A data-driven algorithm to optimize the placement of continuous monitoring sensors on oil and gas sites

Meng Jia, Troy Sorensen, Dorit Hammerling

Applied Mathematics and Statistics

Colorado School of Mines

COLORADO SCHOOL OF MINES
EARTH • ENERGY • ENVIRONMENT

November 15, 2023
Durham, NC

Outline

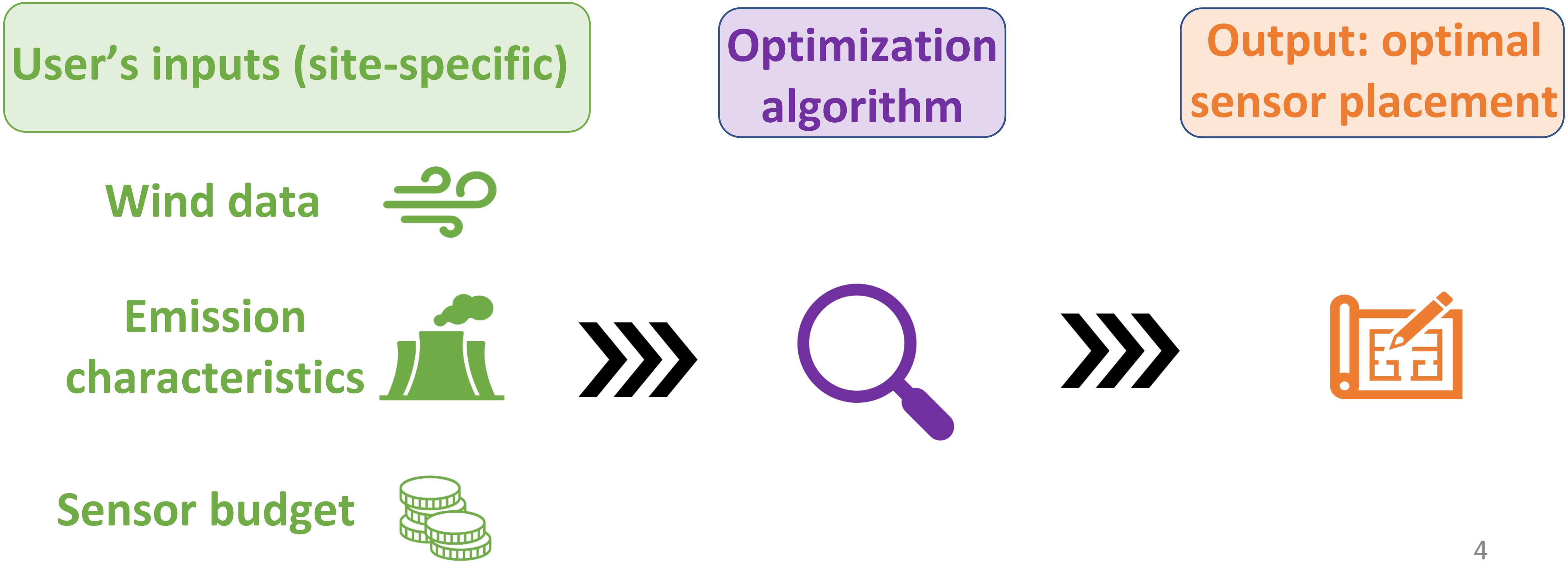
- Problem Setup
- Algorithm
- Experiments & Results
- Conclusions & Future Work

Outline

- Problem Setup
- Algorithm
- Experiments & Results
- Conclusions & Future Work

Problem Setup

- We propose a data-driven algorithm to optimize sensor placement for continuous monitoring systems (CMS).



Problem Setup

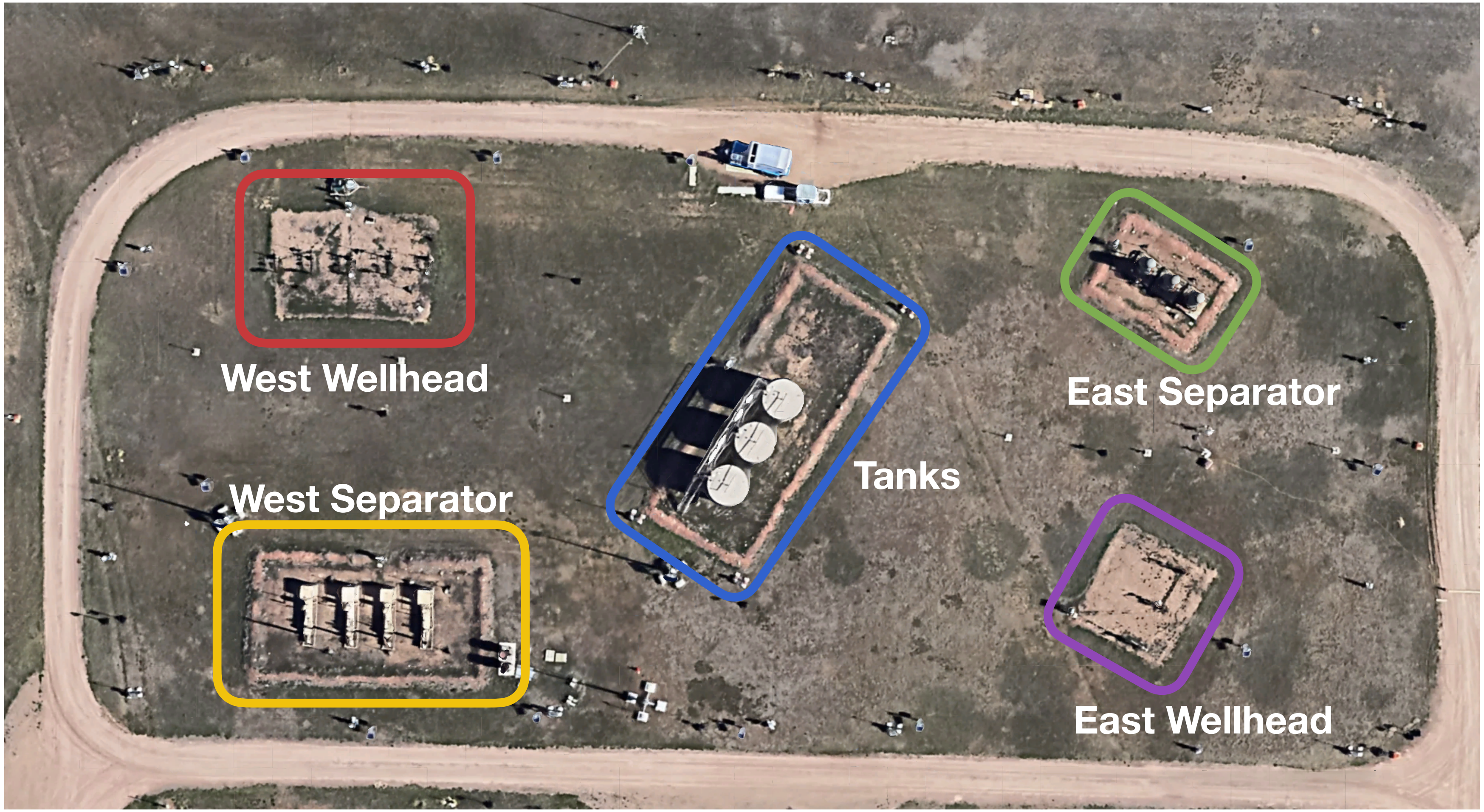
- General idea
  - Use historical wind data and available emission information to simulate M many emission scenarios.
  - Prescribe N possible sensor locations.
  - Find optimal sensor placement from all feasible configurations, under the given budget k, to maximize the detection from all M emission scenarios.

Outline

- Problem Setup
- Algorithm
- Experiments & Results
- Conclusions & Future Work

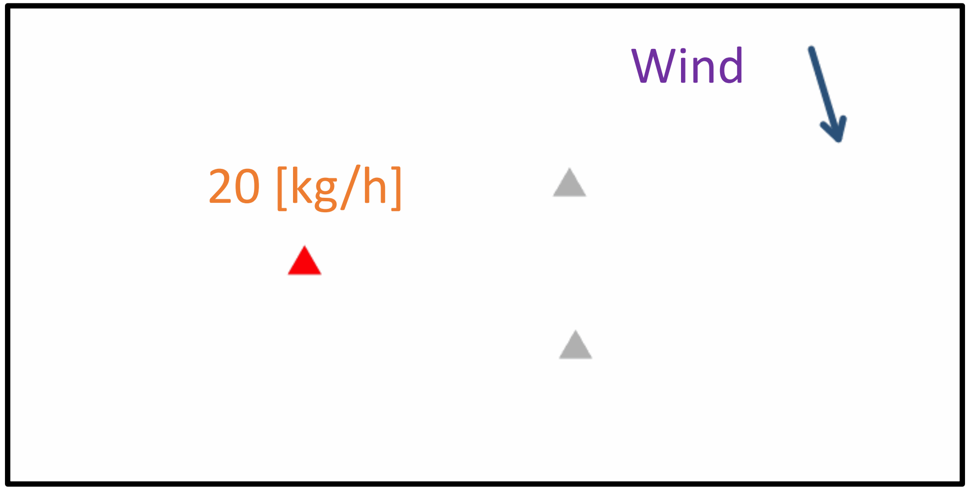
Algorithm

- Step 1: generate emission scenarios.
- Step 2: prescribe potential sensor locations and simulate concentrations.
- Step 3: check detection status.
- Step 4: solve an optimization problem to find the best sensor placement.



Methane Emissions Technology Evaluation Center (METEC), Colorado State University

Algorithm



- Step 1: generate emission scenarios

A combination of

- wind speed time series
- wind direction time series

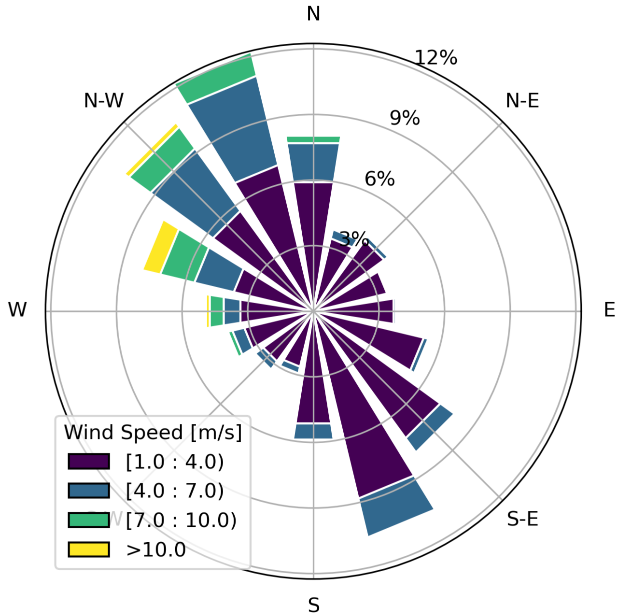
Random sample segments from historical wind time series

- emission source location
- emission rate

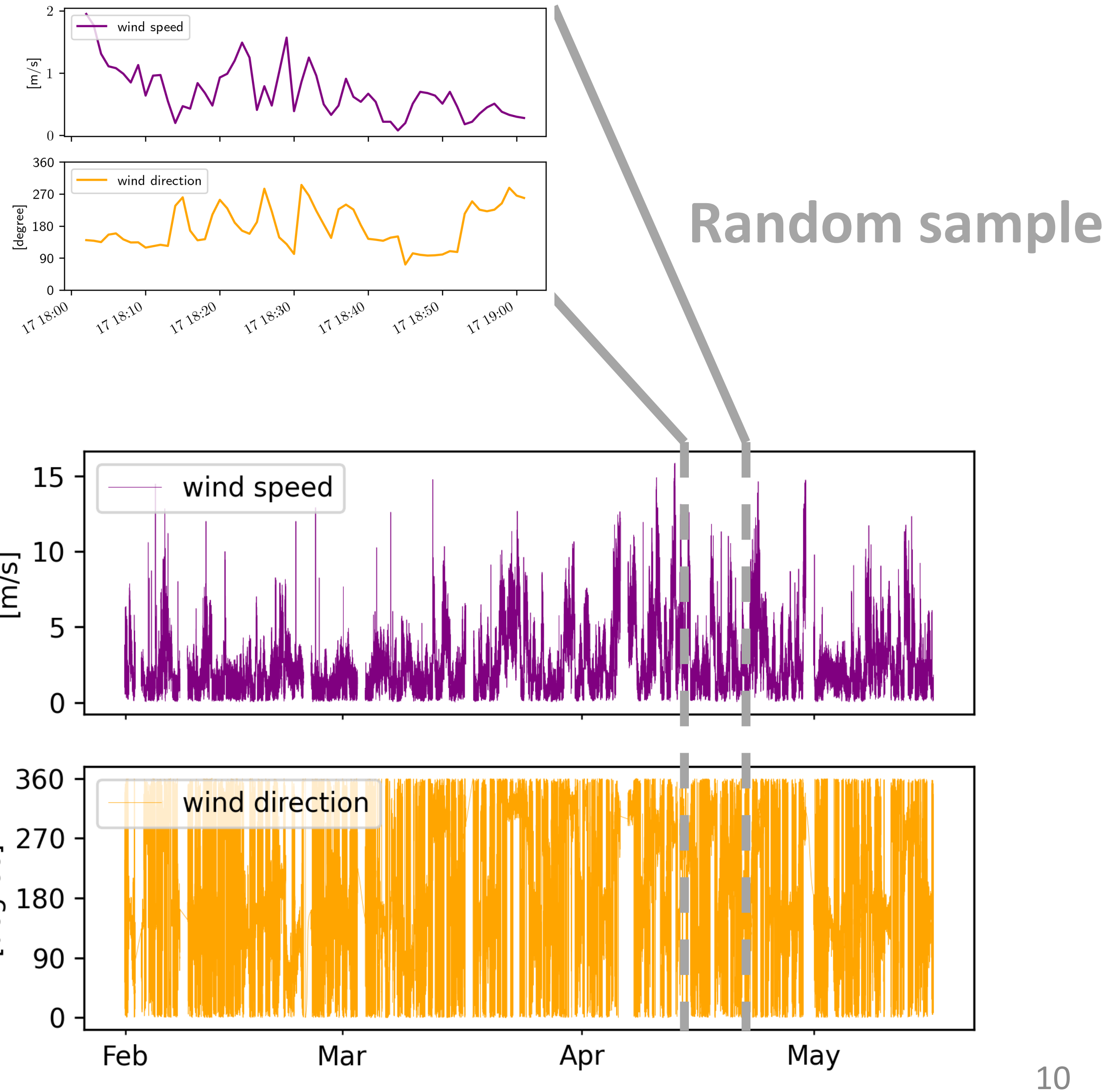
1. Estimate a joint distribution of emission location and rate from prior knowledge
2. Random sample

defines an emission scenario.

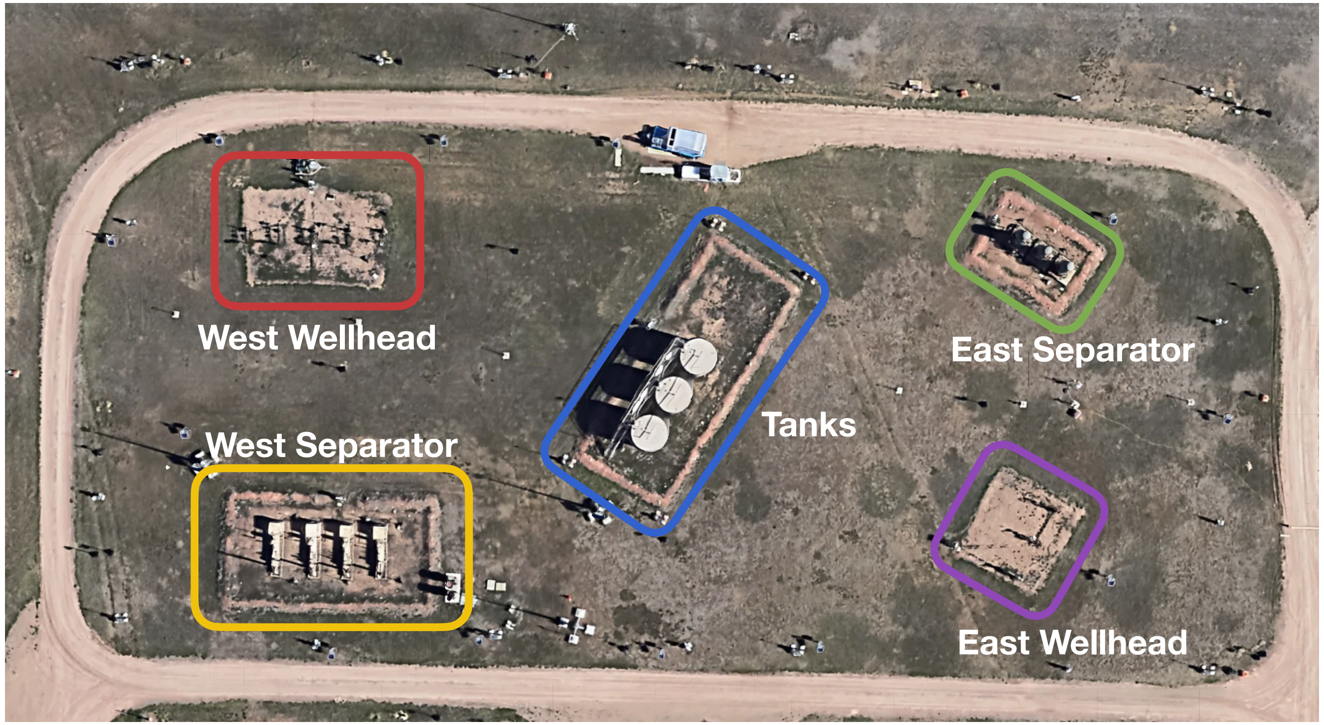
Step 1.1 Wind Data



Four-month wind distribution



Step 1.2 Emission Information



Potential emission sources on METEC

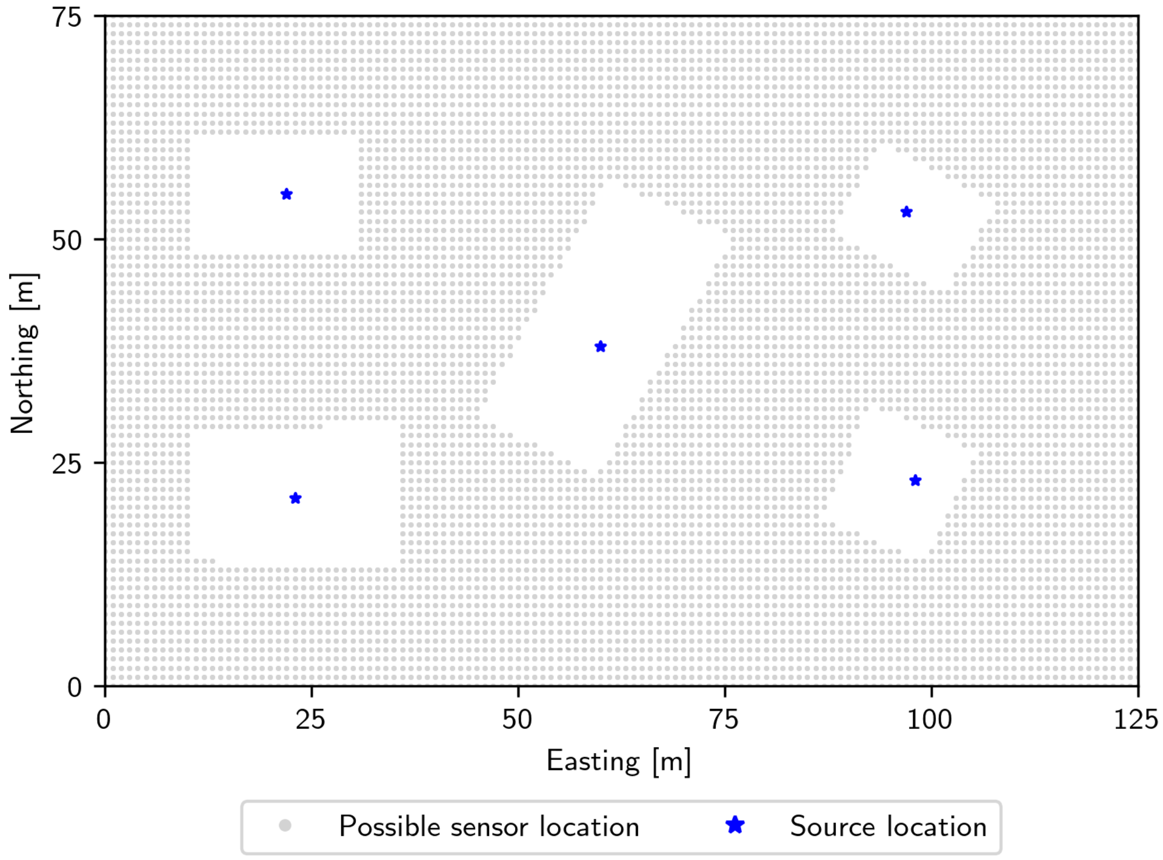
	10 kg/h	20 kg/h	50 kg/h
W.W	1/15	1/15	1/15
W.S	1/15	1/15	1/15
T	1/15	1/15	1/15
E.S	1/15	1/15	1/15
E.W	1/15	1/15	1/15

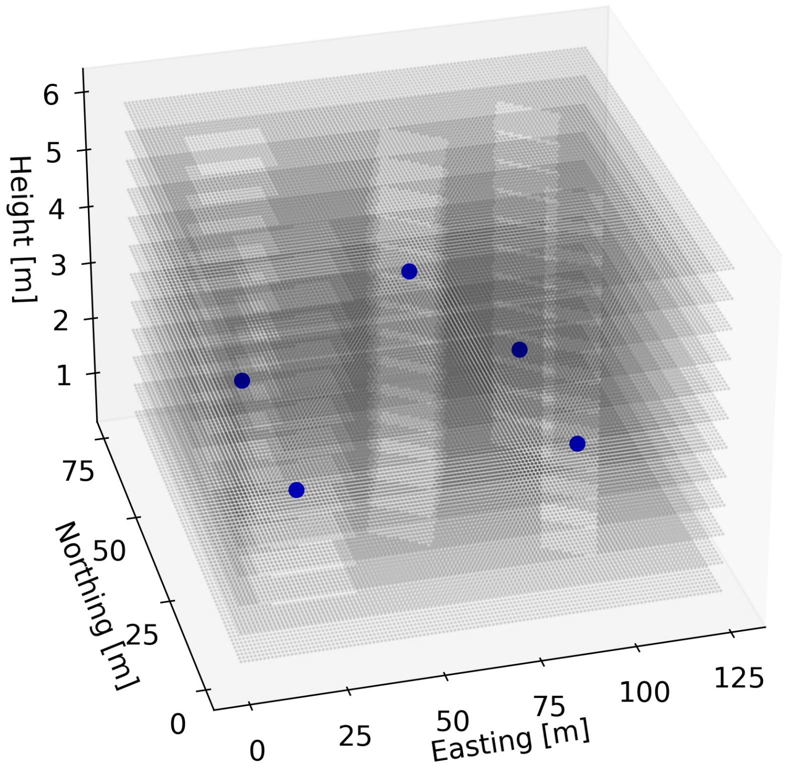
Probability of emission location and rate pair

Algorithm

- Step 2: prescribe possible sensor locations and simulate concentrations
  - Set possible sensor locations by gridding the site in 3D.
  - Filter out invalid locations.
  - For each (emission scenario, sensor location) pair, run Gaussian puff model to compute the CH4 concentration time series.

Step 2.1 Sensor locations

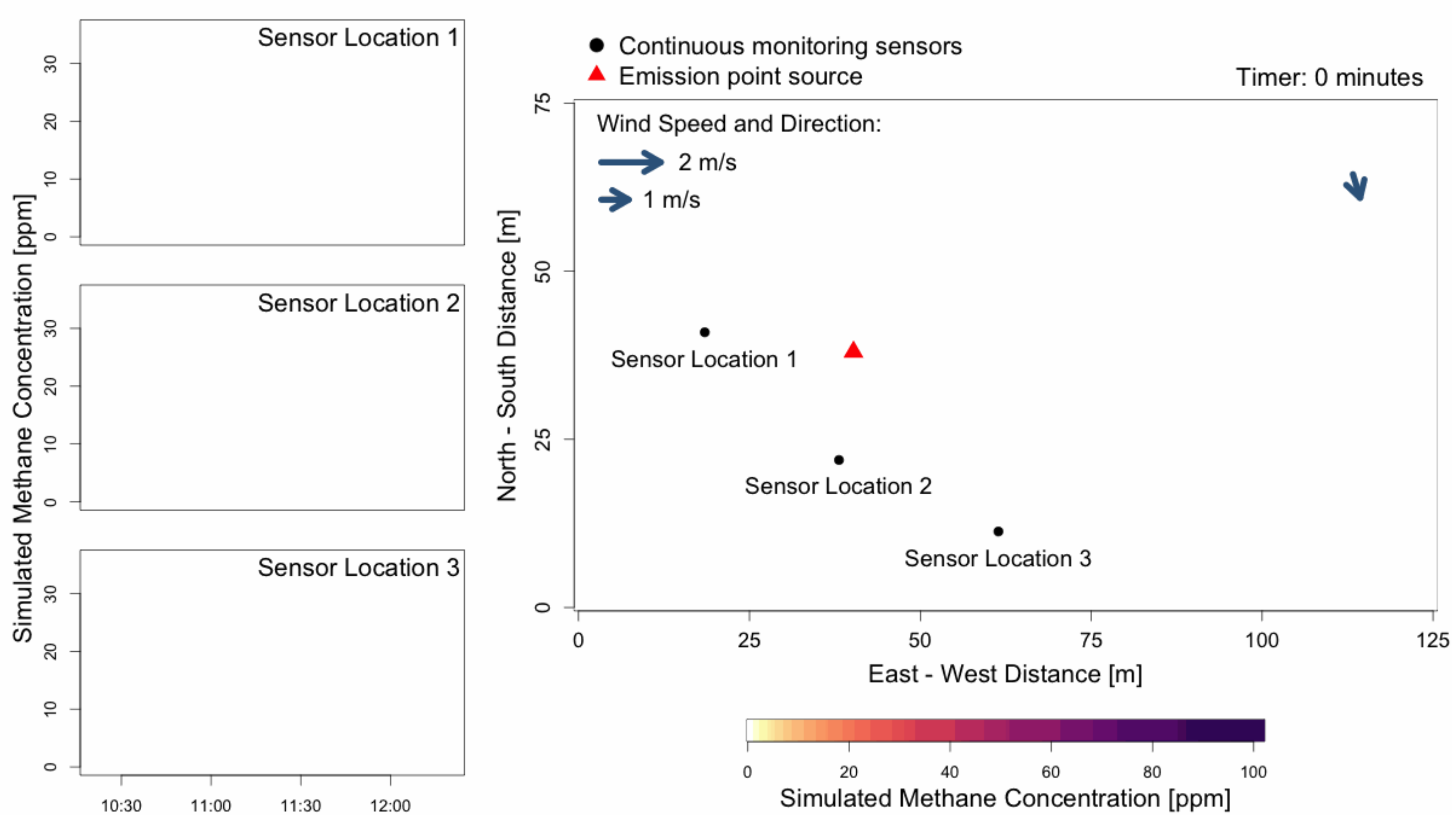




Resolution in Easting and Northing directions = 1 m

Resolution in vertical direction = 0.5 m

Step 2.2 Gaussian puff simulation

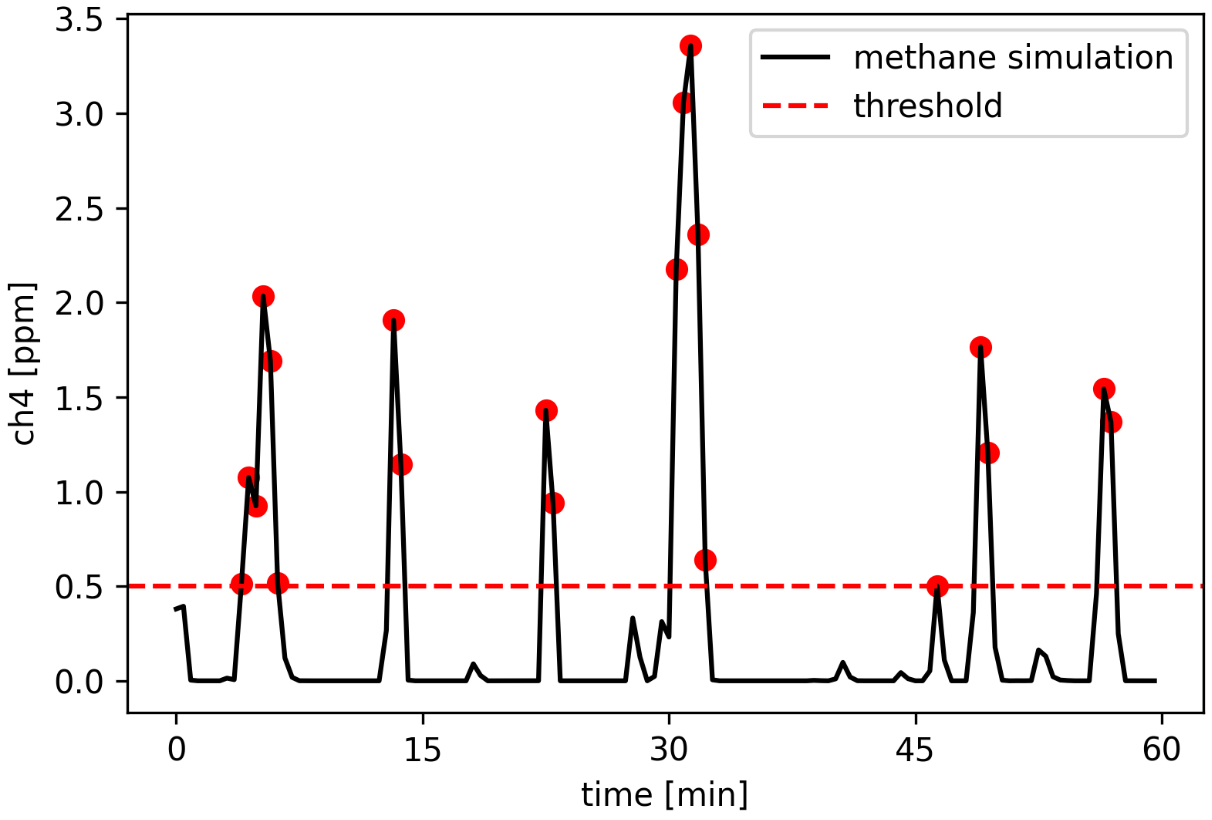


Algorithm

- Step 3: check detection status
  - Use thresholds to determine if a sensor detects an emission scenario.
  - Create a detection matrix.

Step 3.1 Thresholds

- Detection = {elevated concentration above A [ppm] for more than B% of time steps within the period}
- A is determined by sensor sensitivity
- B is set based on the tolerance for false positives



An example of a successful detection.

$A = 0.5$ [ppm] $B = 20\%$

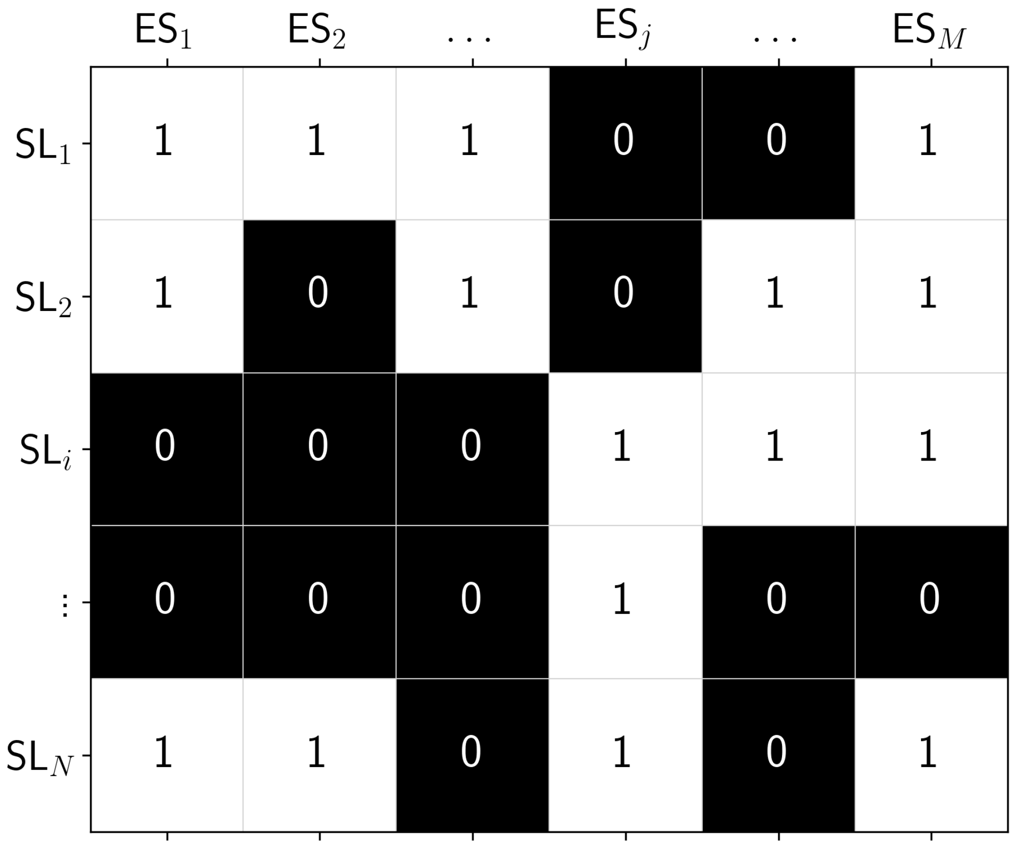
20 out of 60 (33%) points ≥ 0.5 [ppm].

Step 3.2 Detection Matrix

Rows of D: Sensor Locations (SL)

Cols of D: Emission Scenarios (ES)

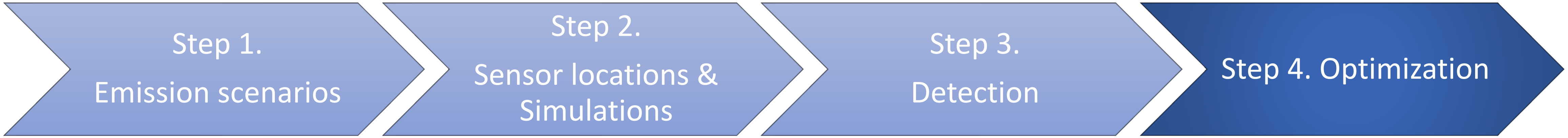
$D_{ij} = 0$, if $\mathrm{SL_i}$ can detect ES_j;
$D_{ij} = 1$, otherwise



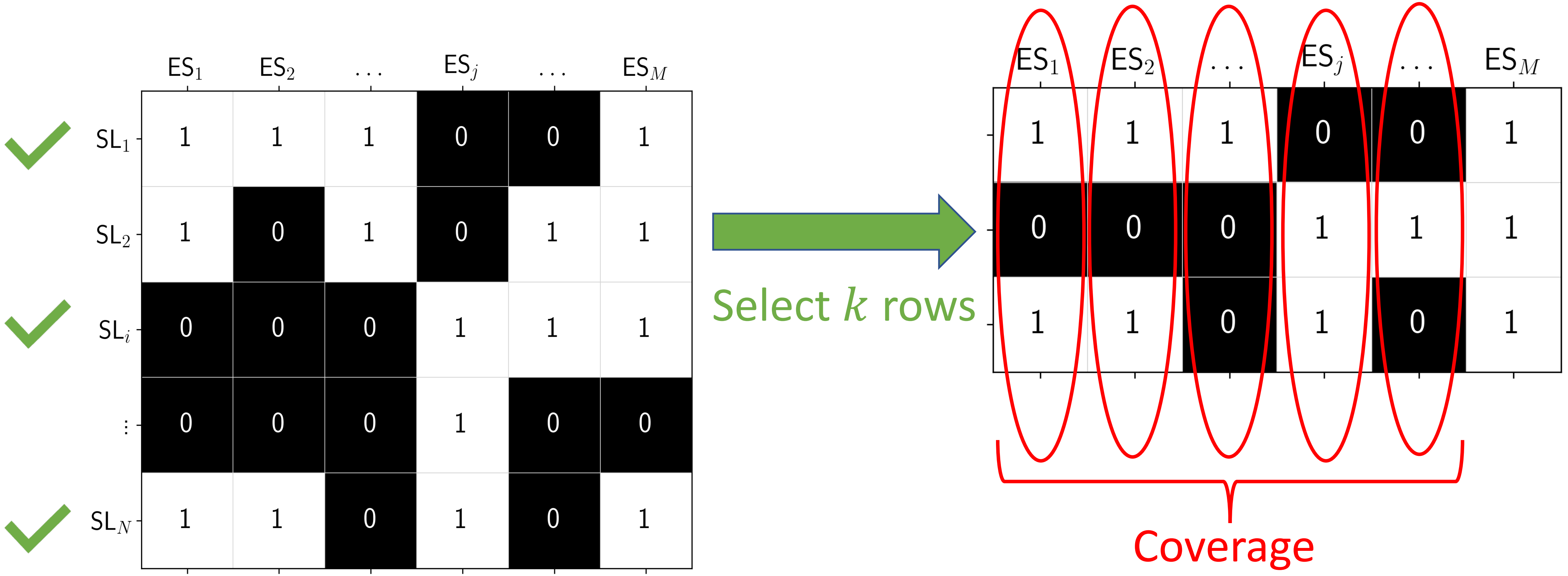
Detection Matrix D

Algorithm

- Step 4: solve an optimization problem to find the best sensor combination
  - Formulate the problem as a best subset selection task.
  - Employ the Pareto optimization framework with evolutionary algorithms (EA) to find the optimal solution.



Step 4.1 Best Subset Selection



Step 4.2 Pareto Optimization & EA

Pareto Optimization

Objectives:

Find a subset of rows (a solution) from the detection matrix to

- maximize emission scenario coverage.
- minimize the size of the subset.

Exhaustive search and standard linear programming algorithms are impossible for large-scale problem!

Evolutionary Algorithms

Process:

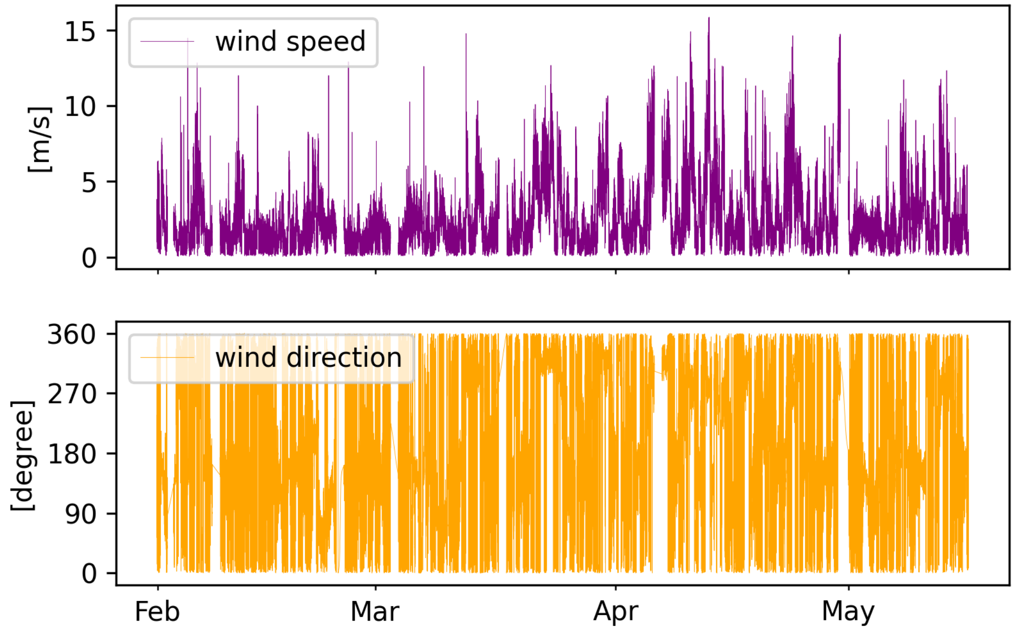
1. Randomly initialize a population of solutions.
2. Propose new solutions by perturbing existing solutions.
3. Update the population by eliminating worse solutions.
4. Repeat Step 2 & 3 until converge.
5. Return the best k-size solution.

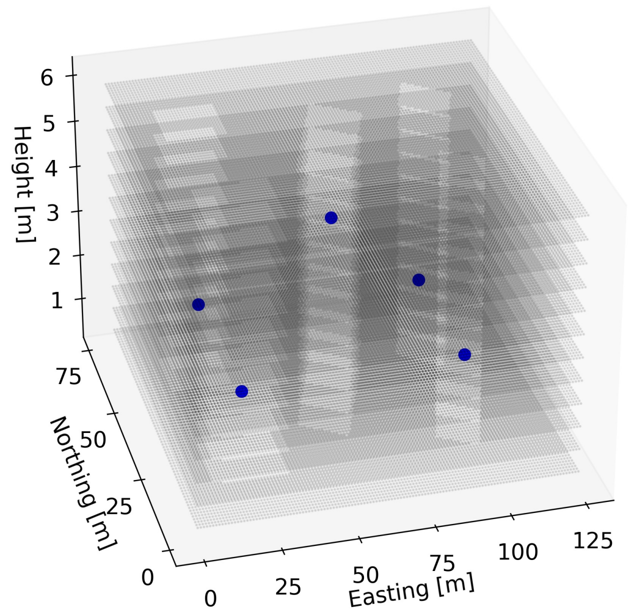
Outline

- Problem Setup
- Algorithm
- Experiments & Results
- Conclusions & Future Work

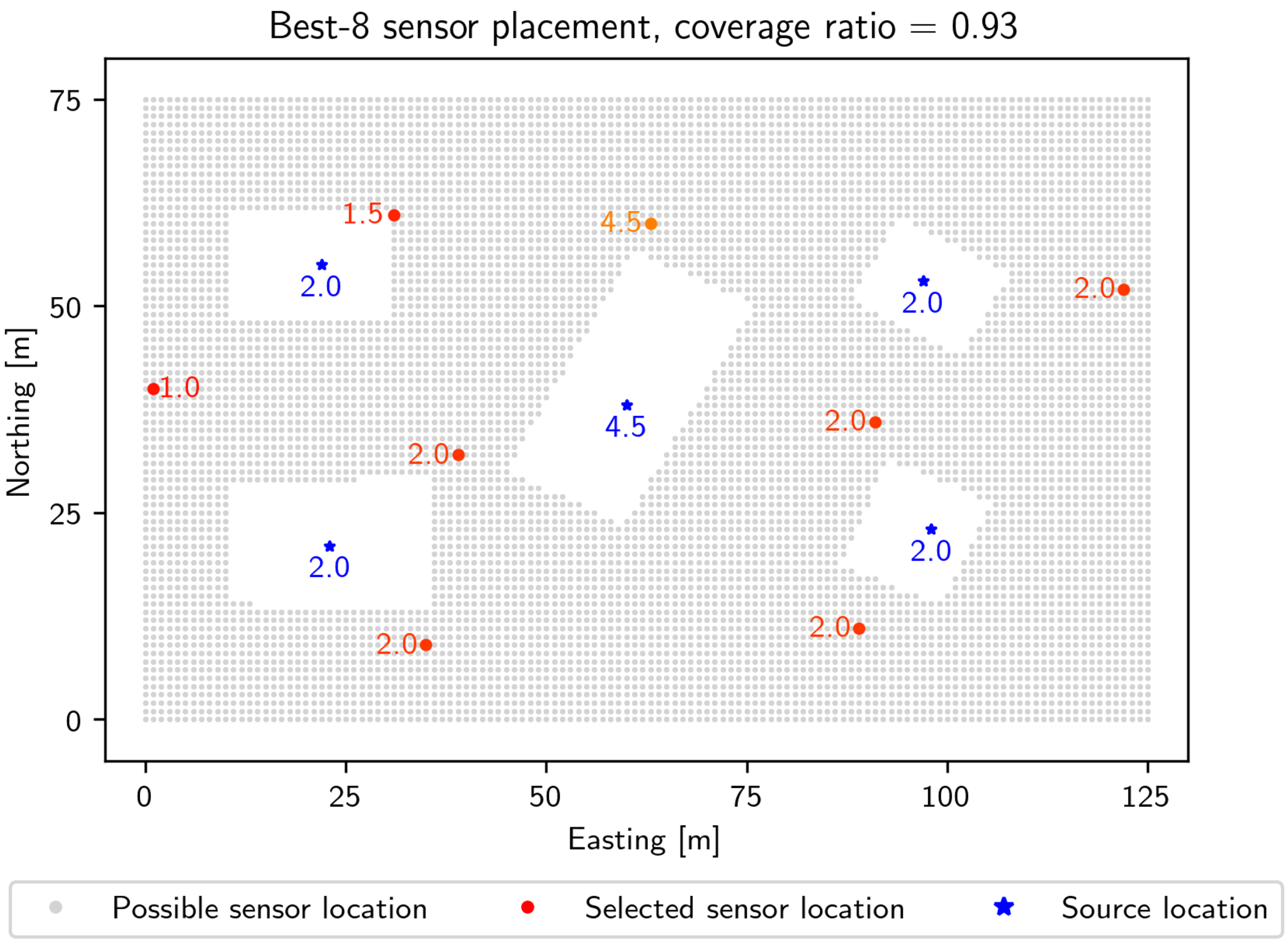
Experiments & Results

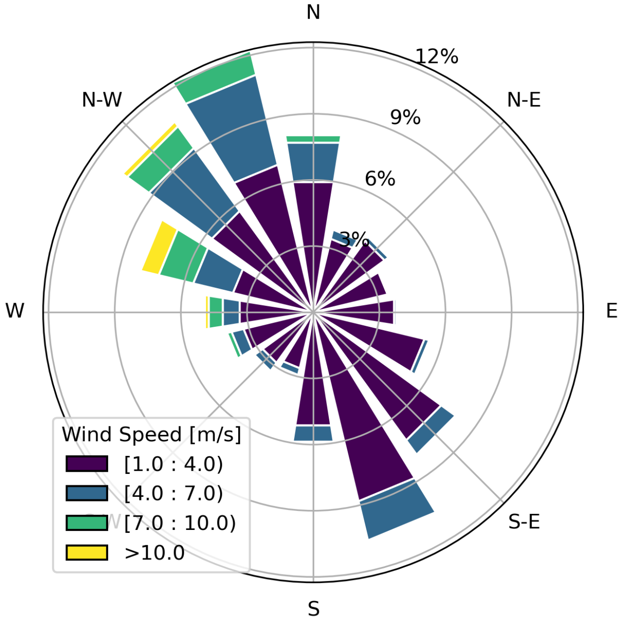
- Emission Scenarios
  - 4-month wind data
  - 5 potential source locations
  - 3 possible emission rates: {10, 20, 50} [kg/h]
  - ⟹ 38,130 emission scenarios
- Sensor locations
  - 1 [m] resolution in easting & northing
  - 0.5 [m] resolution in vertical
  - ⟹ 96,840 sensor locations



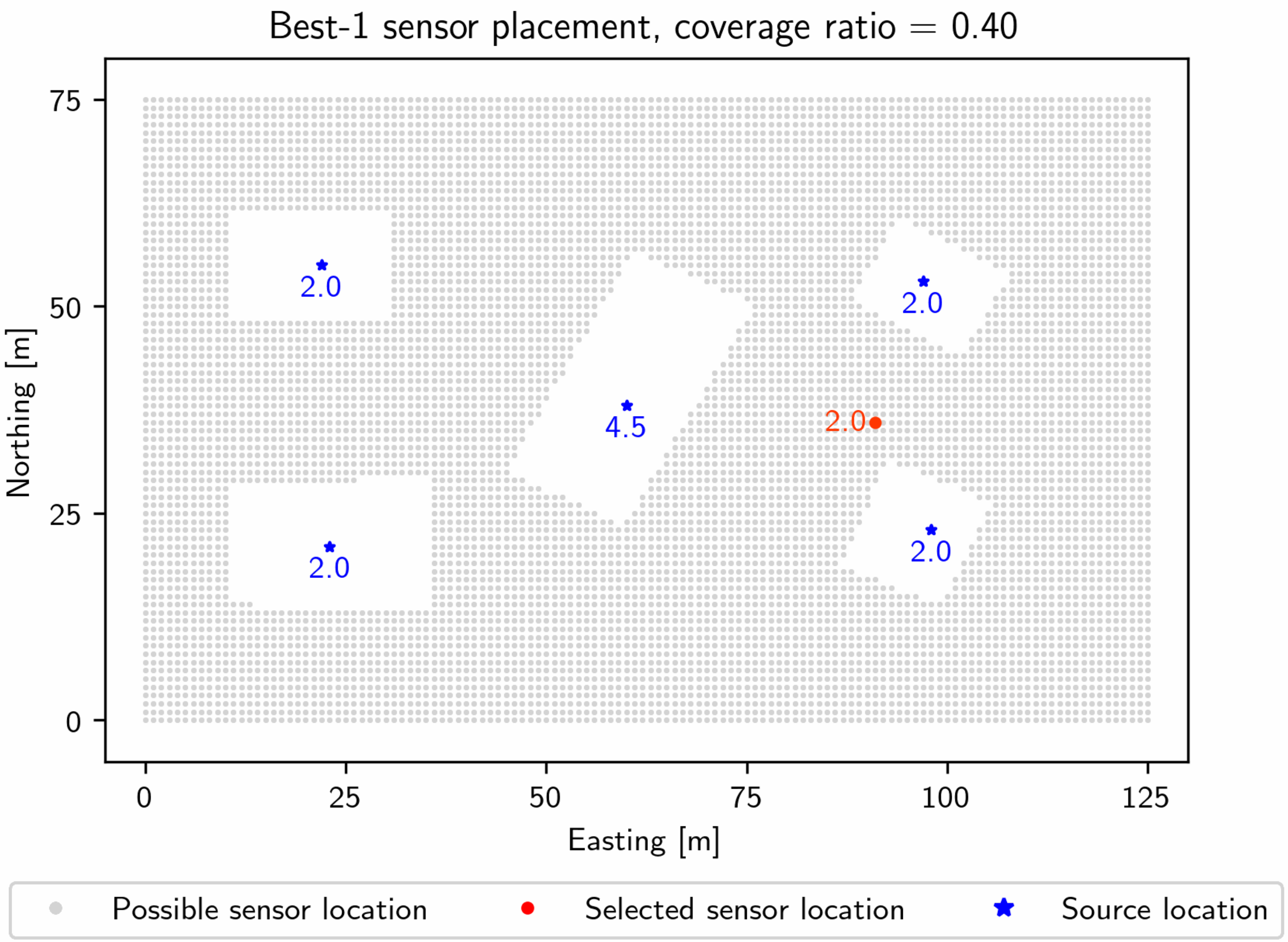


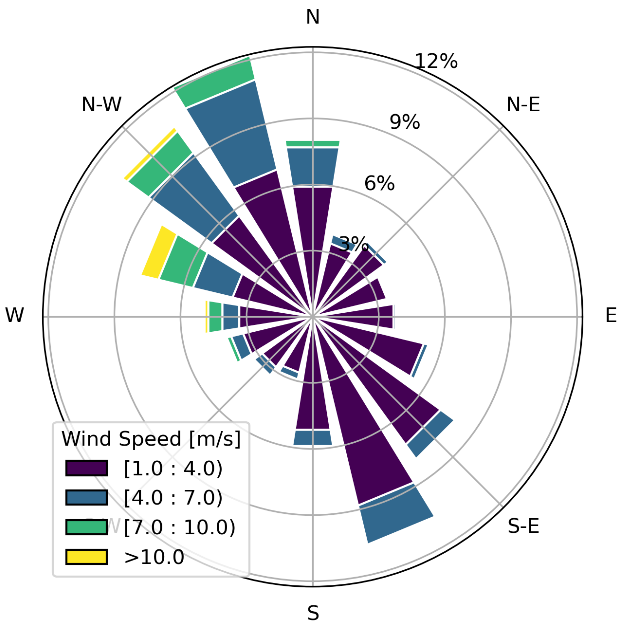
Experiments & Results – best-8 sensor placement



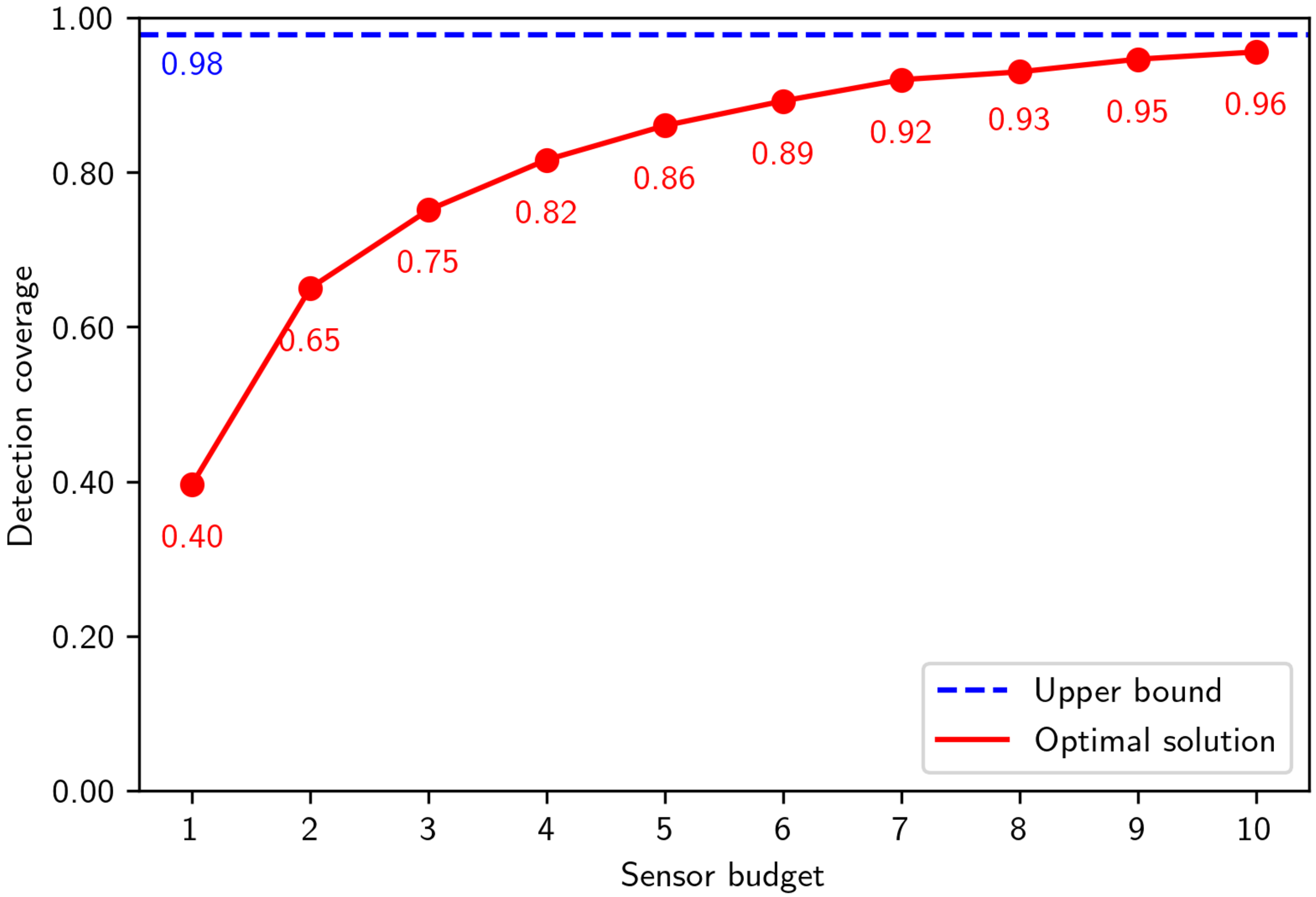


Experiments & Results – best-k sensor placement





Experiments & Results – budget vs. coverage



Outline

- Problem Setup
- Algorithm
- Experiments & Results
- Conclusions & Future Work

Conclusions & Future Work

- Developed a data-driven algorithm for sensor placement more accurate and efficient than traditional methods.
- The algorithm's modularity ensures adaptability to various monitoring needs.
- Optimized for solving large-scale problems efficiently.
- To implement a generative model for better approximation of wind distributions, thereby expanding the emission scenario database.
- To investigate advanced data embedding techniques to manage and solve problems of greater scale.

References

- Klise, Katherine A., et al. "Sensor placement optimization software applied to site-scale methane-emissions monitoring." *Journal of Environmental Engineering* 146.7 (2020): 04020054.

- Qian, Chao, Chao Bian, and Chao Feng. "Subset selection by pareto optimization with recombination." *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 34. No. 03. 2020.

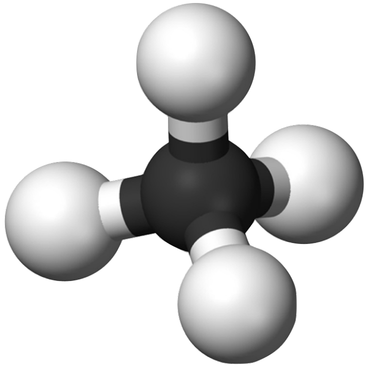
Questions?

Thank you for attending!

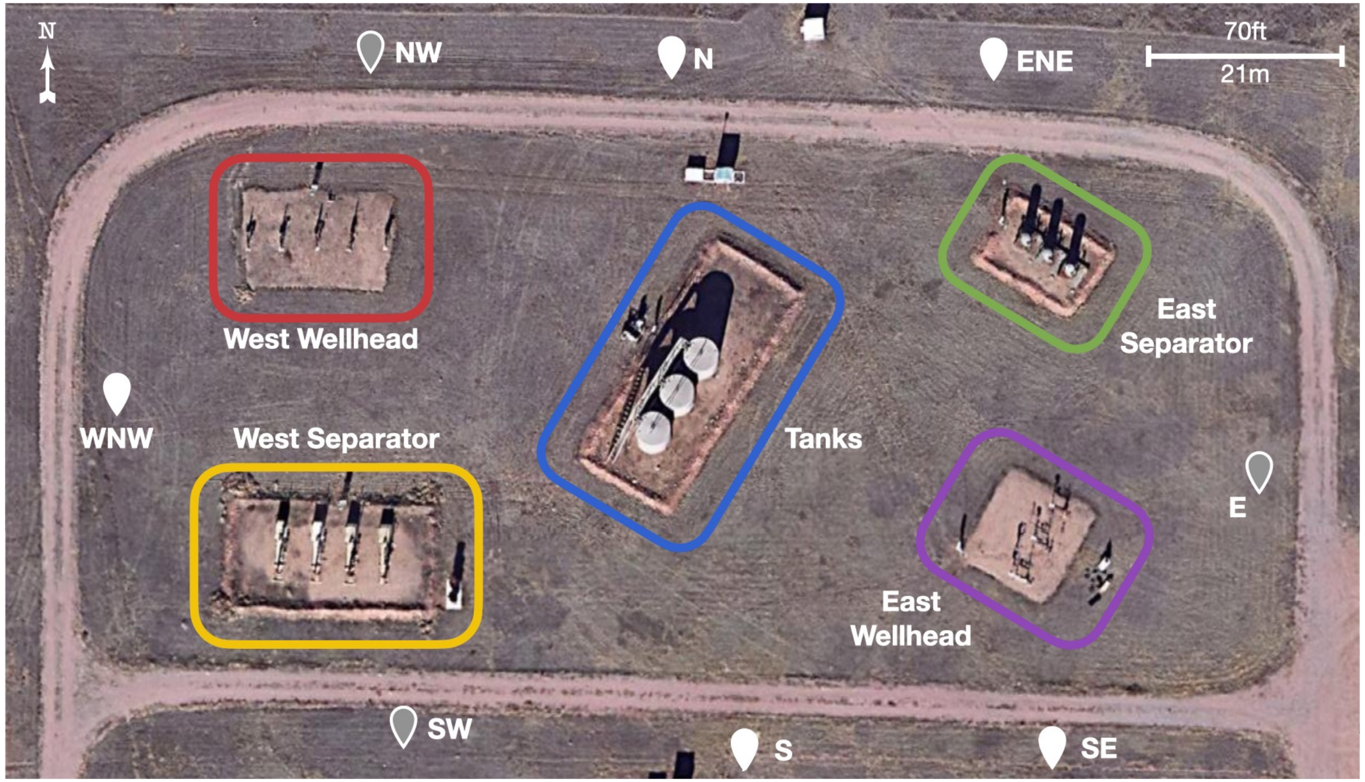
Back up

Background

- Methane, CH_4, is the second biggest cause of climate change after CO_2.
- Oil & gas sector gives off 15-20% of total methane.
- Characters of emission from oil & gas
  - high temporal variability
  - infrequent, short-lived super emitters
- Continuous monitoring system (CMS) is necessary



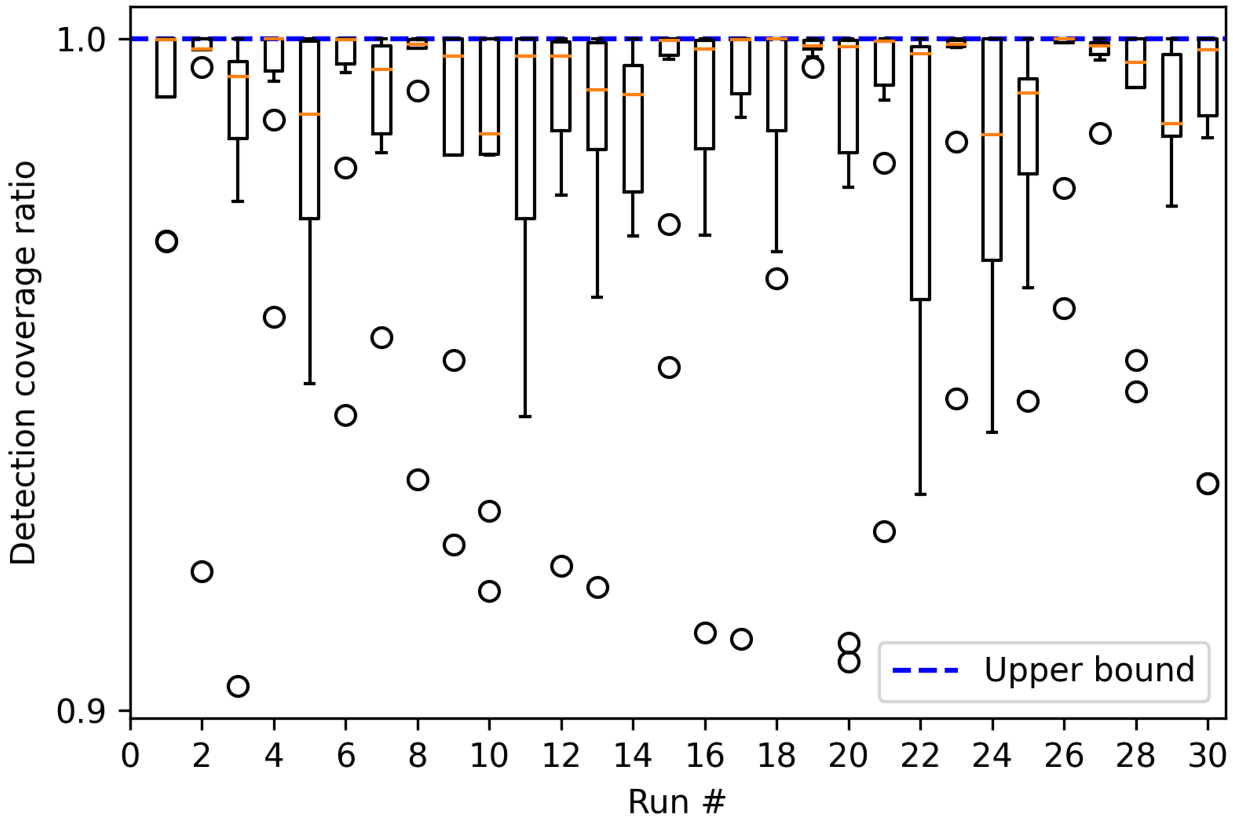
Fence Line Placement on METEC



Continuous monitoring (CM) sensor

Test EA on synthetic large matrix

- nrows = ncols = 100,000
- k = 10, randomly placement in the big matrix
- Test on 30 cases and run 10 EA algorithm for each case



Optimality Guarantee

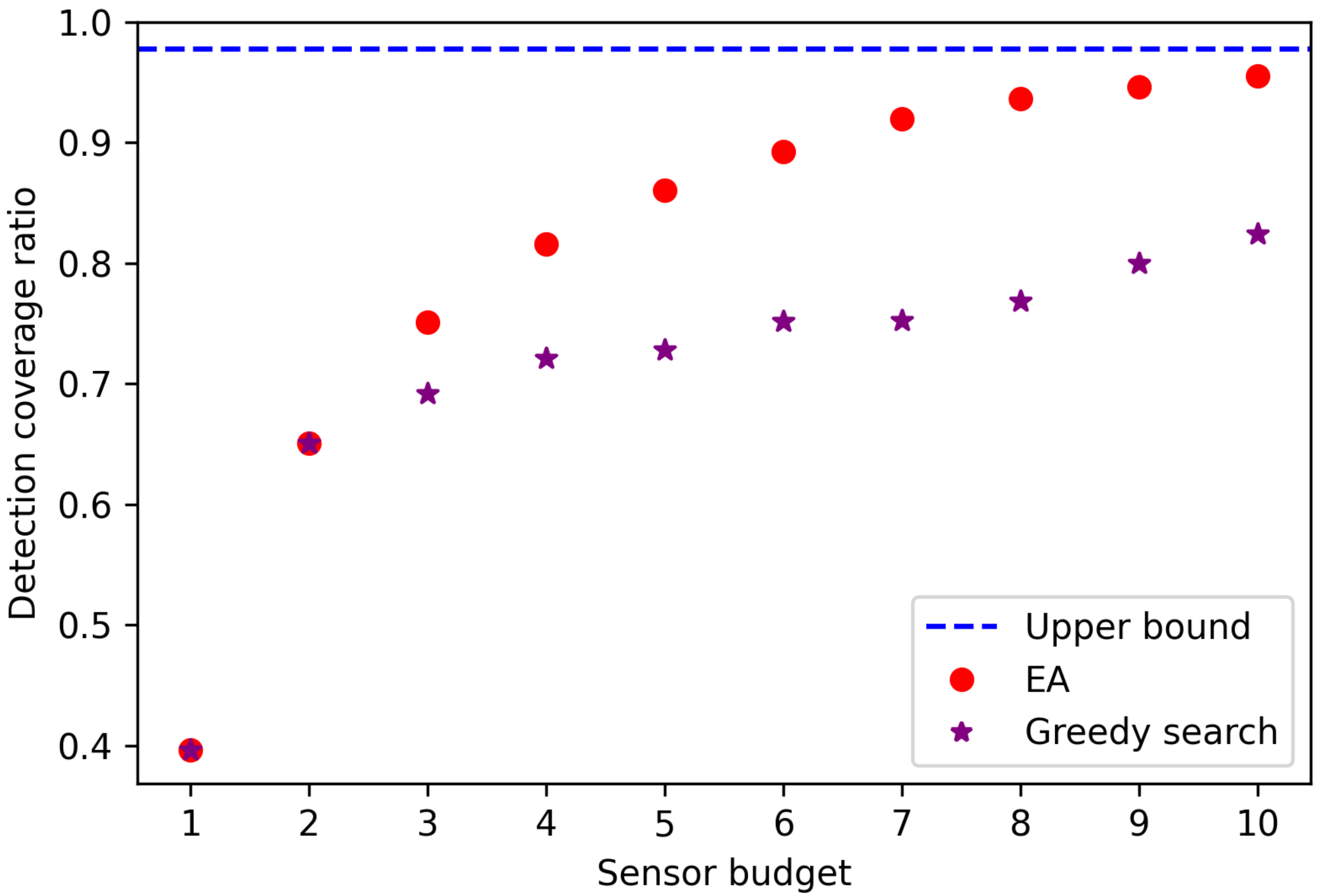
- In theory, we prove that for subset selection with monotone objective functions, PORSS can achieve the optimal polynomial-time approximation guarantee, $1 - e^{\gamma}$ where γ is the submodularity ratio measuring how close your objective function is to submodularity.

Related Work

	Klise et al. (2020)	Our work
# emission scenarios	1,200	$\approx$ 40,000
# possible sensor locations	$\approx$ 2,500	$\approx$ 100,000
Forward model	Gaussian plume	Gaussian puff
Optimization algorithm	Mixed-integer linear programming	Pareto optimization using evolutionary algorithm (EA)

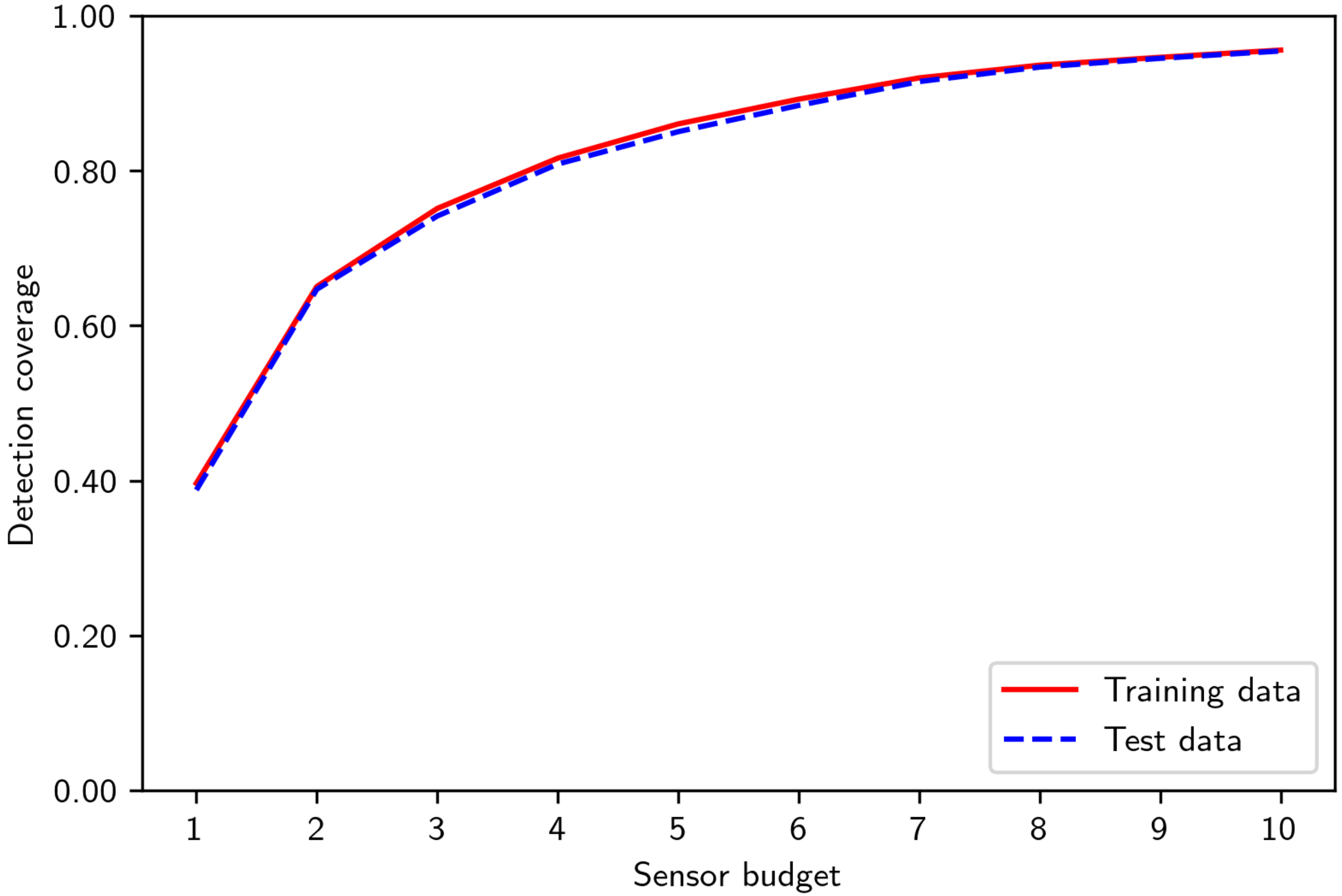
EA vs. Greedy Search

- EA vs. greedy search

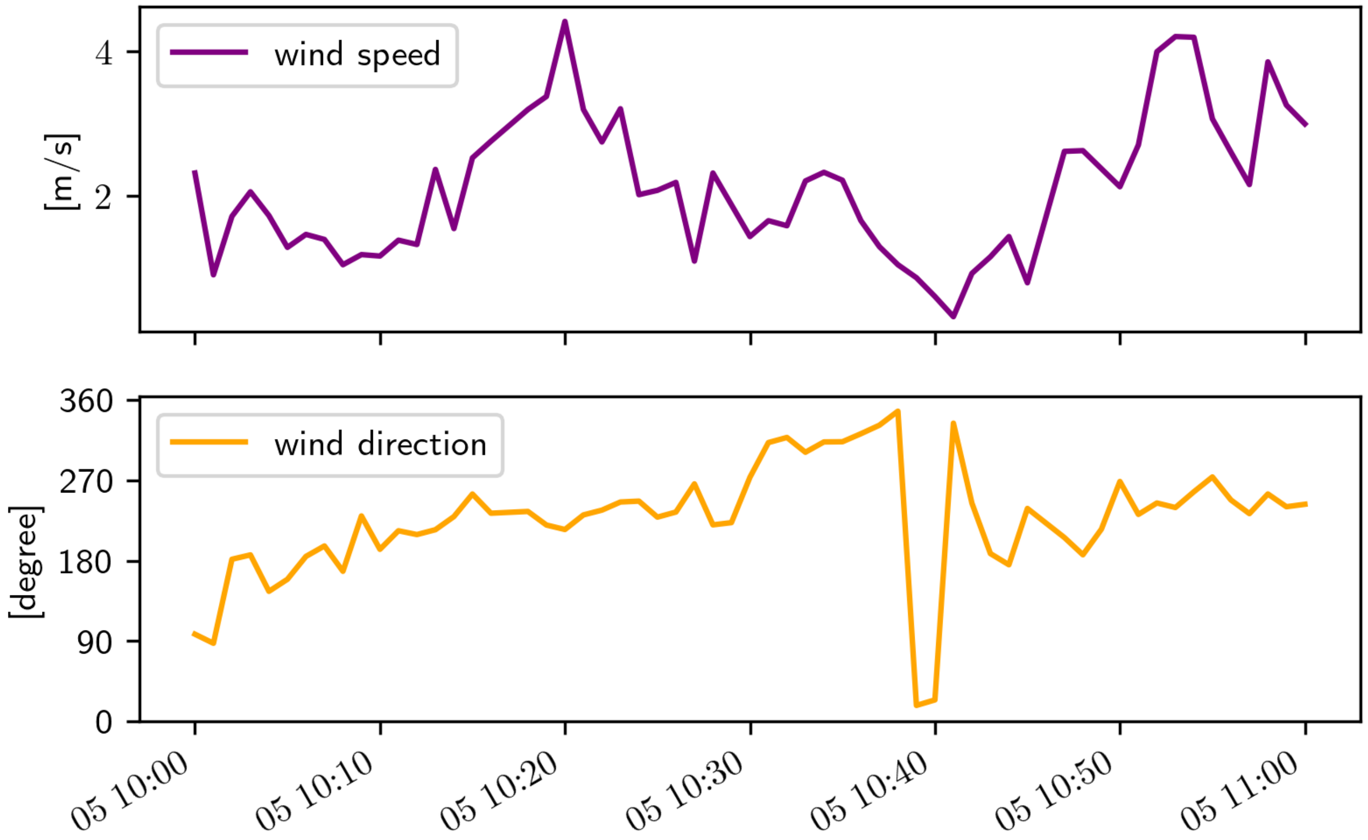


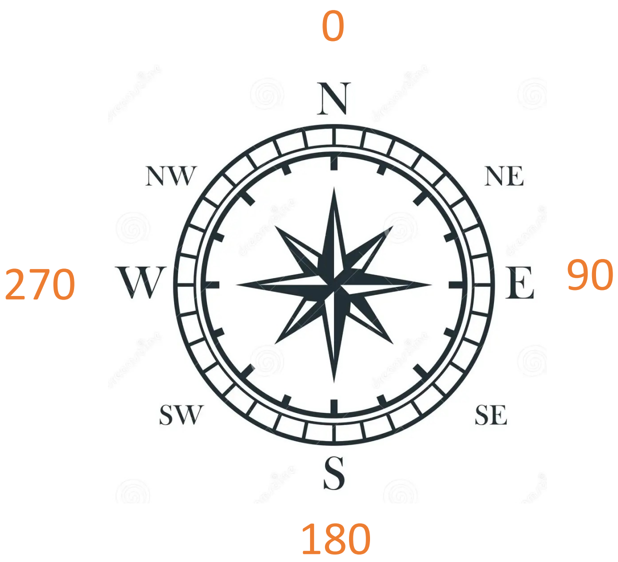
Experiments & Results - robustness

Use a different set of 10,000 emission scenarios to validate the performance of the optimal sensor placement.

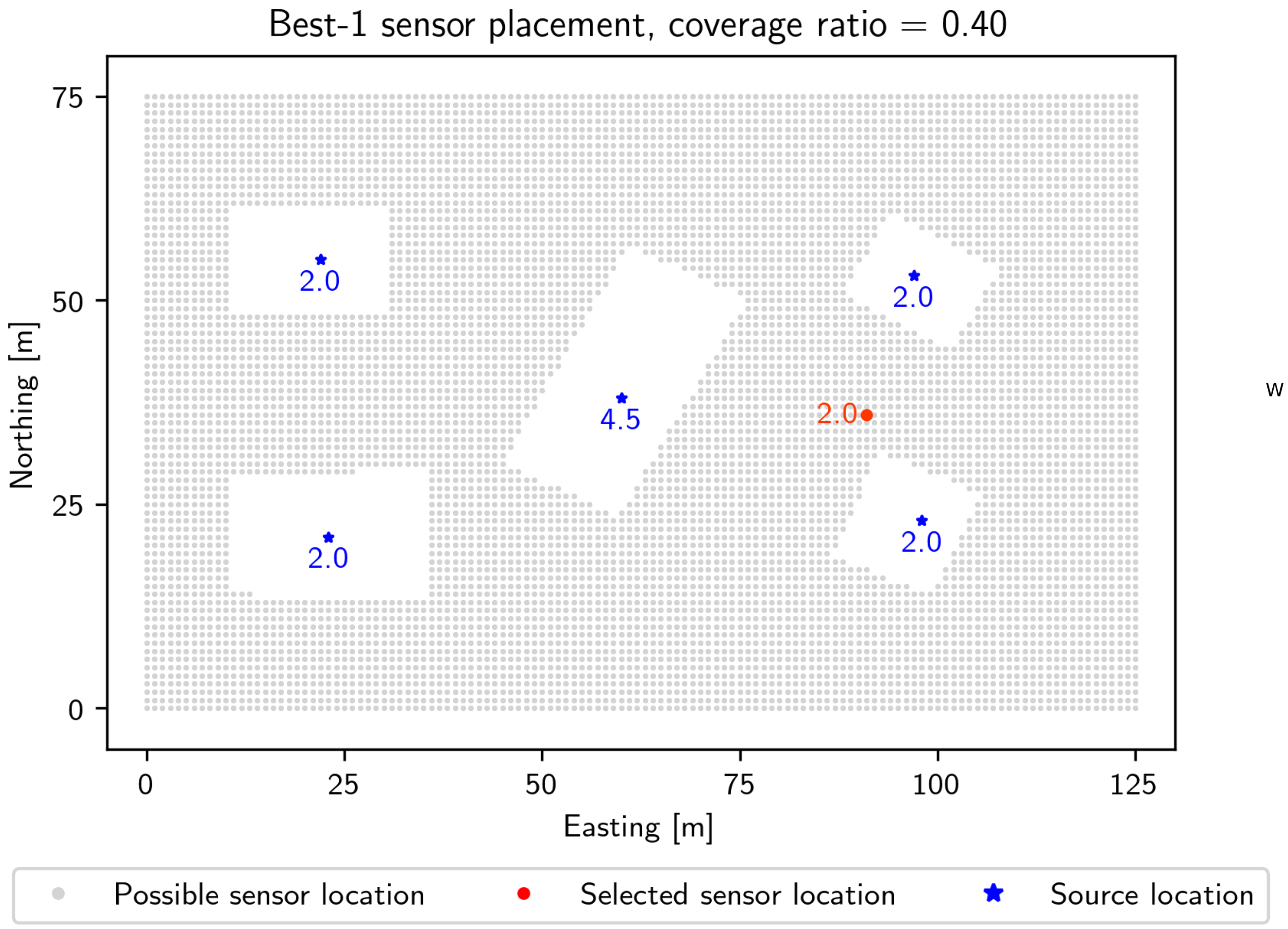


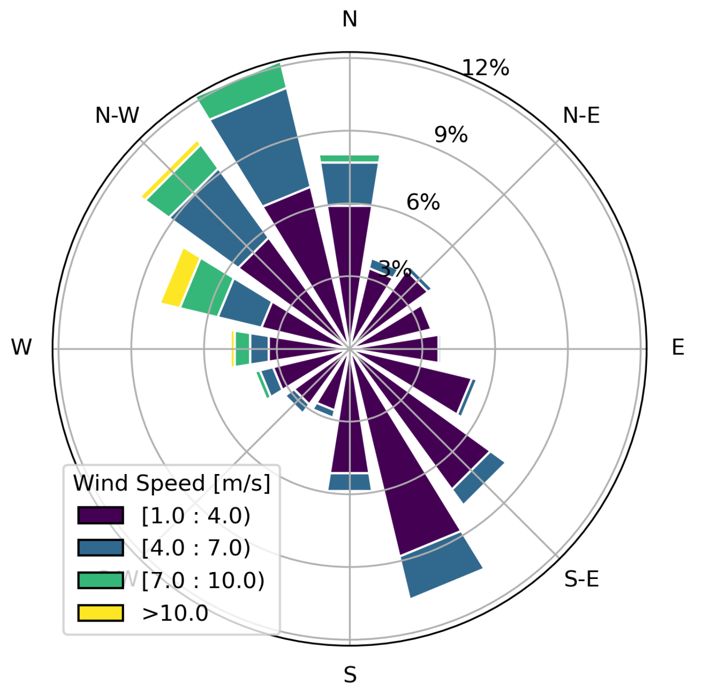
Why some scenarios are always undetected?



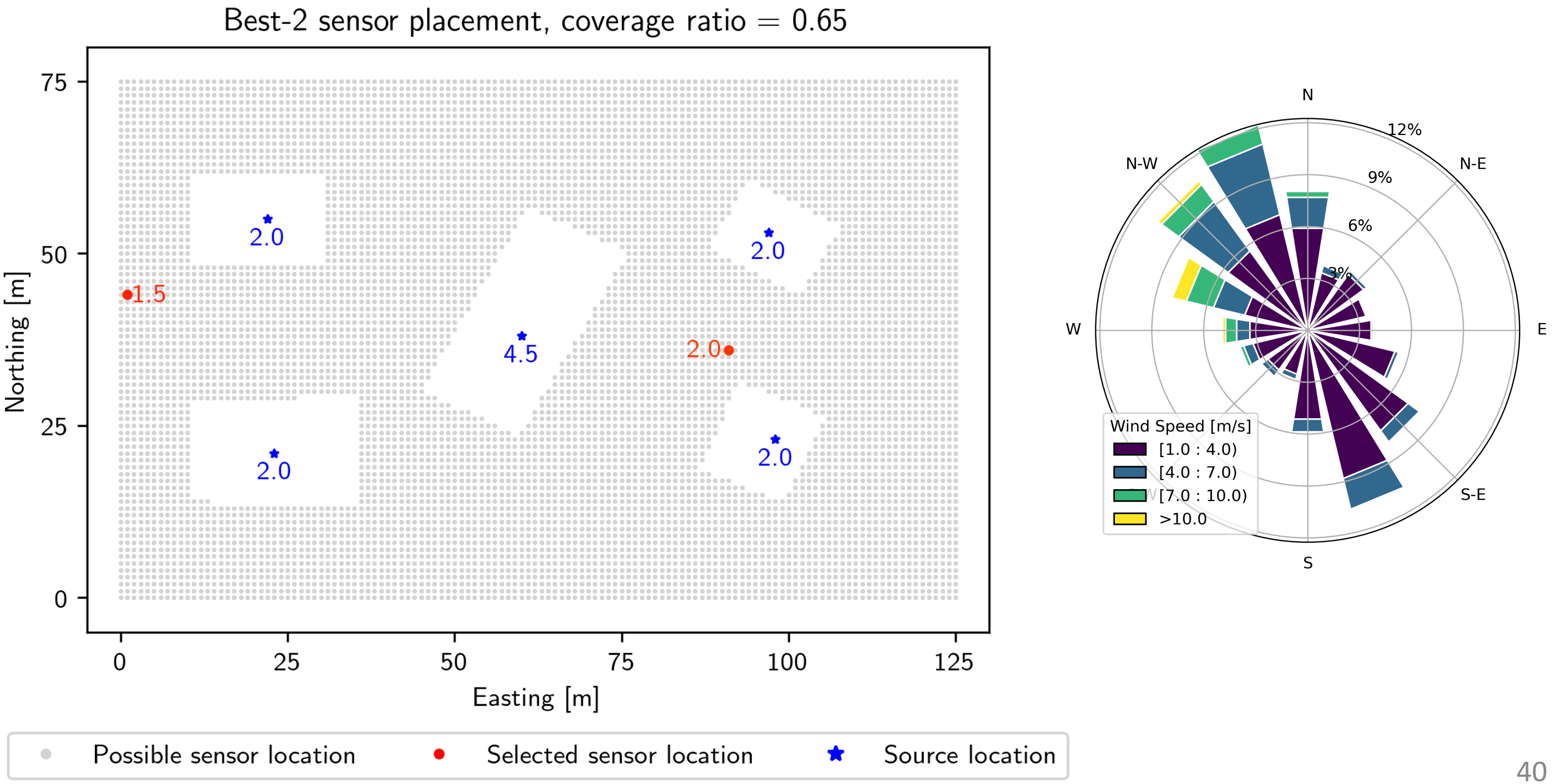


Best-1 Sensor Placement

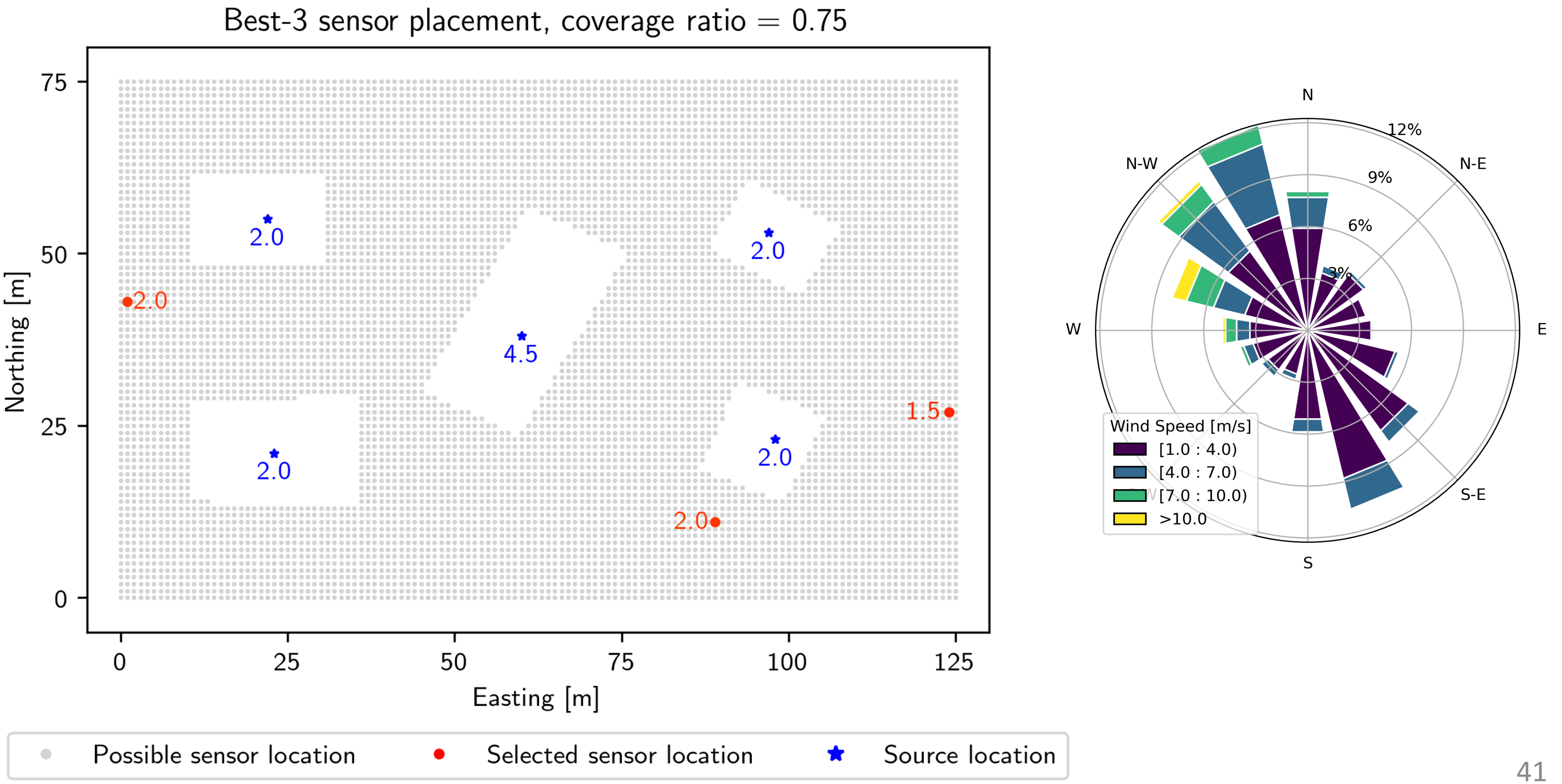




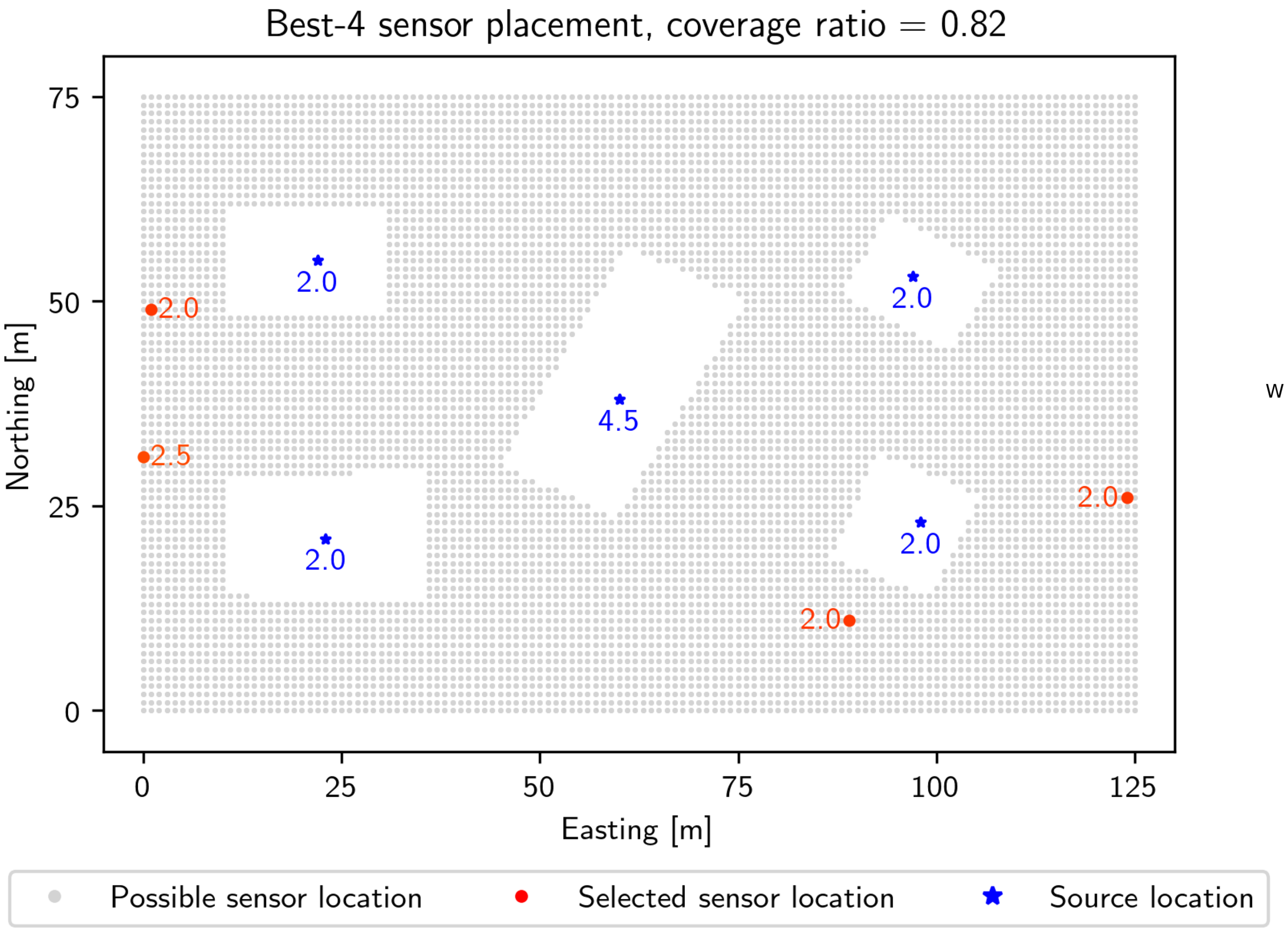
Best-2 Sensor Placement

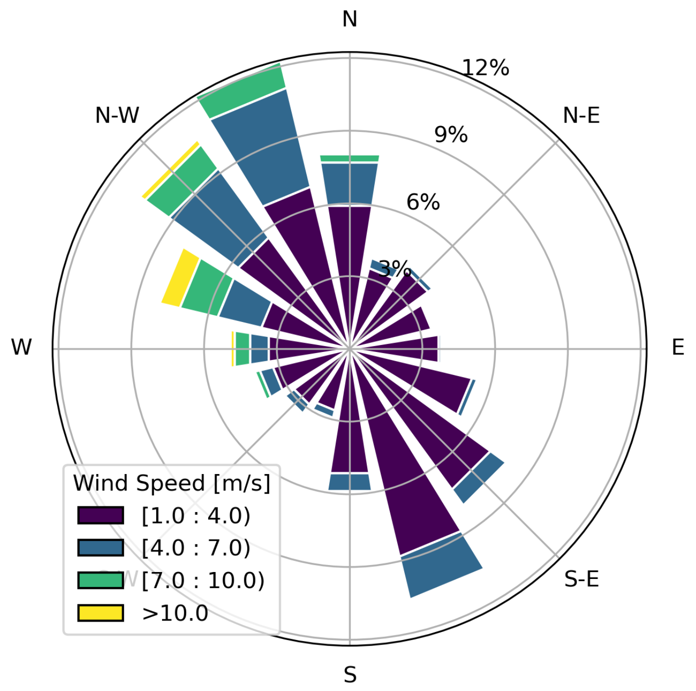


Best-3 Sensor Placement

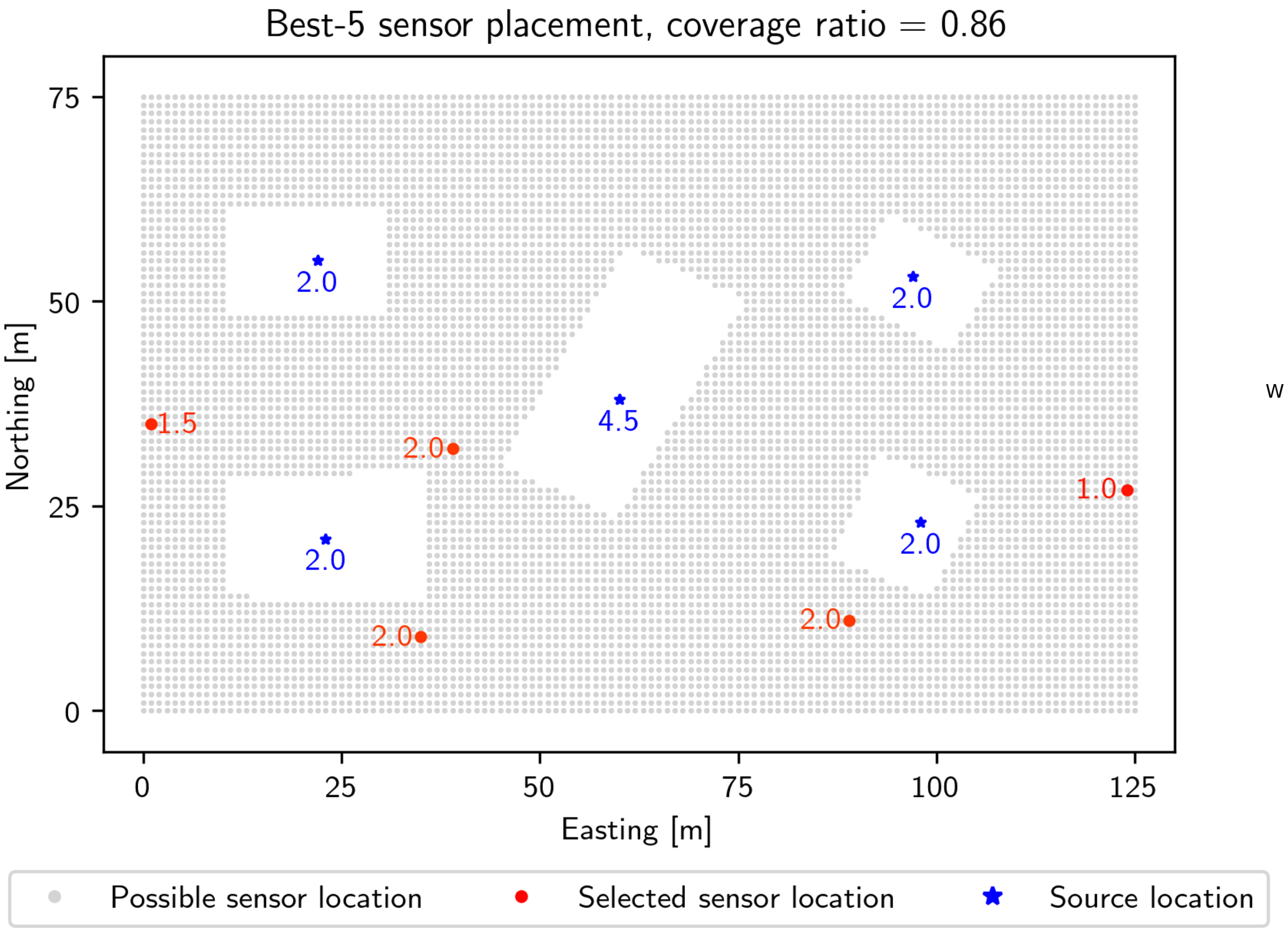


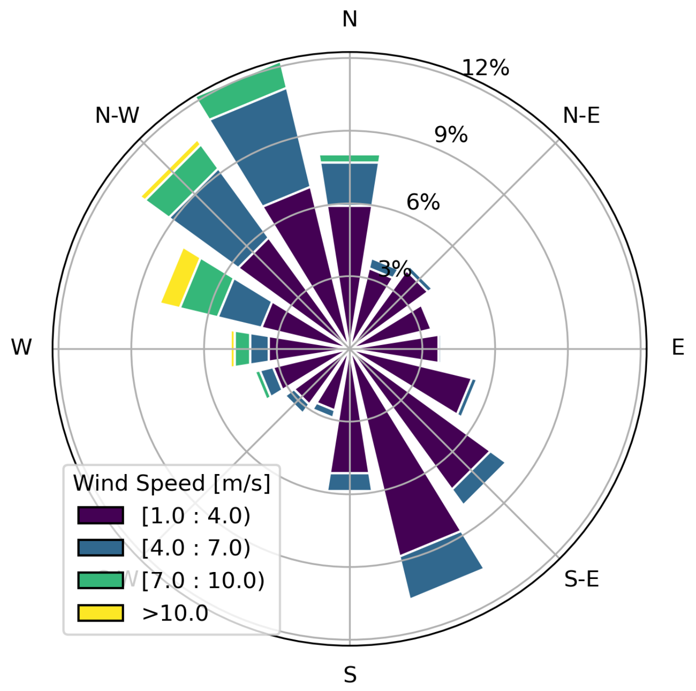
Best-4 Sensor Placement



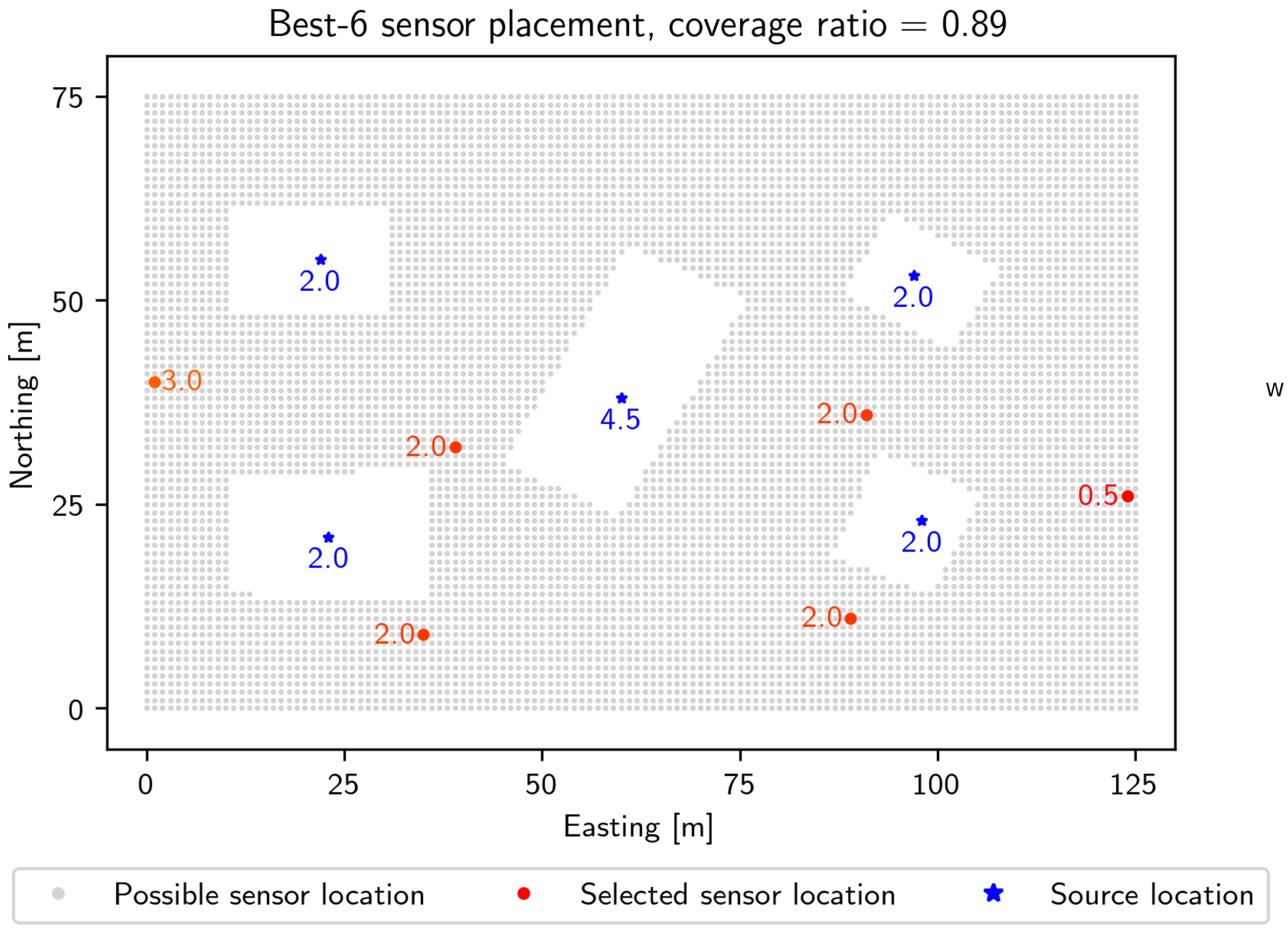


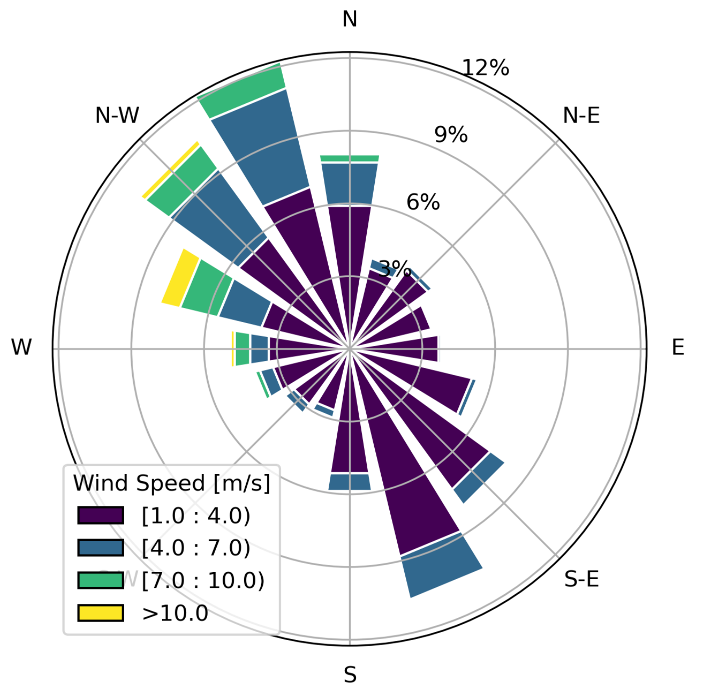
Best-5 Sensor Placement



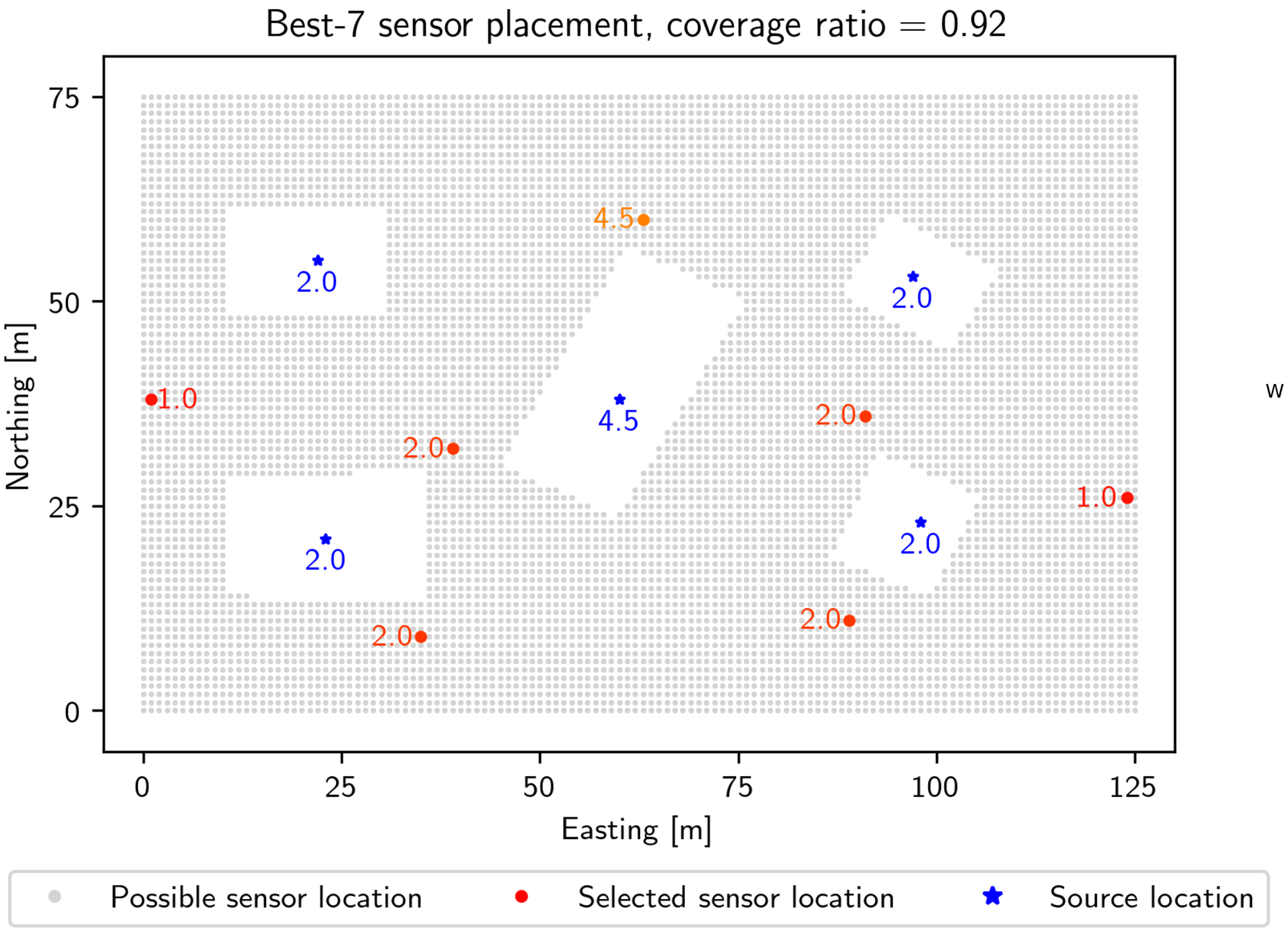


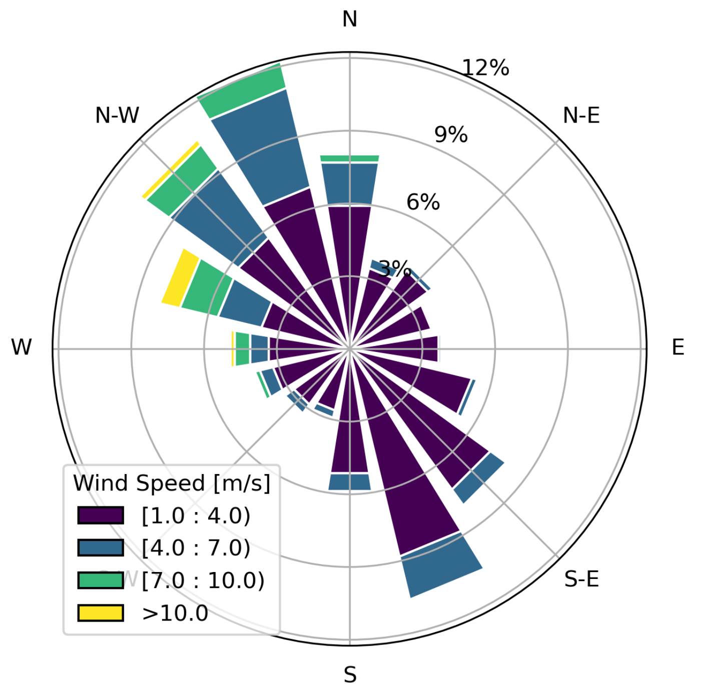
Best-6 Sensor Placement



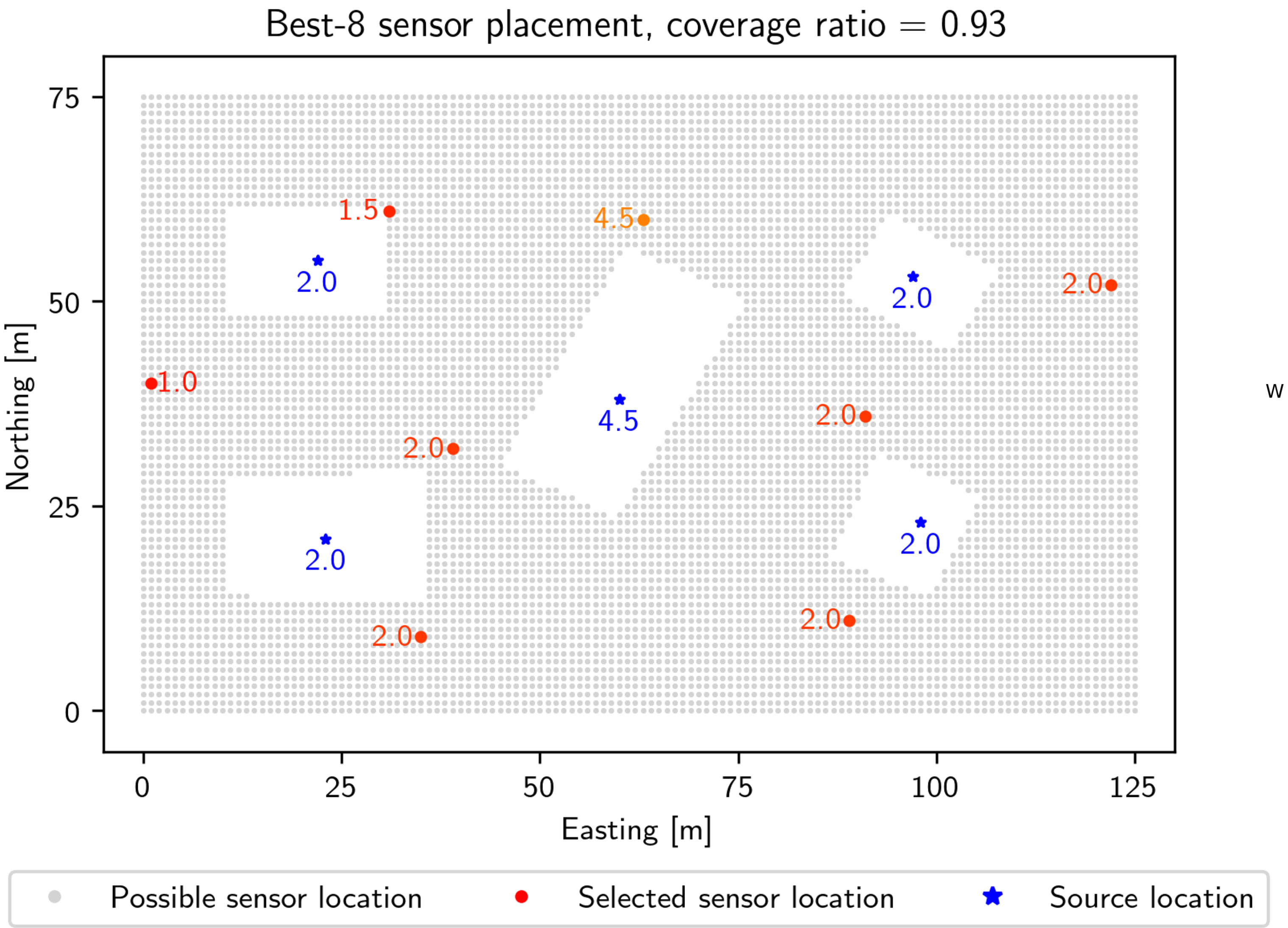


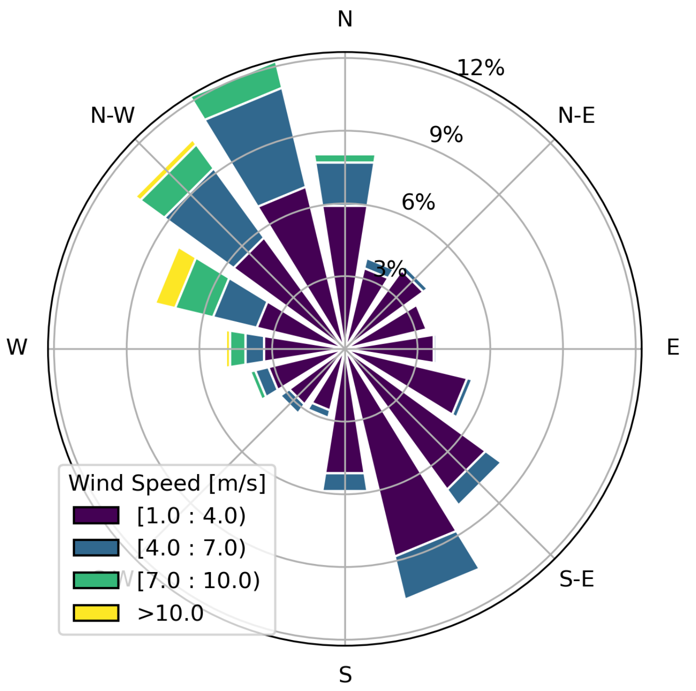
Best-7 Sensor Placement



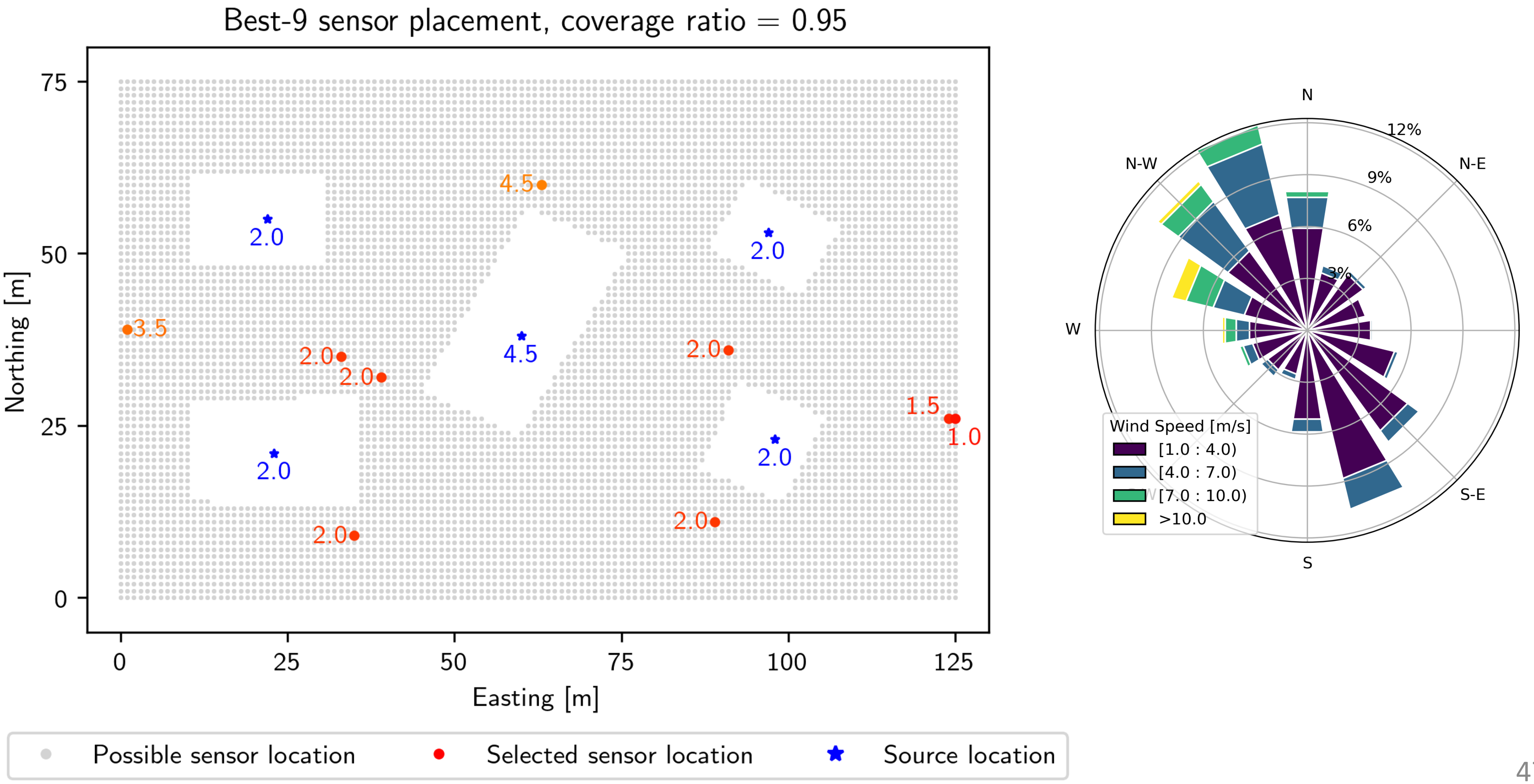


Best-8 Sensor Placement





Best-9 Sensor Placement



Best-10 Sensor Placement