A simulation using a synthetic set of stacks was performed in order to evaluate the proposed method for sensors’ placement. Usually, several constraints exist when placing static sensors; i) law requires sensors to be deployed in populated areas, i.e., inside neighborhoods, ii) placing sensors inside the industrial region could have an effect on the network ability to separate between sources (not sure?), iii) other physical constraints that may characterize the given study area (vegetation, public buildings).

This is especially important in the case of complex industrial site, where access to the leak source might require significant amount of resource such as trained personal, protection gear etc.

Basic idea:  using PED, using winds and atmospheric stability states that characterize the area.

**Simulation 1:** all sensors are the same and have perfect ( detection threshold is zero ?) sensitivity and infinite dynamic range (?).

**Simulation 2:** running a simulation for reasonable sets of detection thresholds and dynamic ranges. See what type of sensor suits.

**Simulation 3:** ~~change the density of the stacks inside a defined area of constant size.~~ Run for average yearly emissions Vs. table of factors emissions (my problem is with stacks that don’t work and then my PED calculation is for 5 and not less stacks). (in a different simulation I take the hourly averages taken from the ministry of environmental protection?).

Assumptions: the optimization process is based on average emission rates of the stacks.So basically we don’t take into consideration activity hours of the plants. These are indirectly considered during the optimization process by PED calculations. Technically we could correlate weather conditions with stacks’ working hours.

Such work may serve regulators and stakeholders.

**What to display:**

1. Pollution dense map for an average wind direction, speed, and average emission rates of stacks for substance X.
2. Distribution of wind speed and direction of the area.
3. Distribution of stability classes.
4. Assuming a sensor in each node of the grid, what is the distribution of my concentrations?
5. Show a map of the weighted PED values (maybe higher resolution than the grid I run).
6. Show a map of solutions for X sensors - an average of 10 runs. Is it clustered in groups?
7. Perform random seed analysis
8. Show the average pareto front of 10 runs.
9. Show pareto fronts of different sensors types.
10. Results for different density of stacks.
11. What is the PED of my set of sensors I just chose, for different emission scenarios?
12. להראות את מפת הריכוזים הממוצעים לעומת מפת הPED. לאו דווקא PED גבוה קורלטיבי עם ריכוזים גבוהים...

**הערות**

זה אולי לא המודל הכי מדויק אבל לצורך הבעיה שלנו הוא יכול להיות מספיק בהחלט בגלל שמסתכלים על ממוצעים שנתיים.

*The pollutant concentrations in the plume are inversely proportional to the wind speed along the x axis, as can be seen from equation X. For example, if we double the wind speed, the concentration in the plume would be half of that with a lighter wind.*

*Incorporating considerations of both…*

*השיטה של מהירות הרוח בעצם לוקחת בחשבון יותר טורבולנציה מכאנית בפני הקרקע. החלוקה ליום ולילה*

**איורים**

A close up of a map

Description automatically generated

A close up of a map

Description automatically generated Pasquill stability class categories: A screenshot of a cell phone

Description automatically generated

The criteria in Tables X and X are for data collected at 10 m and a roughness length of 0.15 m. Wind

A screenshot of a cell phone

Description automatically generated

A screenshot of a social media post

Description automatically generated

קטעים אולי לזרוק

In the past couple of years, citizen science projects had become more common, and large amounts of air quality data are being collected today by individuals, usually using low-cost sensors of various types (cite). Using such data however requires an exhaustive preprocessing work as reliability of some of the measurements might be questionable.

Most plants are given a certain emission permit that refer to a certain average emission rate. Certain standard deviations are allowed, with a unique value for each pollutant. Enforcement of these permits relies on the ministry’s inspection.

אולי מעניין להגיד ששימוש נוסף יכול להיות עבור קמפיינים זמניים בשביל לשפר שיטות של קרינגינג וכאלה.

מממ... למה בעצם אנחנו לא לומדים מהסימולציה את הקורלציות בין הnodes השונים עבור תנאי מזג אוויר מסוימים ופליטות מסוימות. ואז, אני עושה source term estimation כשאני בעצם יודעת את הקורלציות שלי עם מקומות אחרים ויכולה להשתמש בזה כמידע עם אמינות נמוכה. :/

להגיד משהו על חיישנים?

אולי אני אייצר אחר כך מזג אוויר כדי לבחון את ההצבה שלי?

או אעשה source term estimation?

אולי אוסיף כ-constrain את ה-source-term estimation. אפשר להוסיף מגבלות גם על האזור.

צריך לזכור שמה שאני עושה זה יוצרת מפות ממוצעות של ריכוזים. אני יכולה לעשות את זה גם עם להריץ 30 שנה של דאטה.

אפשר להוסיף כ-constrain את העובדה שבהכרח אני יכולה לפתור את בעיית ה-STE.

או להכניס כobjective את היכולת לעשות STE בהכי הרבה סנריואים של חריגה.

Validation - לבדוק האם ה-PED של כל פיתרון של מספר חיישנים משמעותית גבוה מבחירה אקראית של מספר חיישנים זהה ולקחית PED

אפליקציה של העבודה בשדות אחרים

Two types of emissions exist: point source emissions, channeled through a pipe such as a chimney stack, and non-point source emissions, originate mainly from field sources such as tanks, pools or leaks from equipment (valves, taps, etc.) [34].

This produces an inherent problem in their model; e.g., in a hypothetical case where sources are located at the eastern part of the study area and typical winds are blowing from the west, placing sensors west to the sources is not effective.

Distinguishing between daytime and nighttime takes into consideration buoyant turbulence generated by … all kinds of processes….NOT SURE ABOUT THIS

טורבולנציה מכאנית (u ממוצע) – du/dz. בכיוון מעלה (w) אפשר להגיד שממוצע הפלקטואציות הוא אפס ולכן זניח.

טורבולנציה תרמית – dT/dz (T ממוצע)

אולי אוריה צריך למדוד לחות עם רחפן

הפרעות במישור איקס וואי יוצרים שובלים וטורבולנציה downwind (אותו גובה z). זה גזירת רוח במישור. בכיוון הורטיקלי כתוצאה מחיספוס עם הקרקע.

עוד תופעה שיוצרת טורבולנציה זה ציפה תרמית, כתוצאה משינוים תרמיים. רק בכיוון z.

המכאני תמיד מעודד טורבולנציה. תרמי יכול או לעודד טורבולנציה או לדכא טורבולנציה, תלוי במצב האטמוספירי.

אולי הקשר בין u ו-W שדוד הראה זה ההסבר לשיטת חישוב סיגמא A.

אטמוספירה אדיאבטית = אטמוספירה ניטרלית.

In an absolutely **unstable** atmosphere, the environmental lapse rate (the rate at which air temperature changes with altitude) is greater than the dry adiabatic lapse rate (100 for every 1000 m). 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.

In an absolutely **stable** atmosphere, the environmental lapse rate is less than the moist adiabatic rate (60 for every 1000 m) and 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, which is a form of an **absolutely** **stable** atmosphere as well. Such conditions might lead to high pollution concentrations and minimal amount of dispersion. Between the above two extremes is the condition characterized as **neutral**. In this condition, the environmental lapse rate is characterized by a slight decrease of temperature with height, close to the dry adiabatic rate. These conditions can be caused for example by transitional situations near sunrise or sunset when the atmosphere changes its state from stable to unstable or vice versa. With neutral atmospheric conditions, intermediate level of dispersion takes place

It should be mentioned that meteorologists define a **conditionally unstable** atmosphere to distinguish between the absolutely stable and absolutely unstable conditions. In this situation, the environmental lapse rate lies between the moist and dry adiabatic rates, and the stability of the atmosphere depends on whether the rising air is saturated or not. When the air parcel is not saturated, then the dry adiabatic lapse rate is the relevant reference state and the atmosphere is stable. For a saturated air parcel inside a cloud, the reference is the moist adiabatic rate and the atmosphere is unstable

**Research plan (not here)**

* Add changing sources – either as a constraint (must solve the STE problem) or as an objective – find most scenarios (out of different combinations of the 5 sources) of deviation from emission permits.
* Take into consideration background concentrations/roads.
* לשכלל את השיטה שתכלול כמה מקורות של זיהום שונה (סוגי מזהמים שונים ולכן להתאים להם סנסורים שונים)
* Add sensors of different types (sensitivity and dynamic range)
* Evaluate the deployment model on a real-world data set. ﻿+ compare to existing generic deployment strategies
* Do something with Chicago database – 1) Evaluating how well the current placement is. 2) suggest where to place additional sensors
* Perform measurements myself
* Mobile deploy of sensors
* 3D deployment scheme
* How would weather scenarios and emission variability impact the deployment results?
* Add noise and background concentrations
* תרומה – אולי אפשר לעשות רידיפלוימנט כבר לרשתות קיימות לפי שיטה זו

Dumped from 2 draft

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) [43].

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

The multitude of potential solutions with varying degrees of tradeoff between the objectives force decision makers to explore this set of potential solutions and identify the solution(s) to be implemented. While ultimately the selection of the final solution is the responsibility of the decision maker, optimization tools should assist this decision process to the best of their ability.

It is common to define the notion of Pareto optimality (cite), which considers solutions to be superior or inferior to another solution only when it is superior in all objectives or inferior in all objectives, respectively.

* Design 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. The model will capture the distribution of weather conditions that characterizes the regime and sensors of various attributes.
* Design a *temporal* optimization model that computes sensors’ redeployment strategy in case of changing weather conditions of different time scales (e.g., 1 hour, 1 day, season), 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.
* Design a *spatial-temporal* optimization model. In this phase, the findings of the two previous objectives will be integrated, and probability of change in short-term weather conditions will be considered, so it is most probable that future deployments require minimum transfer efforts.

בהנחה שאין אינטראקציה - אפשר לפתור את זה כבעיות נפרדות – עבור מקורות שפולטים X ועבור מקורות שפולטים Y (חלקם פולטים את שניהם). ואז יש לי שני סטים שונים של PED ואני ממקסמת סכום PED יחד עם ממנממת מספר חיישנים של x+y.

אם קיימת אינטראקציה, אני יכולה להניח שיש לי רק סוג אחד של חיישן (מצב א) או כמה חיישנים ולהם תכונות שונות של רגישות וקרוס-רגישות (מצב ב). במצב א, אני משתמשת רק בסוג אחד של חיישן ובידע שלי על קרוס-רגישות של החיישן בשביל להסיק על מפות PED שונות לכל סוג של מזהם (החיישן שלי מודד מצוין NO2 ומדד 10 ppm, אני מסיקה לפי קרוס-רגישות של מינוס 20 אחוז על אוזון, שיש שם 8 ppm אוזון). בגלל שהקרוס-רגישות היא קשר לינארי קבוע, אין צורך להתחשב בזה באופטימיזציה. במצב ב, אני אולי צריכה לבנות כבר פונקציית מטרה מורכבת יותר שתכלול שיקלול כלשהו של החיישנים למשל להגדיר cost שמבוסס על תכונות החיישנים ולנסות למנמם אותו.

Another aspect to be considered is the common methods used to solve the problem of network deployment. 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 [40]. In the “0-1 knapsack” problem, a subset of items, out of n items, possessing each some value and some cost , should be selected such that the sum of the values is maximized, while keeping the summed cost within some capacity . The knapsack problem is NP-complete [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.

The term global optimization refers to the process of attempting to find the solution out of a set of possible solutions {that has the optimal value for some fitness function f, such that [42]. Two main categories of methods to find exist; the first are **deterministic methods**, which find by an exhaustive search over the set { while making certain assumptions about the fitness function to avoid huge calculations. **Stochastic methods** involve random elements to determine the global optimum point, each one trying to reduce the computational burden of pure random search. ﻿At the outset, a random sample of points in the set { is picked. Then, each method manipulates the sample points in a different manner, using different **heuristics** [43]**. Heuristics** may be thought of as sets of rules for deciding which potential solution out of S, should next be generated and tested (i.e., an intelligent search in space). For some randomized heuristics, such as simulated annealing and certain variants of Evolutionary Algorithms (EAs), convergence proofs exist. The problem in these algorithms however is that they will not identify the suggested solution as being globally optimal, rather as simply the best solution seen so far. **Local search algorithms** (often referred to as **hill climbers**), such as “gradient descent”, work by taking a starting solution x, and then searching the candidate solutions in the ﻿neighboring environment for one x’ that performs better than x. Although they may be quick to identify a good solution, this process will eventually lead to the identification of a local optimum, and no guarantee can be offered for the quality of the solution found, compared to . As a result, local searches are usually incorporated in stochastic methods to yield candidate global optimum, from which the best point is eventually picked.

EAs are Meta (i.e., problem-independent)-heuristic optimization algorithms. Inspired by the biological theory, if given a population of individuals (i.e., a set of candidate solutions), the environmental pressure causes natural selection and 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 or more offsprings (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 EAs to maintain a diverse set of solutions, by creating new solutions from a non-uniform distribution, not only provides a means of escaping from one local optimum [23] and handling poor initial estimates; it provides a means of coping with large and discontinuous search spaces. As a consequence, EAs were shown to provide near-optimal results in many studies (e.g., [30], [44]).