

Pre-Thesis(research proposal?)

In fulfillment of the requirements for the PhD candidacy examination

Optimal deployment in time and space of heterogenous sensor array for better representation of air pollution and exposure analysis

פריסה אופטימלית בזמן ובמרחב של רשת חיישנים הטרוגנית לייצוג טוב יותר של זיהום האוויר והערכת חשיפה

**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**

TBD.

# Introduction

## 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, wildfires or vegetation. 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, NO2; carbon monoxide (CO) and dioxide (CO2); 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 the ozone (O3), produced naturally in the troposphere by photochemical oxidation of primary pollutants and mostly a summer problem in urban environments [2]. Many other anthropogenic pollutants are emitted from various industrial processes, among them are ammonia, methane and fluorinated gases, emitted for example from the fertilizer 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. Pollutants with the strongest evidence for public health concern include fine and ultrafine particulate matter (less than 2.5 and 0.1 in diameter, respectively), ozone, nitrogen dioxide and sulfur dioxide [5]. 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].

## 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 a regular basis by regulatory authorities. However, these tend to suffer from a few apparent flaws: i) their span 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 a higher spatial resolution of the desired region of research, but for a limited time only, and thus, would fail to account for seasonality or any phenomenon that presents longer cycles than the time of the campaign.

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 large numbers with far less operational requirements in various sites. These properties enable covering a larger area and obtaining a 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 has been shown that they are able to effectively capture the spatial and temporal variability of the pollution [10], [11], [13]–[15].

## Air pollution representation in space

Spatial interpolation methods (Kriging, inverse distance weighting (IDW), 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, and possibly as predictors in time as well [17]. Nevertheless, atmospheric transport and dispersion (ATD) models are the only models that can quantify the deterministic relationships between sources’ emissions and concentrations in space and in time. 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 [18].

## Source term estimation

For an accurate forecast, 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. PRTRs however, usually contain only a fraction of the actual emissions. In addition, reliability is often questionable, not necessarily because of any bias, but due to inherent difficulties in the application of measurement techniques [19]–[21].

As a consequence, sensor measurements are often being used to determine source parameters in an inverse modelling problem of source-term estimation (STE) [22]. 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.

## Sensor network deployment

When resources are unlimited, the challenge of representing well the pollution field or sufficiently estimate source parameters can be simple, requiring only the vast deployment of sensors in space to reach 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 preferable cost.

The following proposal offers a multi-objective optimization model 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 representation of the pollution field is obtained. The offered tool can serve stakeholders when either establishing a new network or planning a span of sensors in the case of an increase in industrial gas emissions or during a routine sampling task. 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.

# Research objectives (TBD)

היפותזה ומטרת המחקר (רצוי להציג את מטרת העל ולפרט את מטרות המשנה הנגזרות ממנה).

The general objective of the proposed work is to find a set of optimal solutions for the deployment of a WDESN, in order to obtain early warning in the case of increase in industrial gas emission or for generating pollution dense maps. The set of solutions will be obtained using multi-objective optimization techniques that balance between detection capabilities of the sensors (e.g. sensitivity and dynamic range) and operational requirements (e.g. cost, trained personal, protection gear and physical field constraints).

such as in the case of a complex industrial site, where access to the leak source might require significant amount of resources such as trained personal, protection gear, etc.

The more specific objectives are:

Operational requirements –

* This is especially important in the case of complex industrial site, where access to the leak source might require significant amount of resources such as trained personal, protection gear, etc.
* Source term estimation is a crucial step in assessing the pollution dense map and the resulted impact on the environment. It is also an essential step for taking mitigating acts – mainly allocating resources to stop the gas leak. These tasks get complicated as the number of possible leaks in a certain area is increased especially with a sparse sensor network. It is also noted that in many industrial sites sensors locations may be limited due to economic and operational considerations.
* All the aforementioned statistical and optimization studies seek to minimize a single objective, the discrepancy between the estimated and measured pollution levels. While commonly used, this approach is inherently limited because it assumes that accuracy is the only objective important to decision makers. In contrast, multiobjective optimization allows decision makers to simultaneously optimize across multiple performance criteria discover sets of optimal tradeoffs. Decision makers may then apply preferences across objectives *a posteriori*, overcoming cognitive myopia to discover alternatives that may not have been apparent prior to the analysis.

The cost function will be calculated according to the number of sensors and their attributes (e.g. sensitivity and dynamic range) and locations. The latter element in this objective is critical in many cases as the deployment of sensors close to a leak may be costly due to additional operational costs to cope with the expected harsh environment, but, on the other hand, will enable using a less sensitive sensor that will probably be cheaper.

* Design a *spatial* optimization model that computes sensor deployments which captures the coverage of pollution time-variation of weather conditions and emission rates.
* Design a *temporal* optimization model that computes sensor deployments, which minimizes the change between two deployments.
* Design a spatio-temporal optimization model. In this phase, the findings of the two previous objectives are integrated, so not only current conditions are considered, but the network deployment is sought, so it is most probable that for future deployments minimum changes are required.
* Conduct a set of simulations to derive engineering insights for effective deployments of air pollution sensors at various leak / sensor configurations representing possible cases.

# Research contribution (TBD)

חשיבות המחקר ותרומותיו מבחינה מדעית/עיונית והנדסית/יישומית.

# Literature review

## The problem of sensors placement

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 [23]. ﻿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 [24]. ﻿Zhang and Liu [25] 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 coverage area and data accuracy only, since most MSUs in use today are independent units which transmit data directly to a centralized computer [26], [27]. Hence, the following review regards works that dealt with optimizing the property of coverage as the main objective.

## The application of the network

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 [22] 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. [28] used population **exposure assessment** as an objective. In their method, they first generate a “demand surface”, which represents the spatial variability in pollution concentrations. Then, their “demand surface” is modified to also consider the density of a population of interest, such as children or elderly. Eventually, the “demand surface” is used as input to an algorithm that solves ﻿a constrained optimization problem from the general family of location-allocation problems, for a predefined number of air pollution sensors. Their methodology however is limited to already monitored areas, since it relies on existing measurements from monitoring stations. Lerner et. al. [27] developed a method to optimally deploy a network of low-cost sensors of two types for the general purpose of **environmental monitoring** of ozone, NO and NO2 emissions. The optimal locations were found using an optimization process that seeks for the set of locations, constrained by available locations and a given budget, that maximizes the overall utility of the sensor network. The latter is comprised of the suitability of the type of sensor to the location of deployment and the rank of that location, relative to its surrounding. Lerner et. al. focused on the land use of the region of interest for specific atmospheric conditions.Carter and Ragade [29] proposed a probabilistic model which takes 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 utilizes 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 physical base. Boubrima et. al. [30] 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, sensors are eventually placed at the most common pollution zones and only one objective is applied – minimizing the cost of the network.

In another work, Boubrima et. al. [31] 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. [32] wished 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 considers 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 evaluates 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 ﻿considers locations that improve prediction uncertainty and provide high prediction accuracy. Their method however requires a robust preliminary deployed network that already captures the spatial variability adequately, 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.

Another aspect of network deployment dealt with in many works, is the use of a network of sensors mounted on mobile platforms [22], which enables a real-time adaptive deployment of sensors according to the spreading of the pollutant. Kuroki et al. [33] developed an expert system for navigating unmanned aerial vehicles for optimal contamination mapping that enables estimation of source parameters. Belkhiri et. al. [34] proposed several deployment models, that simulates air pollution concentrations as ground truth and estimates concentrations where no sensors are deployed using IDW interpolation. Sensors are redeployed so estimation error (i.e., ﻿the absolute difference between ground truth and estimation) is minimized and execution time is considered.

## Optimization aspect

﻿Another aspect to be considered is the optimization method used to solve the problem of network deployment. A number of algorithms have been presented in the literature, some of them were used in the above reviewed works. These include: Integer linear programing [30], [34], Location-allocation methods [28], gradient-descent based methods (cite) and Meta-heuristic optimization algorithms such as simulated annealing and evolutionary and genetic algorithms (cite). Meta-heuristic optimization algorithms are considered global search algorithms and their uniqueness is in alternating parameter estimates to generate new solution candidates. They benefit over other methods as they can handle poor initial estimates and employ methods to prevent becoming stuck in local minima [22], and were shown to provide near-optimal results in many studies (e.g., [29], [35]). It should be noted that the optimization techniques may be implemented as a single objective scheme (usually minimizing the cost) or a multi-objective scheme.

# Methods and Research plan

## The Gaussian plume model

Atmospheric transport and dispersion (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 [36]. 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 [22].

The Gaussian plume model is one of the simplest and widely used models that offers an analytical 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 (see illustration is Figure X). The advection-diffusion equation can be derived by the mass conservation equation for the pollutant concentration C [kg/m3]:

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

where S [kg/m3s] is a source or sink term and the vector function represents the mass flux [kg/m2s] of a pollutant owing to the combined effects of diffusion and advection. ﻿The diffusive flux () is proportional to the concentration gradient through the diffusion/turbulence coefficient K [m2/s]:

|  |  |  |
| --- | --- | --- |
|  |  | (2) |

The second contribution to the flux is due to simple linear advection by the wind, which can be expressed as:

|  |  |  |
| --- | --- | --- |
|  |  | (3) |

Substitution Eq.2 and Eq.3 into the equation of mass conservation (Eq. 1) yields the three-dimensional advection-diffusion equation:

|  |  |  |
| --- | --- | --- |
|  |  | (4) |

The Gaussian plume model, which is the solution of the three-dimensional advection-diffusion equation (Eq. 4), eventually describes the pollutants’ concentration C [kg/m3] in a certain position in space:

|  |  |  |
| --- | --- | --- |
|  |  | (5) |

Where Q [kg/sec] 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 [m/s] is the mean wind speed at the height h of the release. and [m] are key parameters in the model and represent the standard deviation of the Gaussian concentration distribution in the crosswind and vertical direction (Figure X).

For simplicity, the solution steps of Eq. 4 were not specified. However, it is important to mention the simplifying assumptions that had to be made in order to reach the Gaussian plume model equation (Eq. 5) [36]:

1. The pollutant is emitted in a constant rate Q from a single point source located at height H above ground surface.
2. The wind velocity is constant and aligned with the positive x-axis.
3. The solution is a steady state.
4. The diffusion/turbulence coefficient K is a function of the downwind distance x only, and diffusion is isotropic so that .
5. Wind velocity is sufficiently large that diffusion in the x-direction is much smaller than advection.
6. Variations in topography are negligible so that the ground surface can be taken as the plane z=0.
7. The pollutant does not penetrate the ground.

To these assumptions we add that the material diffused is a stable gas or aerosol, with a negligible deposition rate and that background pollution is negligible. parameters are used instead of the diffusion/turbulence coefficient , 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 (cite):

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, and the most used values are those obtained by Turner [37], using Pasquill-Gifford stability class categories [38]. These coefficients are specified in [Table X](http://homepages.see.leeds.ac.uk/~lecrrb/dispersion/index5.html).

## Meteorology

### Background

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, 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. If the temperature of the atmosphere increases with altitude in a certain layer in the atmosphere, then this layer is called an **inversion,** an absolutely stable condition. 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**. In this condition, temperatures decrease slightly with height, in a rate close to the dry adiabatic rate (about 100 for every 1000 m).

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 meter 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], [37], [39].

### Determining the atmospheric stability class

In order to apply the suitable parameters to the Gaussian plume model, the Pasquill-Gifford stability category which characterizes the simulated set, should first be determined. The original method to do so, developed by Turner [40], requires knowledge on cloud cover and cloud ceiling. Alternative methods were developed for situations where these data are not available. They include a radiation-based method which uses measurements of solar radiation during the day and delta-T at night and turbulence-based methods which use wind fluctuation statistics [41]. For the current simulation, we chose to use a turbulence-based method ( method), because of its simplicity and our currently available data. ﻿The method (sometimes referred to as ) uses the standard deviation of the wind direction () in combination with the scalar mean wind speed (), to determine the Pasquill-Gifford stability category (see Tables X and X). Wind direction standard deviation values of 10-min resolution were obtained from the Israel Meteorological Service (see section X). To minimize the effects of wind meander (﻿long period oscillations associated with light wind speed conditions), it is recommended to calculate the 1-hour value using 10-min or 15-min averages, as specified in Eq.6 (i.e., calculating the root mean square) [41].

|  |  |
| --- | --- |
|  | (6) |

## Evolutionary algorithms and Borg MOEA (TBD)

אלגוריתם אבולוציוניים

כלי האופטימיזציה בורג

## Methodology – problem formulation and optimization

### General formulation

Let be the region of interest, i.e., the industrial area and its surroundings, where we wish to place our network of sensors in. Let be the set of sources, where each source, is located in and for a specific time t, emits [Kg/sec]. Similarly, the set is the set of sensors, where each sensor, , is located in and records a pollution level of . The location of the sources are known, these are the locations of the industrial plants.

Let be the pollution transfer function of the dispersion model, which associates sensor’s r readings, , with the emissions of source s. Thus, the model’s estimated contribution of source to the pollution level in , is given by:

|  |  |
| --- | --- |
|  | (7) |

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, , i.e.:

|  |  |
| --- | --- |
|  | (8) |

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

### Pairwise Euclidean distance (PED)

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 computed using (9):

|  |  |
| --- | --- |
|  | (9) |

In a generated network set that is able to distinguish well between overlapping plumes (if given a source term estimation problem), a change in the number of active sources is expected to have a dramatic effect on the sensor network readings, resulting in high PED values. On the other hand, low PED values will characterize a network of sensors that will not be able to sufficiently resolve a problem of source separation and estimation. An example of a matrix of mean PED values is shown in Figure X for different combinations of 1-5 active sources.

### Problem formulation

In our problem, we aim to find an optimal deployment (number of sensors and their locations?) of a set of sensors . In general, when the number of sensors () increases, the PED values, calculated between two sets of active sources with different sizes {S}' and {S}", increases as well (see Eq. 9). Given that the span of a set of static 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 form the highest PED value, which represents the best ability of those sensors to separate the different sources. The first objective is then to find the minimal number of sensors deployed:

|  |  |
| --- | --- |
|  | (10) |

It is possible to phrase several different objective functions that maximizes the PED value. Our algorithm seeks to find the optimal set of sensors , so that the percentage of cases where PED values are greater than a certain threshold is maximized, considering all possible combinations of two sets of number of active sources with different sizes {S}' and {S}". For example, in the case of 5 sources, 31 options of 1-5 active sources exist, creating a total number of possible combinations (NPC) 355 of two sets of number of active sources with different sizes {S}' and {S}". The second objective is::

|  |  |
| --- | --- |
|  | (11) |

The network of sensors deployed should consider the varying meteorological conditions. Hence, our objective takes into account the discrete frequencies of each possible meteorological combination of wind speed, wind direction and Pasquill–Gifford atmospheric stability class. Hourly mean wind directions were binned into 16 categories (22.5o each) and hourly mean wind speeds into 7 categories (<1, 1-2, 2-4, 4-6, 6-8, 8-10, >10 m/s). With six possible Pasquill–Gifford classes for atmospheric stability (A-F), 672 weather states exist, out of them only X were found to be physically meaningful according to the data archive used (see section X).

All possible situations (X weather states with all options of active/non-active sources) are applied in the Gaussian plume model to obtain steady state solutions of a pollution concentration map. Each concentration map is then given a weight according to the frequency distribution derived.

### Tri-objective problem

The second objective can be separated to two different objectives, one which considers PED values obtained during **night** hours and one which considers PED values obtained during the **day**. This forms a tri-objective optimization problem and requires creating a separated frequency distribution of weather states for day and for night, in the same way described above.

### Optimization procedure

Once we obtained all weighted maps of concentrations, the optimization process starts. The Borg MOEA algorithm (see section X) searches for the optimal set of sensors by starting with an arbitrary guess. Then the algorithm projects the set on and evaluates all possible PEDvalues of that set (Eq. 9). Then, based on the two or three objectives ( and ), the Borg refines the set of decision variables, to locate a set of solutions on the Pareto frontier. Technically, we assume that each location may host a sensor. The decision variables are held in a data array, each value represents a sensor, linked to a certain location . The sensor is “placed” in that location if the decision variable is 1 and is not “placed” if the decision variable is 0.

### Simulation set

We simulated a 1000x1000 meter flat area. The simulated sensors were assumed to be situated at ground level, and so Eq. 5 was reduced to:

|  |  |
| --- | --- |
|  | (12) |

Five point-sources (stacks) were given average yearly emission rates of 1000, 1500, 600, 1900 and 300 . All model runs used the same average yearly release rates of the sources, and not some momentary emission rates (), assuming these are usually not available. Figure X depicts some of the configurations. Sources are marked in red circles (o) and optional locations of sensors are marked in black crosses (+). A total of X optional locations are spread in a grid, 50 m apart. Stack heights (effective heights) were taken at 10 m, following the height of the obtained wind data.

## Meteorological Data

For our simulation set, we used 10 min wind data obtained from the Israel Meteorological Service (IMS) for the years X and for one station Nevatim/Negev junction. This station was chosen because of its relatively idealized location: a desert area with little land cover, low topography and very few pollution sources except for Ramat-Hovav industrial zone. In future work, we may wish to implement our methodology on a real-world problem, possibly in that area.

# Research plan (TBD)

* **Evaluation and Validation** – using real-world data sets or generate synthetic data of emissions.
* **Improve the modelling tool** – use as a constraint or as an objective, the demand for solving the STE problem at least for at least X percent of the scenarios and of maximum deviations from emission permits.
* **Create more complex scenarios** – allow different types of sensors with different sensitivities and dynamic ranges. possibly consider different pollution sources.
* **Conduct measurements**
* **Chicago database** - Evaluate how well the current placement of sensors is, suggest where to place additional sensors.
* **Mobile sensors and online deployment**

# Initial results (TBD)

# Work schedule (TBD)

# 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-2018].

[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] 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.

[19] 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.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

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

[25] 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.

[26] 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.

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

[28] 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.

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

[30] 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.

[31] 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.

[32] 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.

[33] 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.

[34] 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.

[35] 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.

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

[37] D. B. Turner, “Workbook of Atmospheric Dispersion Estimates; 2nd ed;” 1994.

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

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

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

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