, while the summed cost is kept within some capacity c_{max}. In both problems, candidate solutions can be represented as binary strings of length n, where a 1 in a given position indicates that an item is included in the knapsack, or that a sensor is placed in that position, and a 0 indicated the item is omitted, or the sensor is not placed. The solution-space size is then 2^n. This basic problem is NP-hard [41], meaning that the time required to solve the problem, using any currently known algorithm, increases rapidly as the size of the problem grows. As a result, approximation algorithms, which focuses on finding good solutions ("global optimums") instead of provably optimal solutions, are required beyond a certain size of problem.

Global optimization refers to the process of attempting to find the solution x^*, out of a set of possible solutions $\{X\}$, that has the optimal value for some fitness function f, such that $x \neq x^* \rightarrow f(x^*) \geq f(x)$ [42]. Two main categories of methods to find x^* exist; the first are deterministic methods, which find x^* by an exhaustive search over the set $\{X\}$. Stochastic methods involve random elements that try to reduce the computational burden of pure random search. At the outset, a random sample of points in the set $\{X\}$ is picked. Then, each method manipulates the sample points using different heuristics, which are sets of rules for deciding which potential solution out of $\{X\}$ should next be generated and tested [43]. Local search

algorithms (often referred to as hill climbers), work by taking a starting solution x, and then searching $\{X\}$ in the neighboring environment for one x' that fulfils $f(x') > f(x)$. The process will eventually lead to the identification of a local optimum, with no guarantee for the quality of the solution found, compared to x^*. As a result, local searches are usually incorporated in stochastic methods to yield candidate global optimums, from which the best point is eventually picked [44].

Evolutionary Algorithms (EA) are classified as Meta (i.e., problem-independent)-heuristic optimization algorithms. Inspired by the biological theory [45], if given a population of individuals (i.e., a set of candidate solutions), the environmental pressure causes natural selection and so according to a fitness measure (i.e., an objective function), the better candidates have a higher chance to survive and reproduce (i.e., to stay in the set of candidate solutions and generate new solutions by variation operators such as crossover and mutation). Crossover of two or more selected parents (i.e., selected solutions) may result in one offspring or more (i.e., new solutions) that based on their fitness will compete with the old candidates for a place in the next generation [42].

The ability of EA to maintain a diverse set of solutions, by creating new solutions from a non-uniform distribution, not only provides means of escaping from one local optimum [23] and handling poor initial estimates; it provides means of coping with large and discontinuous search spaces. As a consequence, EA were shown to provide near-optimal results in many studies (e.g., [30], [46]–[48]). EA were applied to many STE problems; Haupt et al. was the first to demonstrate the ability of an EA to optimize the source calibration factors that couple a Gaussian plume ATD model with readings from receptors [49], [50]. Later, Allen et al. [51] used this method to characterize a pollutant source by estimating its location, strength and the surface wind direction, this way accounting for the sparse resolution of meteorological wind field. Other modifications to this methods included the use of more complex ATD models [52] or the calculations of the number of sensors necessary to identify multiple source and wind parameters [53].

This proposed work offers new approaches to real-world air pollution problems, by combining environmental aspects with state-of-the-art optimization technique. Using the proposed methodology would allow decision makers to retrieve sets of optimal solutions for network deployment, each suited for different use-cases. To the best of our knowledge, and as reviewed above, no other work has offered such perspective to solve similar problems. The basic proposed methodology will be reviewed thoroughly in the next section.

5 Methods and Research plan

5.1 Background

The proposed methodology requires the use of an atmospheric transport and dispersion (ATD) model in order to connect sources' emissions and predicted concentrations in all potential locations to place the sensors at. Several types of fundamental dispersion models exist: box models, Gaussian plume models, Lagrangian models, Eulerian dispersion models and Dense gas models as well as extensions and combinations of all of the above [23]. High resolution modeling tools for complex environments are available for most model types, such as the Gaussian plume AERMOD [54], [55] and the Eulerian grid-based CMAQ [19], both developed by the US Environmental Protection Agency (EPA). Other commonly used tools are the HYSPLIT [19] and GRAL [56], [57] models, which simulate pollutant emission by releasing a fixed number of Lagrangian particles from the source. At the first stages, we will use the Gaussian plume model and will not account for topography or land use. Nevertheless, our proposed methodology is invariant to the selected ATD model, and can easily be replaced in the future, as all computations of the various weather conditions and source combinations are computed in advance. Once our method is proved to be effective, a more complex ATD model will be considered.

5.1.1 The Gaussian plume model

The Gaussian plume model is one of the simplest and widely used models that offers an analytical steady state solution to the advection-diffusion equation for idealized circumstances, corresponding to a continuous point-source that emits pollutants into a unidirectional wind blowing in a domain of infinite extent. The Gaussian plume model (Eq.1) eventually describes the pollutants' concentration C [kg m^{-3}] in a certain position in space:

$$C(x,y,z,Q) = \frac{Q}{\bar{u}\sigma_y\sigma_z 2\pi}\exp\left(\frac{-y^2}{2\sigma_y^2}\right)\left[\exp\left(\frac{-(z-H)^2}{2\sigma_z^2}\right) + \exp\left(\frac{-(z+H)^2}{2\sigma_z^2}\right)\right] \tag{1}$$

Where Q [kg s^{-1}] is the release rate from the stack, H [m] is the effective height (i.e., the sum of the actual stack height h and the plume rise Δh), x, y and z [m] are the downwind, crosswind and vertical distances, respectively, and $\bar{u}$ [m s^{-1}] is the mean wind speed at the height h of the release. σ_y and σ_z [m] are key parameters in the model and represent the standard deviation of the Gaussian concentration distribution in the crosswind and vertical direction. Several simplifying assumptions are made in order to reach the Gaussian plume model equation (Eq. 1) [18]:
  i)   The pollutant is emitted in a constant rate Q from a single point source located at height H above ground surface.
  ii) The wind velocity is constant and aligned with the positive x-axis.
  iii) The solution is steady state.

iv) The eddy diffusion coefficient K [m² s⁻¹] is a function of the downwind distance x only, and diffusion is isotropic so that $K_x(x) = K_y(x) = K_z(x) =: K(x)$.

v) Wind velocity is sufficiently large that diffusion in the x-direction is much smaller than advection.

vi) Variations in topography are negligible so that the ground surface can be taken as the plane z=0.

vii) The pollutant does not penetrate the ground.

To these assumptions we add that the material diffused is non-reactive in the atmosphere, with a negligible deposition rate and that background pollution is negligible as well. σ parameters are used instead of the eddy diffusion coefficient $K(x)$, due to the fact that they are much easier to determine experimentally and can usually be described by a simple power law of the form:

$$\sigma_y = ax^b \tag{2a}$$

$$\sigma_z = cx^d + f \tag{2b}$$

This kind of dependence on downwind distance (x) expresses the widening of the plume when moving further from the source. Experimental measurements have been used to estimate the coefficients a, b, c, d and f under a variety of atmospheric conditions, using Pasquill-Gifford stability class categories [58]. The coefficients used in this work [59] are specified in Table 1.

5.1.2 Meteorology

As reviewed above, the meteorological parameters that play a role in any gas dispersion model are wind speed, wind direction, and the atmospheric stability condition. The stability condition is a function of the vertical temperature profile of the atmosphere, which is govern by processes such as solar heating,

Table 1. Constants of empirical relationships between x and σ_y and σ_z (Eq. 2) for different stability classes: A – extremely unstable, B – moderately unstable, C – slightly unstable, D – neutral, E – slightly stable and F – moderately stable. b parameter is taken as 0.894. From Martin (1976).

Stability class	a	$x \leq 1$ km		
		c	d	f
A	213	440.8	1.941	9.27
B	156	106.6	1.149	3.3
C	104	61.0	0.911	0
D	68	33.2	0.725	−1.7
E	50.5	22.8	0.678	−1.3
F	34	14.35	0.740	−.035

radiative cooling or winds (cold or warm advection). In general, three regimes of atmospheric stability exist: unstable, neutral and stable. In an unstable atmosphere, an air parcel that starts to move upward will continue to rise because it is warmer and less dense than the air around it. Pollutants will be mixed rapidly in extreme vertical motions due to thermal turbulence. At a stable atmosphere, an air parcel will resist an upward vertical motion and will tend to spread out

horizontally. In stable conditions, the reduction of vertical exchange reduces pollutant mixing, and may even damp out some of the mechanical turbulence, caused by the friction with Earth's surface. Between the above two conditions is the condition characterized as neutral, when temperatures decrease slightly with height, in a rate close to the dry adiabatic rate.

Typical diurnal changes in the stability of the lower atmosphere exist. During the night, especially when winds are light and skies are clear, the radiative cooling of the ground surface often leads to surface air that is colder than the air above it. A stable layer thus exists in the lower hundred or so meters in the atmosphere. Pollutants emitted during the night inside this shallow layer get trapped and can reach relatively high concentrations. As the sun rises, the ground and the air next to it start warming up and the temperature profile corresponding to an unstable atmosphere is established. This change occurs over a period of a few hours in the morning and results in breaking of the inversion usually before noon [1], [60], [61].

In order to apply the suitable parameters to the Gaussian plume model (Table 1), the Pasquill-Gifford stability category (A-F) which characterizes the simulated set, should first be determined. The original method to do so, developed by Turner [62], requires knowledge on cloud cover and cloud ceiling. Alternative empirical methods were developed for situations where these data are not available [63]. For the proposed work, we chose to use a turbulence-based method (σ_A method), because of its simplicity and our currently available data. The σ_A method uses the standard deviation of the wind direction (σ_θ) in combination with the scalar mean wind speed ($\bar{u}$), to determine the Pasquill-Gifford stability category, as specified in Table 2 and Table 3.

5.1.3 The Borg Multi-Objective Evolutionary Algorithm (MOEA) framework

Unlike single objective optimization, which employs a single criterion for identifying the best solution among a set of alternatives, multi-objective optimization employs two or more criteria, i.e., two or more objectives. As multiple objectives can conflict with one another such that improving one objective leads to the deterioration of another - there is no single optimal solution to multi-objective problems. Tradeoffs are captured by solutions which are superior in some objectives but inferior in others. Such pairs of solutions which are both superior and inferior with respect to certain objectives are called non-dominated. The set of all nondominated solutions in a given generation is referred to as the Pareto front [64].

For the proposed work, the self-adaptive Borg MOEA [65] is used. The Borg is classified as a hyper-heuristic global multi-objective search tool [66], as it uses internal feedbacks during the search progress to dynamically adapt an ensemble of genetic operators, rewarding those that maximize search progress (i.e., the dominance of new generated solutions). A stochastic restart mechanism exploiting Uniform Mutation is built into the Borg MOEA to automatically detect and avoid pre-mature convergence to a local optimum and to achieve a diverse set of solutions [64], [65].

5.2   Problem formulation

5.2.1   Notation

Let Ω be the region of interest, i.e., the industrial area and its surroundings, where we wish to place our network of sensors at. Let $\{S\}$ be the set of sources, where each source, $s \in \{S\}$ is located in $\omega_s \in \Omega$ and for a specific time t, emits $q_s(t)$ [kg s⁻¹]. Similarly, the set $\{R\}$ is the set of sensors, where each sensor, $r \in \{R\}$, is located in $\omega_r \in \Omega$ and records a pollution level of $c_r(t)$ [μg m⁻³]. The location of the sources $\{\omega_s\}$ are known, these are the locations of the industrial plants.

Let m_{sr} be the pollution transfer function of the dispersion model, which associates sensor's r readings, $c_r(t)$, with the emissions of source s. Thus, the model's estimated contribution of source $s \in \{S\}$ to the pollution level in $\omega \in \Omega$, is given by:

$$\hat{c}_\omega(t) = m_{s\omega} \cdot q_s(t) \mid \forall \omega \in \Omega, \forall s \in \{S\} \tag{3}$$

For multiple sources scenario, each sensor readings are simply the sum of all sources' contributions, which are the sources' emissions multiplied by their corresponding transfer function, $m_{s\omega}$, i.e.:

$$\hat{c}_\omega(t) = \sum_s m_{s\omega} \cdot q_s(t) \, \Bigg| \, \forall \omega \in \{\Omega\}, \forall s \in \{S\} \tag{4}$$

As each source owns its unique parameters and location with respect to the sensors, the values of m_{sr} for each source-sensor combination is determined by the dispersion model, such as the Gaussian plume model.

5.2.2   Pairwise Euclidean Distance (PED)

With the objective of finding a possible increase in industrial gas emissions, a criterion that incorporates the sensitivity of the deployed network to changes in the source term needs to be defined. We propose a new criterion for evaluating how well a hypothetical network of sensors, placed in hypothetical locations, would respond to slight changes in the source term. The criterion is the pairwise Euclidean distance (PED) between calculated readings of two sensor network sets with different number of active sources. That PED value of two sets of active sources, as measured with the same WDESN having different $\{S'\}$ and $\{S''\}$, is given by:

$$\text{PED}_{\{S'\},\{S''\}} = \sqrt{\sum_{r=1}^{|R|} \left(C_r^{\{S'\}} - C_r^{\{S''\}} \right)^2} \tag{5}$$

A network set that is sensitive to changes in the source term will produce high PED values. On the other hand, low PED values will characterize a non-sensitive network of sensors and might indicate on low ability of such network to resolve the specific source term estimation problem.

In general, when the number of sensors $|R|$ increases, the PED values, calculated between two sets of active sources with different sizes $\{S\}'$ and $\{S\}''$, increases as well.

5.2.3 Optimization

In the following, we describe the optimization problem of the first goal of the proposed work, i.e., the design of a spatial optimization model, where we aim at finding an optimal deployment (number and locations) of a set of sensors $\{R\}$, considering a set of potential positions in the region of interest.

The network of sensors to be deployed, considers the varying meteorological conditions. Hence, our objective takes into account the discrete frequencies of each possible meteorological combination of wind speed (WS), wind direction (WD) and Pasquill–Gifford atmospheric stability class (SC). Each combination (WS, WD, SC) is assumed to have a probability of p to occur, such that $\sum p = 1$. The Gaussian plume model is applied with all meteorological combinations, to obtain steady state solutions of pollution concentration maps. Each concentration map is given a weight according to the meteorological frequencies derived and so the average concentration map is generated to all sets of active/non-active combinations of sources.

Given that the deployment of a set of sensors is an expensive task due to sensors' cost and maintenance, we may wish to minimize the number of sensors in use, and place them in a set of locations that together generate the highest PED value, which represents the best ability of those sensors to detect changes in industrial gas emissions. The first objective is then to minimize the number of sensors deployed:

$$\Psi_{\#r} = \min_{\{R\}}|R| \tag{6}$$

It is possible to phrase several different objective functions that maximize the PED value and conflicts with the minimization of the number of sensors. We suggest here several possible objectives:

a) Taking the average of minimal PED values out of each transformation pair $\{S'\} \leftrightarrow \{S''\}$:

$$\Psi_{\text{minimal\_PED}} = \max\{\frac{\sum \min_{\{S'\}\leftrightarrow\{S''\}}\left(\text{PED}_{\{S'\},\{S''\}}\right)}{\text{NCTP}}\} \tag{7}$$

For example, in the case of 5 sources, 31 options of 1-5 active sources exist, creating in total 355 possible combinations of two sets of active sources with different sizes {S}' and {S}''. In this case, an array of 355 PED values are calculated for each optional solution. Out of 355 PED values, Number of 10 Categories of Transformation Pairs (NCTP) between same number of sources exist (1<->2 ,1<->3, etc.).

b) Taking the xth percentile PED value, considering all possible combinations:

$$\Psi_{percentile\_PED} = \max\{x^{th}\ percentile(PED_{\{S'\},\{S''\}})\} \qquad (8)$$

c) Splitting the second objective to two or more objectives, where each considers PED values obtained during different hours of the day. In areas where wind field, stability class or industrial activity is substantially different between day time and night time for example, more conflicting objectives are formed, and other considerations should be taken into account in order to achieve an optimal deployment.

Other objectives will further be examined as the research progresses, including the addition of more PED-related objectives.

5.2.4   Optimization procedure

Once all pre-calculations are obtained, the optimization process starts. The Borg MOEA algorithm (see Section 5.1.3) initiates its search by starting with a uniform random generation of its initial population of candidate solutions. Then, the algorithm projects the set $\{R\}$ on Ω and evaluates all possible PED values of that set (Eq. 5). Then, based on the defined objectives, the Borg MOEA rewards those sets of decision variables (i.e., possible solutions) that dominate competing alternatives (i.e., better in all objectives) until a high-quality approximation of the Pareto frontier is attained. Technically, we assume that each location $\omega \in \Omega$ may host a sensor. The decision variables are held in a data array, each value represents a sensor, linked to a certain location $\omega \in \Omega$. The sensor is "placed" in that location if the decision variable is 1 and is not "placed" if the decision variable is 0.

5.3   Research plan

The research will include three main stages:

i) In the first stage, a *spatial* optimization model will be designed, that computes the tradeoff between sensor array effectivity, i.e. PED and cost. The model will capture the distribution of weather conditions that characterize the regime and sensors of various attributes. Evaluation of the model will include optimization performance evaluations and model evaluation through the examinations of different scenarios.

ii) In the second stage, a *temporal* optimization model will be designed, that computes sensors' redeployment strategy in case of changing weather conditions of different time scales (e.g., 1 hour, 1 day, season, etc.), which induces minimal transfer effort. Various distance metrics as well as the number of sensors to relocate will be considered as objective functions to be minimized.

iii) The third stage will consolidate the two proposed schemes into one generalized model. The probability of change in short-term weather conditions will be considered, so it is most probable that future deployments require minimum transfer efforts.

During each of the specified phases, several extensions will be considered, including the use of other more sophisticated dispersion models, other forms of objective functions, related to the maximization of the PED values, and real-world test cases. In the second and third stages, a case of sudden changes in number of sources or emission patterns will potentially be examined as a trigger for network redeployment.

6   Preliminary results

6.1   Simulation setup

An initial simulation setup was explored. A 1000x1000 m flat area was taken as the simulated arena, with sensors situated at ground level (i.e., z=0 in Eq. 1). Five point-sources (stacks) were arbitrarily given average emission rates of 0.47, 0.51, 0.38, 0.9 and 0.19 [kg s⁻¹]. Figure 1 illustrates the simulation set with pollutant concentrations obtained by the Gaussian plume model for light (4 m s⁻¹) west winds (270 degrees) and stability class C (slightly unstable). Sources are marked in red dots and optional locations of sensors are marked in black crosses. 396 optional locations were spread in a grid, 50 m apart, and at least 50 meters from the location of each source. Effective heights of stacks were taken at 10 m.
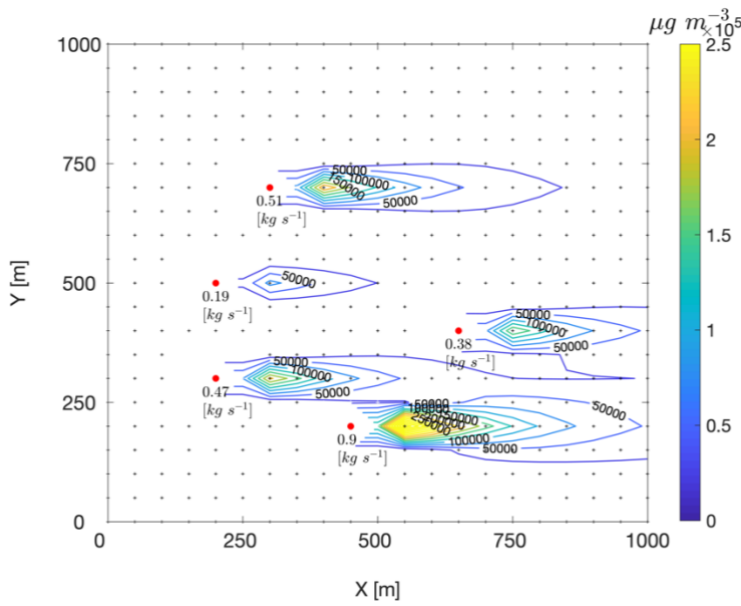

Figure 1. Simulated arena, showing sources (red dots) with their average emission rates [kg s⁻¹], optional sensor locations (black crosses) and steady-state concentrations [$\mu g\ m^{-3}$], obtained by the Gaussian plume model for weather state of (WS, WD, SC) = (4 m s⁻¹, 270 degrees, C class).
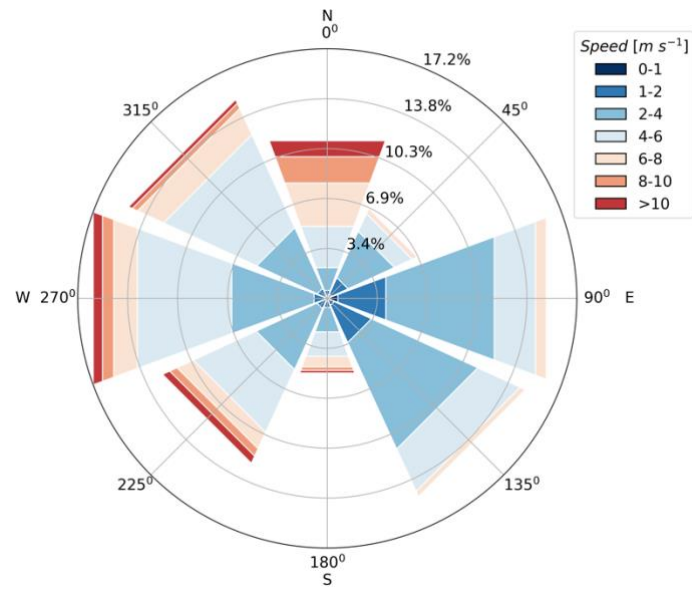
Figure 2. Windrose distribution of wind data used in the initial simulation. Color represents speed distribution (in percentage) in each direction category. Data obtained from IMS for Hadera-port station.

Hourly mean wind directions were binned into 8 categories (45 degrees each) and hourly mean wind speeds into 7 categories (<1, 1-2, 2-4, 4-6, 6-8, 8-10, >10 m s⁻¹). With six possible Pasquill–Gifford classes for atmospheric stability (A-F), 336 weather states exist. The Gaussian plume model was applied with each weather state combination, using the mean value of each wind speed and direction category. For our initial simulation, we used 10-minutes wind data from Hadera-port station, obtained from the Israel Meteorological Service (IMS) for the years 2004-2018, to derive frequencies of each binned category (see Section 5.2.3). Figure 2 shows the windrose distribution of winds used in the initial simulation. The taken wind data are characterized by a bi-model distribution - comprised of west winds of relatively high speeds compared to east winds of lower speeds. The vector-average wind speed and direction is 0.9 m s⁻¹ and 298 degrees, respectively. Scalar-average wind speed is 4.2 m s⁻¹. Distribution of stability

classes is as follows: A-1%, B-1%, C-6%, D-78%, E-11%, F-3%, as determined according to the σ_A method (see Section 5.1.2).

6.2 Obtained solutions

The optimization process was constrained to find the optimal locations of 2-300 sensors. Figure 3 illustrates the Pareto frontier received by two optimization runs, using the number of sensors ($\Psi_{\#r}$) as the first objective (Eq. 6) and average minimal PED values ($\Psi_{minimal\_PED}$) or the 25th

percentile of all PED values ($\Psi_{percentile\_PED}$) as the second objective (Eq.7 and Eq.8, respectively). The ultimate optimal solution in Fig. 3 would be the most upper left corner of the graph, which represents zero cost and highest PED. However, as this goal cannot be achieved due to the constraints of the problem, each dot in Fig. 3 represents a non-dominated solution (see Section 5.1.3), i.e., an optimal solution for a given number of sensors. The conflict between the cost of the network and its performance is reflected in the Pareto frontier - higher PED values and better network performance are achieved by larger number of sensors but result in a more expensive sensor array.
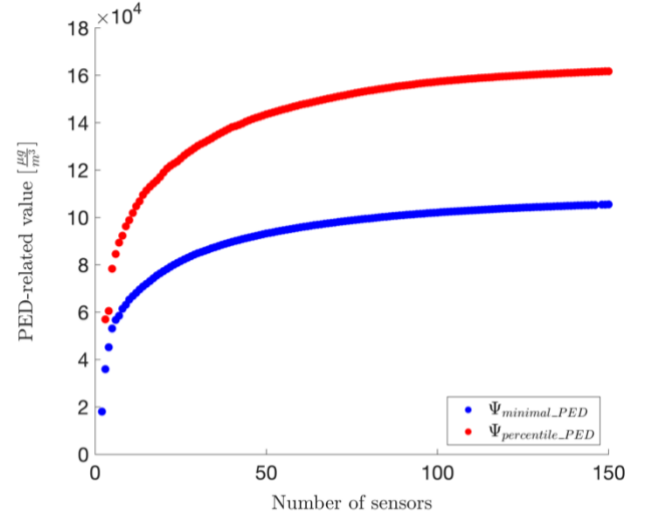

Figure 3. Pareto frontier of two optimization runs with two different PED-related objectives - $\Psi_{minimal\_PED}$ (blue) and $\Psi_{percentile\_PED}$ (red). Number of sensors were constrained to 2-300.

For the above-mentioned obtained solutions, Fig. 4 displays a heatmap of preferable locations sensors were placed at. Each colored pixel on the 1000x1000 m simulation-map
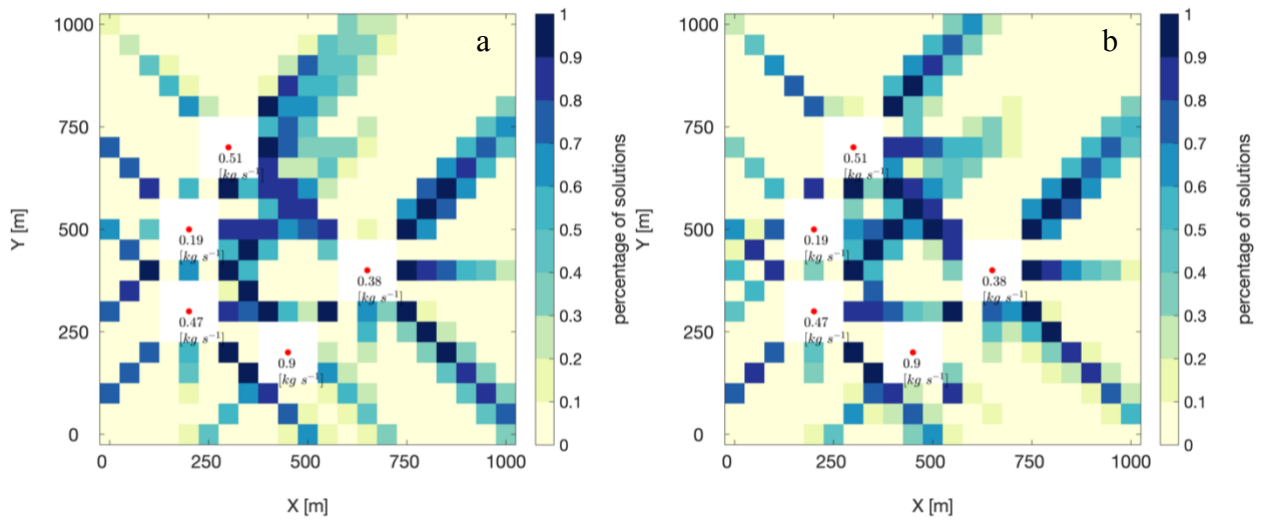

Figure 4. Preferable sensor locations, represented by the percentage of solutions each specific location was chosen, for two objectives - (a) average minimal PED values (172 solutions in total) and (b) 25th percentile of all PED values (167 solutions in total). Sources' locations are marked in red dots and average emission rates are specified below. White areas are restricted locations, due to sources' proximity.

represents the percentage of solutions a sensor was placed at that specific location. It can be seen that the two objectives - $\Psi_{\text{minimal\_PED}}$ (Fig. 4a) and $\Psi_{\text{percentile\_PED}}$ (Fig. 4b) generate quite similar preferable locations, though some differences are seen; in Fig. 4a for example, the most preferable places (colored in dark blue) are between sources and as close as possible to a source. In Fig. 4b on the other hand, preferred locations are in between sources, but not necessarily as close as possible to them. These results are illustrated in order to show the process of finding a suitable objective function. An evaluation process of these solutions is further needed in order to determine which set of solutions performs better in measuring pollutant concentrations and detecting changes in emissions.

For comparison, mean concentrations of the tested simulation-set, considering all possible weather states and their frequencies, are shown in Fig. 5 No clear correlation is seen between high average concentrations and preferable locations of sensors (Fig. 4), showing sensors are not necessarily placed at locations where concentrations are high. Pearson correlation between Fig. 4a and Fig. 5 and between Fig. 4b and Fig. 5 is 0.47 and 0.51, respectively.

Figure 7 (see Appendix C) shows several optimal solutions from one optimization run, using $\Psi_{\#r}$ as the first objective and $\Psi_{\text{minimal\_PED}}$ as the second objective. Solutions for relatively small number of sensors (low-cost networks) - 2, 4, 5, 7, 8 and 15 are displayed, in order to highlight the process of sensors' deployment.

Figure 5. Mean concentration map of simulated emissions [μg m⁻³], considering all possible weather states and their frequencies. White areas are restricted locations, due to sources' proximity.

It seems that preferable locations of sensors follow the highest derivative of the gas concentration in the air as these places are more sensitive to the nature of the source. This assumption will be tested in the suggested study. A colormap of minimal PED values of each transformation pair between each number of active sources are attached to each displayed solution. For example, in the colormap to the right of Fig. 7a, the minimal PED values of the transformation pairs 2<->5 and 1<->4 are around $3 \cdot 10^4$ μg m⁻³ (30 ppm). The higher the number of deployed sensors, the higher the minimal PED values.

7 Work schedule

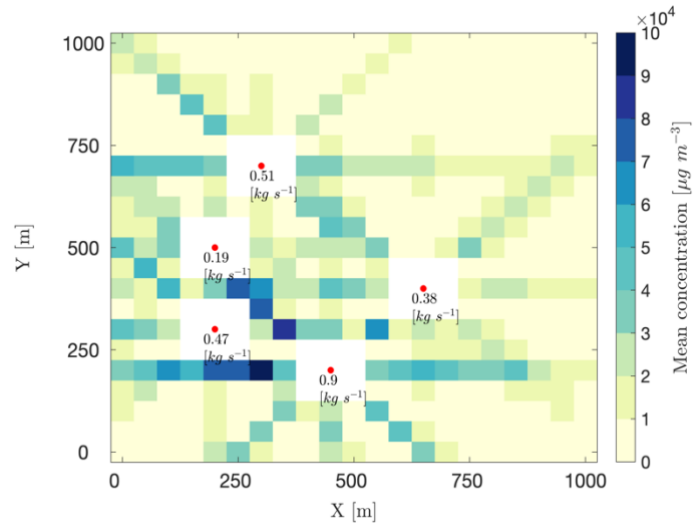The following Gantt diagram (Fig. 6) presents the work schedule of the first two planned modules, as well as work accomplished so far. Milestones are marked in triangles and tasks are marked in rectangles. Each major task is broken down into smaller tasks, which are specified underneath it.
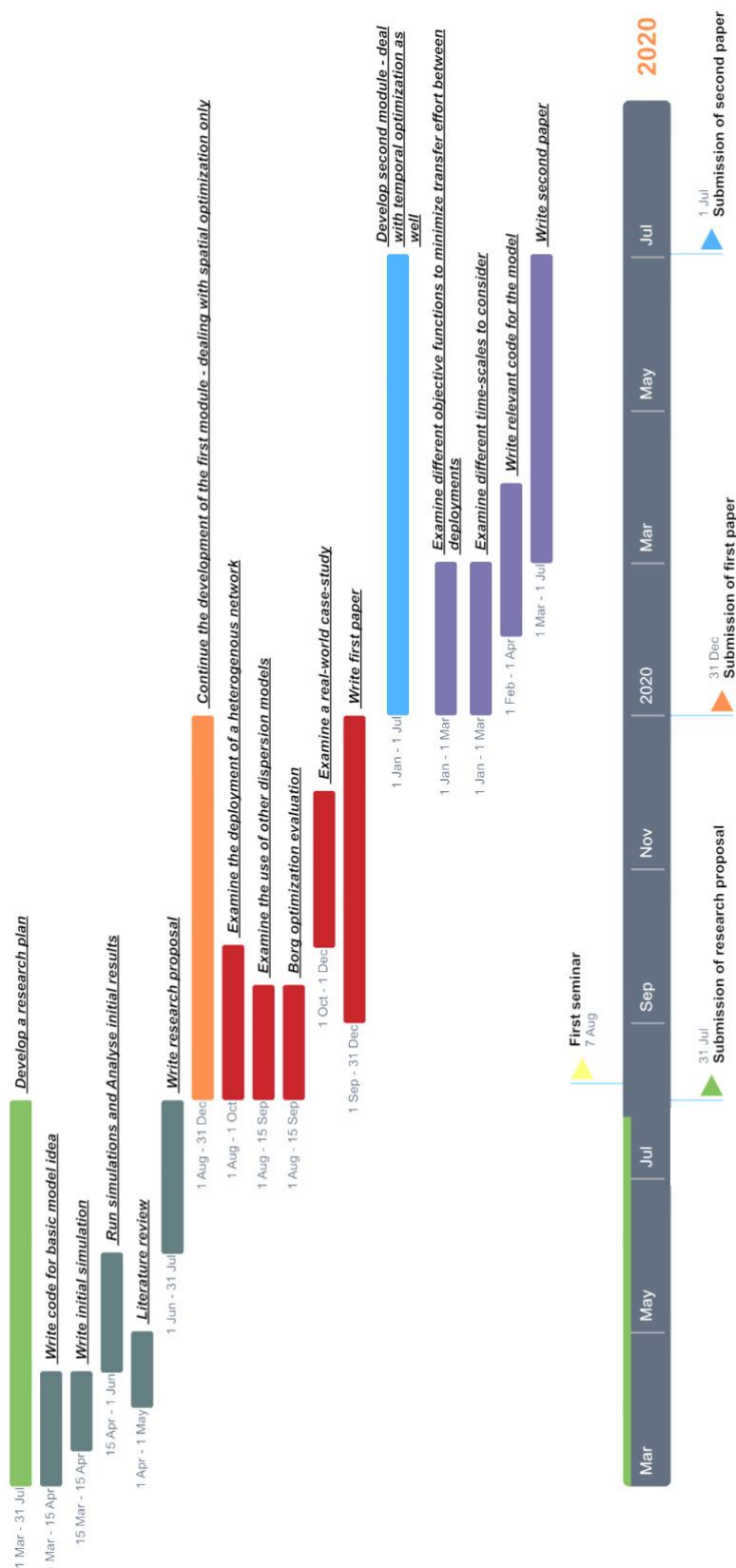

Figure 6. Research Gantt diagram. Milestones are marked in triangles. Tasks are marked in rectangles.

8   References

[1]   J. H. Seinfeld and S. N. Pandis, *Atmospheric chemistry and physics - from air pollution to climate change*, Second edi. Wiley-Interscience, 2006.

[2]   D. J. Jacob and D. A. Winner, "Effect of climate change on air quality - citation," *Atmos. Environ.*, vol. 43, no. 1, pp. 51–63, 2009.

[3]   "Greenhouse Gas Emissions," *EPA, United States Environmental Protection Agency*, 2019. [Online]. Available: https://www.epa.gov/ghgemissions/overview-greenhouse-gases. [Accessed: 23-May-2019].

[4]   "Israel ministry of environmental protection," *The Impact of Air Pollution from Industry*, 2015. [Online]. Available: http://www.sviva.gov.il/English/env_topics/Health-and-Environment/Health-Impact-of-Environmental-Nuisances/Pages/The-Impact-of-Air-Pollution-from-Industry.aspx. [Accessed: 23-May-2019].

[5]   WHO, "Ambient air pollution: Health impacts," 2019. [Online]. Available: https://www.who.int/airpollution/ambient/health-impacts/en/. [Accessed: 10-Mar-2019].

[6]   IPCC, *Climate Change 2013*, vol. 5. 2014.

[7]   A. M. Fiore, V. Naik, and E. M. Leibensperger, "Air quality and climate connections," *J. Air Waste Manag. Assoc.*, vol. 65, no. 6, pp. 645–685, 2015.

[8]   M. Jerrett *et al.*, "A review and evaluation of intraurban air pollution exposure models," *J. Expo. Anal. Environ. Epidemiol.*, vol. 15, no. 2, pp. 185–204, 2005.

[9]   F. Kizel *et al.*, "Node-to-node field calibration of wireless distributed air pollution sensor network," *Environ. Pollut.*, vol. 233, pp. 900–909, 2018.

[10]   U. Lerner, T. Yacobi, I. Levy, S. A. Moltchanov, T. Cole-Hunter, and B. Fishbain, "The effect of ego-motion on environmental monitoring," *Sci. Total Environ.*, vol. 533, pp. 8–16, 2015.

[11]   S. Moltchanov, I. Levy, Y. Etzion, U. Lerner, D. M. Broday, and B. Fishbain, "On the feasibility of measuring urban air pollution by wireless distributed sensor networks," *Sci. Total Environ.*, vol. 502, pp. 537–547, 2015.

[12]   A. Marjovi, A. Arfire, and A. Martinoli, "High Resolution Air Pollution Maps in Urban Environments Using Mobile Sensor Networks," *2015 Int. Conf. Distrib. Comput. Sens. Syst.*, pp. 11–20, 2015.

[13]   M. I. Mead *et al.*, "The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks," *Atmos. Environ.*, vol. 70, pp. 186–203, 2013.

[14]   R. Piedrahita *et al.*, "The next generation of low-cost personal air quality sensors for quantitative exposure monitoring," *Atmos. Meas. Tech.*, vol. 7, no. 10, pp. 3325–3336, 2014.

[15]   N. Castell *et al.*, "Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates?," *Environ. Int.*, vol. 99, pp. 293–302, 2017.

[16]   A. Nebenzal and B. Fishbain, *Hough-Transform-Based Interpolation Scheme for Generating Accurate Dense Spatial Maps of Air Pollutants from Sparse Sensin*, vol. 507. Springer International Publishing, 2017.

[17]   P. H. Ryan and G. K. LeMasters, "A Review of Land-use Regression Models for Characterizing Intraurban Air Pollution Exposure," *Inhal. Toxicol.*, vol. 19, no. sup1, pp. 127–133, Jan. 2007.

[18]   J. M. Stockie, "The Mathematics of Atmospheric Dispersion Modeling," *SIAM Rev.*, vol. 53, no. 2, pp. 349–372, 2011.

[19]   A. F. Stein, V. Isakov, J. Godowitch, and R. R. Draxler, "A hybrid modeling approach to resolve pollutant concentrations in an urban area," *Atmos. Environ.*, vol. 41, no. 40, pp. 9410–9426, 2007.

[20]   R. L. Burritt and C. Saka, "Quality of physical environmental management accounting information, Lessons from Pollutant Release and Transfer Registers," *Sustain. Account. Report.*, pp. 373–407, 2006.

[21]   D. Kerret and G. M. Gray, "What do we learn from emissions reporting? Analytical considerations and comparison of pollutant release and transfer registers in the United States, Canada, England, and Australia," *Risk Anal.*, vol. 27, no. 1, pp. 203–223, 2007.

[22]   R. Sullivan and A. Gouldson, "Pollutant release and transfer registers: Examining the value of government-led reporting on corporate environmental performance," *Corp. Soc. Responsib. Environ. Manag.*, vol. 14, no. 5, pp. 263–273, 2007.

[23]   M. Hutchinson, H. Oh, and W. H. Chen, "A review of source term estimation methods for atmospheric dispersion events using static or mobile sensors," *Inf. Fusion*, vol. 36, pp. 130–148, 2017.

[24]   P. S. Kanaroglou *et al.*, "Establishing an air pollution monitoring network for intra-urban population exposure assessment: A location-allocation approach," *Atmos. Environ.*, vol. 39, no. 13, pp. 2399–2409, 2005.

[25]   B. Liu, O. Dousse, P. Nain, and D. Towsley, "Dynamic coverage of mobile sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 2, pp. 301–311, 2013.

[26]   Q. Zhao and A. Swami, "Coverage and Connectivity in Wireless Sensor Networks," *Adapt. Cross Layer Des. Wirel. Networks*, pp. 301–323, 2010.

[27]   H. Zhang and C. Liu, "A Review on Node Deployment of Wireless Sensor Network," *IJCSI Int. J. Comput. Sci. Issues*, vol. 9, no. 6, pp. 378–383, 2012.

[28]   B. Fishbain, U. Lerner, N. Castell, D. M. B. a Tom Cole-Hunter c, d, Olalekan Popoola e, and A. B. b Tania Martinez Iñiguez c, d, Mark Nieuwenhuijsen c, Milena Jovasevic-Stojanovic f, Dusan Topalovic f, g, Roderic L. Jones e, Karen S. Galea h, YaelEtzion a, FadiKizel a, Yaela N. Golumbic a, i, Ayelet Baram-Tsabari i, Tamar Yacobi a, Dana Drahler a, Johan, "An evaluation tool kit of air quality micro-sensing units," *Sci. Total Environ.*, vol. 575, no. September 2016, pp. 639–648, 2017.

[29]   U. Lerner, O. Hirshfeld, and B. Fishbain, "Optimal Deployment of a Heterogeneous Environmental Sensor Network," *Jorunal Environ. Informatics*, no. X, pp. 1–9, 2018.

[30]   B. Carter and R. Ragade, "A probabilistic model for the deployment of sensors," *SAS 2009 - IEEE Sensors Appl. Symp. Proc.*, pp. 7–12, 2009.

[31]   A. Boubrima, W. Bechkit, and H. Rivano, "Optimal WSN Deployment Models for Air Pollution Monitoring," *IEEE Trans. Wirel. Commun.*, vol. 16, no. 5, pp. 2723–2735, 2017.

[32]   A. Boubrima, W. Bechkit, H. Rivano, and L. Soulhac, "Leveraging the potential of WSN for an efficient correction of air pollution fine-grained simulations," *Proc. - Int. Conf. Comput. Commun. Networks, ICCCN*, vol. 2018-July, 2018.

[33]   J. D. Berman, L. Jin, M. L. Bell, and F. C. Curriero, "Developing a geostatistical simulation method to inform the quantity and placement of new monitors for a follow-up air sampling campaign," *J. Expo. Sci. Environ. Epidemiol.*, vol. 29, no. 2, pp. 248–257, 2019.

[34]   R. Berkowicz, M. Winther, and M. Ketzel, "Traffic pollution modelling and emission data," *Environ. Model. Softw.*, vol. 21, no. 4, pp. 454–460, 2006.

[35]   C. C. Austin, B. Roberge, and N. Goyer, "Cross-sensitivities of electrochemical detectors used to monitor worker exposures to airborne contaminants: False positive responses in the absence of target analytes," *J. Environ. Monit.*, vol. 8, no. 1, pp. 161–166, 2006.

[36]   K. Chakrabarty, S. S. Iyengar, H. Qi, and E. Cho, "Grid coverage for surveillance and target location in distributed sensor networks," *IEEE Trans. Comput.*, vol. 51, no. 12, pp. 1448–1453, 2002.

[37]   I. K. Altinel, N. Aras, E. Güney, and C. Ersoy, "Binary integer programming formulation and heuristics for differentiated coverage in heterogeneous sensor networks," *Comput. Networks*, vol. 52, no. 12, pp. 2419–2431, 2008.

[38]   Y. Kuroki, G. S. Young, and S. E. Haupt, "UAV navigation by an expert system for contaminant mapping with a genetic algorithm," *Expert Syst. Appl.*, vol. 37, no. 6, pp. 4687–4697, 2010.

[39]   A. Belkhiri, W. Bechkit, H. Rivano, and M. Koudil, "Context aware MWSN optimal redeployment strategies for air pollution timely monitoring," *IEEE Int. Conf. Commun.*, vol. 2018-May, 2018.

[40]   H. Kellerer, U. Pferschy, and D. Pisinger, *Knapsack Problems*. Springer, 2004.

[41]   R. M. Karp, "Reducibility among combinatorial problems," in *Complexity of computer computations*, Springer, 1972, pp. 85–103.

[42]   A. E. Eiben and J. E. Smith, *Introduction to evolutionary computing*, 2nd Editio., vol. 53. Springer, 2007.

[43]   G. O. Concepts, *Introduction to Optimum Design, Chapter 16 - Global Optimization Concepts and Methods*, Fourth Edi. Elsevier Inc., 2017.

[44]   S. E. Haupt, G. S. Young, and C. T. Allen, "A genetic algorithm method to assimilate sensor data for a toxic contaminant release," *J. Comput.*, vol. 2, no. 6, pp. 85–93, 2007.

[45]   J. H. Holland, *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT press, 1992.

[46]   R. Ramadan, H. El-Rewini, and K. Abdelghany, "Optimal and approximate approaches for deployment of heterogeneous sensing devices," *Eurasip J. Wirel. Commun. Netw.*, vol. 2007, 2007.

[47]   A. Preis and A. Ostfeld, "Genetic algorithm for contaminant source characterization using imperfect sensors," *Civ. Eng. Environ. Syst.*, vol. 25, no. 1, pp. 29–39, 2008.

[48]   Z. E. Nicklow John, Reed Patrick, Savic Dragan, Dessalegne Tibebe, Harrell Laura, Chan-Hilton Amy, Karamouz Mohammad, Minsker Barbara, Ostfeld Avi, Singh Abhishek, "State of the Art for Genetic Algorithms and Beyond in Water Resources Planning and Management," *J. Water Resour. Plan. Manag.*, vol. 136, no. 4, pp. 412–432, 2010.

[49]   S. E. Haupt, G. S. Young, and C. T. Allen, "Validation of a receptor-dispersion model coupled with a genetic algorithm using synthetic data," *J. Appl. Meteorol. Climatol.*, vol. 45, no. 3, pp. 476–490, 2006.

[50]   S. E. Haupt, "A demonstration of coupled receptor/dispersion modeling with a genetic algorithm," *Atmos. Environ.*, vol. 39, no. 37, pp. 7181–7189, 2005.

[51]   C. T. Allen, G. S. Young, and S. E. Haupt, "Improving pollutant source characterization by better estimating wind direction with a genetic algorithm," *Atmos. Environ.*, vol. 41, no. 11, pp. 2283–2289, 2007.

[52]   P. E. Bieringer *et al.*, "Automated source term and wind parameter estimation for atmospheric transport and dispersion applications," *Atmos. Environ.*, vol. 122, pp. 206–219, 2015.

[53]   K. J. Long, S. E. Haupt, and G. S. Young, "Assessing sensitivity of source term estimation," *Atmos. Environ.*, vol. 44, no. 12, pp. 1558–1567, 2010.

[54]   A. Kumar, S. Dixit, C. Varadarajan, A. Vijayan, and A. Masuraha, "Evaluation of the AERMOD dispersion model as a function of atmospheric stability for an urban area," *Environ. Prog.*, vol. 25, no. 2, pp. 141–151, 2006.

[55]   A. S. Rood, "Performance evaluation of AERMOD, CALPUFF, and legacy air dispersion models using the Winter Validation Tracer Study dataset," *Atmos. Environ.*, vol. 89, pp. 707–720, 2014.

[56]   A. Berchet *et al.*, "A cost-effective method for simulating city-wide air flow and pollutant dispersion at building resolving scale," *Atmos. Environ.*, vol. 158, pp. 181–196, 2017.

[57]   A. Berchet, K. Zink, D. Oettl, J. Brunner, L. Emmenegger, and D. Brunner, "simulations over the city of Zürich , Switzerland," vol. 2, no. 2, pp. 3441–3459, 2017.

[58]   F. Pasquill, "The estimation of the dispersion of windborne material," *Met. Mag.*, vol. 90, no. 1161, pp. 33–49, 1961.

[59]   D. O. Martin, "Comment On 'The Change of Concentration Standard Deviations with Distance,'" *J. Air Pollut. Control Assoc.*, vol. 26, no. 2, pp. 145–147, Feb. 1976.

[60]   D. B. Turner, "Workbook of Atmospheric Dispersion Estimates; 2nd ed;" 1994.

[61]   C. D. Ahrens and R. Henson, *Meteorology Today: An Introduction to Weather, Climate, and the Environment*, 12th ed. Cengage Learning, Inc, 2017.

[62]   D. B. Turner, "A diffusion model for an urban area," *J. Appl. Meteorol.*, vol. 3, no. 1, pp. 83–91, 1964.

[63]   USEPA- United States Environmental Protection Agency, "Meteorological Monitoring Guidance for Regulatory Modeling Applications," *Epa-454/R-99-005*, p. 171, 2000.

[64]   D. Hadka, "Beginner ' s Guide to the MOEA Framework," p. 214, 2017.

[65]   D. Hadka, "Borg : An Auto-Adaptive Many-Objective Evolutionary Computing Framework," vol. 21, no. 2, pp. 231–259, 2013.

[66]   E. K. Burke *et al.*, "Hyper-heuristics: A survey of the state of the art," *J. Oper. Res. Soc.*, vol. 64, no. 12, pp. 1695–1724, 2013.

9   Appendices

9.1   Appendix A

Table 2. Standard deviation of wind azimuth angle (σ_A) for initial estimate of Pasquill-Gifford stability category. For use with Table 3 [63].

Initial estimate of P-G stability category	Standard deviation of wind azimuth angle σ_A
A	$22.5 \leq \sigma_A$
B	$17.5 \leq \sigma_A < 22.5$
C	$12.5 \leq \sigma_A < 17.5$
D	$7.5 \leq \sigma_A < 12.5$
E	$3.8 \leq \sigma_A < 7.5$
F	$\sigma_A < 3.8$

9.2   Appendix B

Table 3. Wind speed adjustments for determining final estimate of Pasquill-Gifford stability category from σ_A. Used with Table 2 [63].

Initial estimate of P-G Category		10-meter wind speed (m/s)	Final estimate of P-G Category
Daytime	A	$u < 3$	A
	A	$3 \le u < 4$	B
	A	$4 \le u < 6$	C
	A	$6 \le u$	D
	B	$u < 4$	B
	B	$4 \le u < 6$	C
	B	$6 \le u$	D
	C	$u < 6$	C
	C	$6 \le u$	D
	D, E, or F	ANY	D
Nighttime	A	$u < 2.9$	F
	A	$2.9 \le u < 3.6$	E
	A	$3.6 \le u$	D
	B	$u < 2.4$	F
	B	$2.4 \le u < 3.0$	E
	B	$3.0 \le u$	D
	C	$u < 2.4$	E
	C	$2.4 \le u$	D
	D	ANY	D
	E	$u < 5$	E
	E	$5 \le u$	D
	E		
	F	$u < 3$	F
	F	$3 \le u < 5$	E
	F	$5 \le u$	D

.

9.3   Appendix C
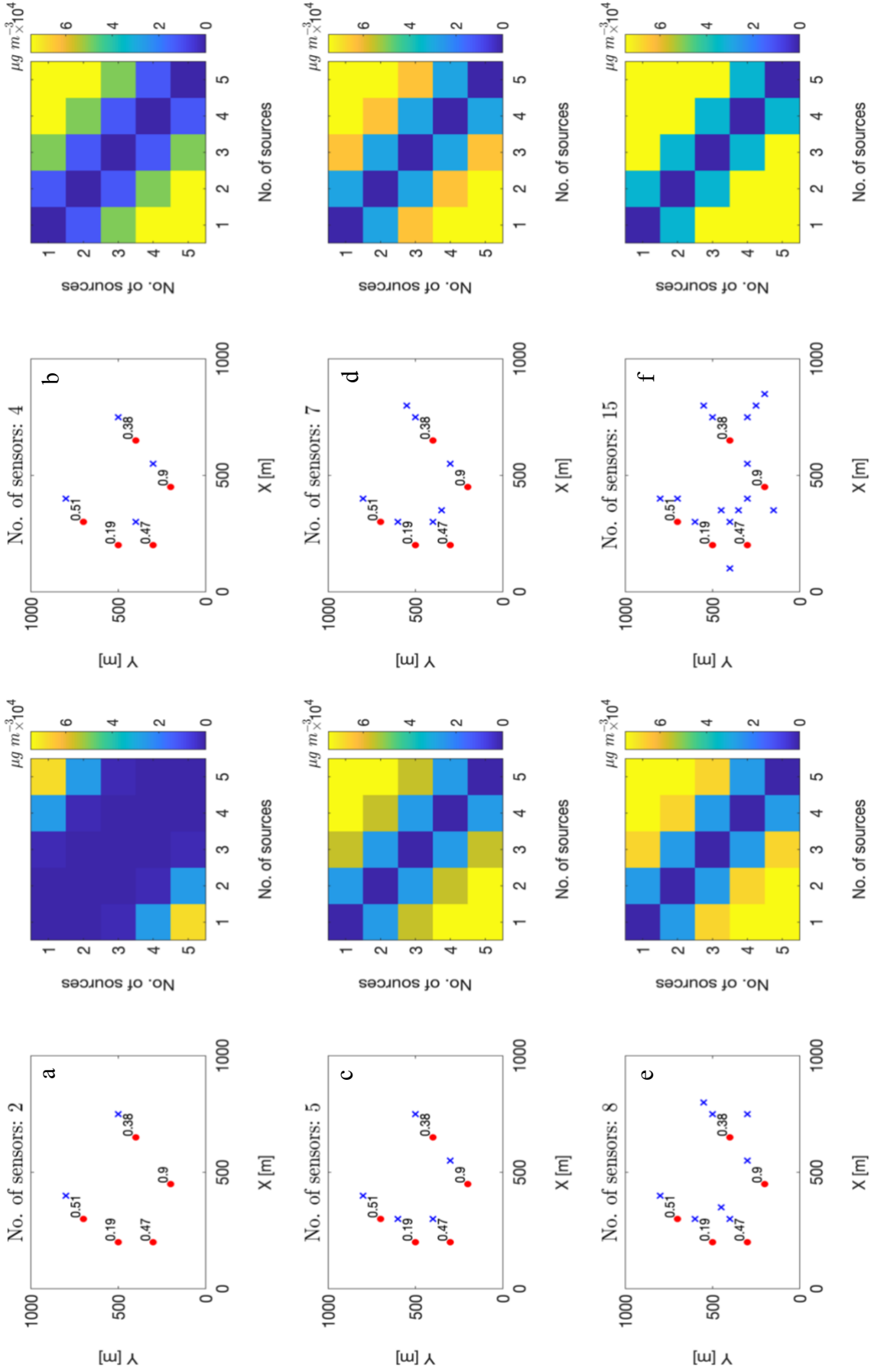
Figure 7. Sample solutions for **2 (a)**, **4 (b)**, **5 (c)**, **7 (d)**, **8 (e)** and **15 (f)** sensors obtained from one optimization run, using number of sensors ($\Psi_{\#r}$) as the first objective (Eq. 6) and average minimal PED values ($\Psi_{minimal\_PED}$) as the second objective (Eq.7). Sources' locations are marked in red dots and sensors' locations in blue crosses. Average emission rate is specified next to each source [μg m⁻³]. To the right of each displayed solution is attached a colormap of minimal PED values of each transformation pair between active sources.