Optimal selection of monitoring sites in cities for SARS-CoV-2 surveillance in sewage networks

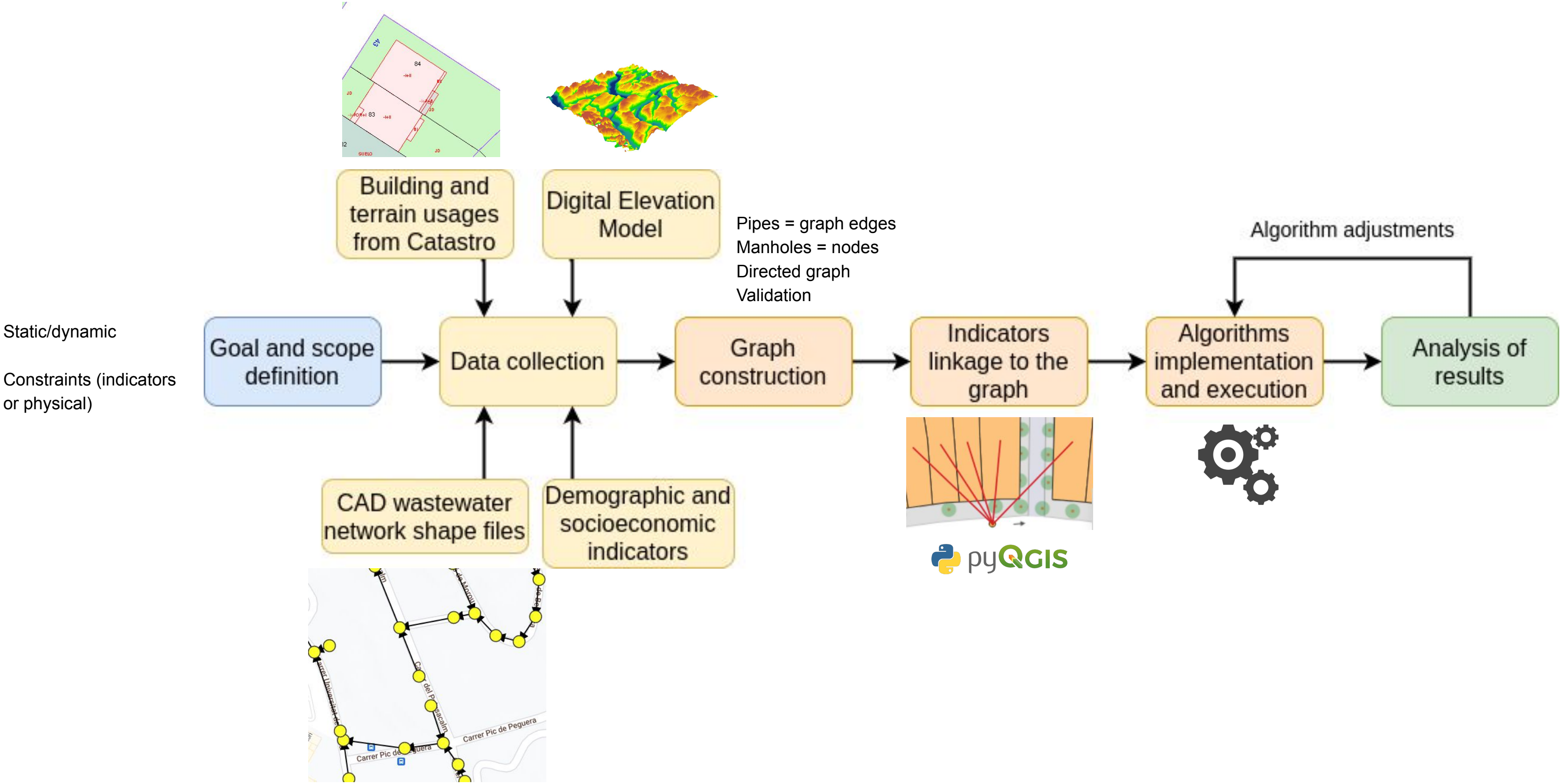
Speaker: Miquel Farreras

Authors: Eusebi Calle, David Martínez, Roser Brugués-i-Pujolràs, Miquel Farreras, Joan Saló-Grau, Josep Pueyo-Ros, Lluís Corominas

Introduction

- Sewage systems are a good, unbiased indicator of the prevalence of a virus in a population.
- Successful monitoring of SARS-CoV-2 in wastewater treatment plant (WWTP), but limited experience at **neighbourhood level**.
- ‘Upstream’ surveillance for SARS-CoV-2 may facilitate finer spatial detection of the virus.
- Sewer systems are long **complex networks**. 100,000 inhabitants ≈ 300 km of small sewer pipes and ≈ 60 km of bigger main pipes.
- It is not evident where to place autosamplers (or sensors if available in the future).
- Selecting sampling locations can follow several criterion:
    - Number of sampling locations to be installed for maximum coverage -> (**static monitoring site selection**)
    - Detect the area where the virus is more present -> (**dynamic monitoring site selection**)
- Based on **graph-theory** and coupled to a **greedy optimization algorithm.**

Methodology



Static monitoring site selection algorithm

Goal: Select the locations for permanent monitoring sites, for maximum coverage at minimal investment.

Interference (only for static monitoring site selection): two or more monitoring sites overlap in covered manholes.

Algorithm **Monitoring Site Selection (MSS)**: greedy approach. **Starts from a random** combination of nodes with placed monitoring sites S, and then iterates to find better combinations by moving each monitoring site through its neighbouring nodes, creating subgraphs. The MSS algorithm halts when it is not possible to find a monitoring site set by which does not improve as compared to the S set.

Evaluation function **Monitoring Site Selection Evaluation (MSE)**: allows an optimization metric after each execution of the MSS algorithm to be estimated. The maximization of the unique coverage U minimizes the interference between coverage areas.

The MSE maximizes the unique coverage U and minimizes the difference between the sizes of the resulting network coverage areas.

Dynamic monitoring site selection algorithm

Using the MSE algorithm and obtaining a subgraph of the sewage network with a monitoring site s as output, the dynamic monitoring approach is developed.

Larson et al.’s approach uses Bayesian probabilities of infection to all possible source nodes and applying a “binary search” using these probabilities. Our implementation assigns the Bayesian probabilities according to inhabitants connected to each node. Higher population, greater chances of finding infected people. **Two goals**:

- **Patient Zero** (PZ) algorithm assumes there exists only one case of COVID-19 in a city and tries to find the minimum sequence of manholes to test in order to locate that first source of infection.
- **Hot Spot** (HS) algorithm assumes that many individuals are already infected and seeks to find the cluster in the sewage network with the largest SARS-CoV-2 RNA load.

If the viral load is high compared to the previously tested manhole, the infected area is upstream from this point and we can discard all the network nodes downstream. Otherwise, the upstream nodes are discarded.

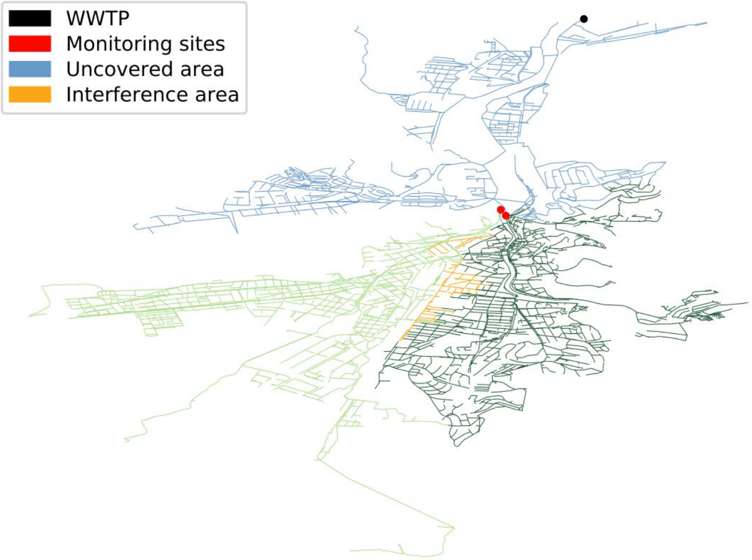
The **PZ algorithm is a special instance of the HS algorithm**, where the stopping rule is reached when there is only one source node left.

Case study: Girona

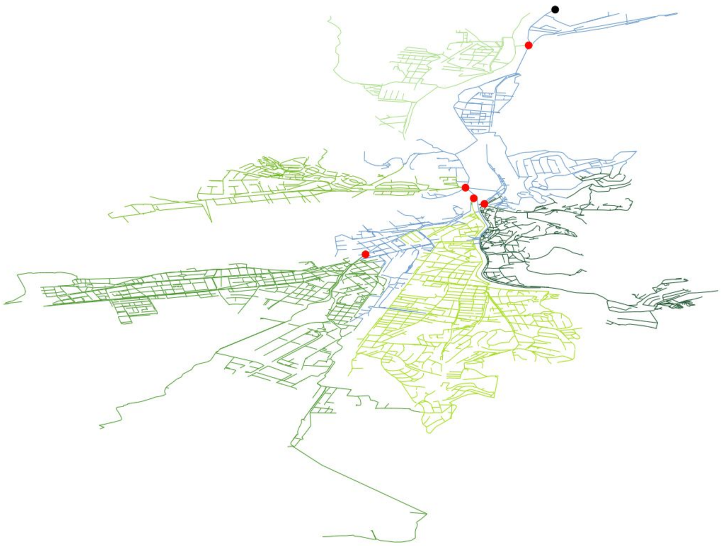
The city of Girona, in northern Catalonia, Spain, is a city with the following characteristics:

- 101,852 inhabitants, with a metropolitan area shaped by seven municipalities with a total of 148,504 inhabitants.
- **9,718 manholes** (nodes)
- 338 km of pipes
- **10,185 edges**
- Average node degree of 2.1 (D)
- Average shortest path length of 9,580 (d)
- **2.7 citizens per household**
- Topological data provided by the municipality through GIS format. Files combined to a GraphML file format, compatible with the **Network Robustness Simulator** (NRS) used for graph analysis.
- Data verification and reconciliation approach, checking for inconsistencies in disconnected nodes and/or edges.

Girona results : Static monitoring site selection

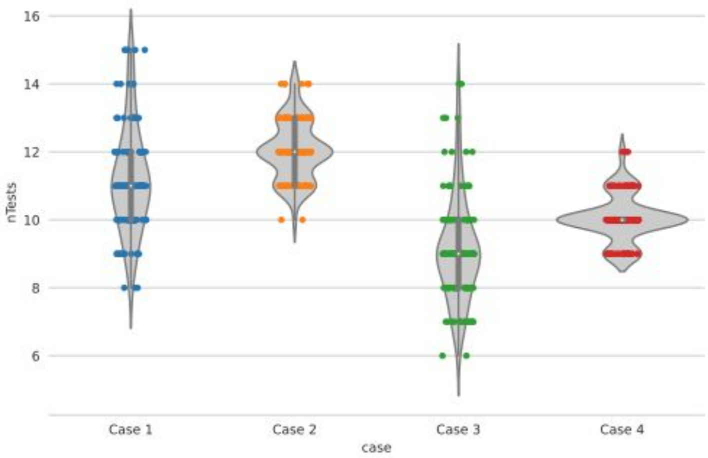


Two monitoring sites (71.51% cov.)

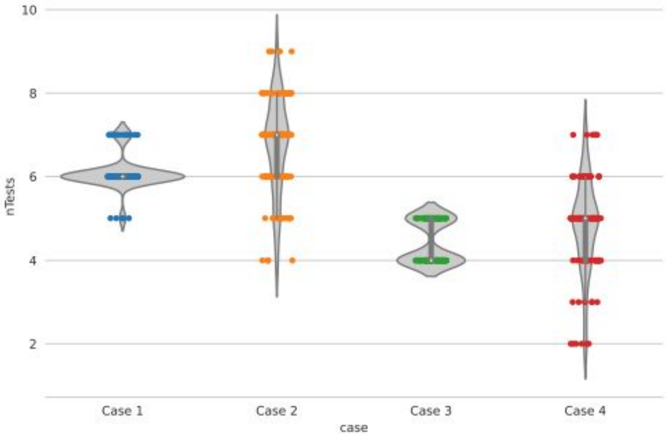


Five monitoring sites (82.97% cov.)

Girona results: Dynamic monitoring site selection



(a) Patient Zero algorithm.



(b) Hot Spot algorithm.

case 1 – entire wastewater network and Bayesian probabilities as a function of population.
case 2 – entire wastewater network and Bayesian probabilities randomly assigned.
case 3 – island from the MSS algorithm and Bayesian probabilities as case 1.
case 4 – island from the MSS algorithm and Bayesian probabilities as case 2.

Conclusions

It is possible to optimally select sampling points for sewage surveillance in cities. **Three algorithms achieved**:

- Static sensor placement
- Dynamic sensor placement:
    - Identify patient zero
    - Hot spots in cities

The best option for detecting a patient zero and a hotspot area implies assigning probabilities as a function of the number of inhabitants connected to each manhole.

For the case study of Girona, a static sensor placement of five monitoring sites (or more) results in a **coverage greater than 80% of both manholes and inhabitants**. 11 iterations would be needed to detect the patient zero, and six iterations for identifying a hotspot of about 3,000 infected inhabitants. In the case of combining both algorithms, the number of iterations can be reduced to nine and four, respectively.

Any questions or comments?

Reference to paper: Optimal selection of monitoring sites in cities for SARS-CoV-2 surveillance in sewage networks

Elsevier Environment International