Pygeostat

Geostatistical data analysis tool for soil and groundwater data from DOV, application to PFAS contamination

Guillaume Vandekerckhove

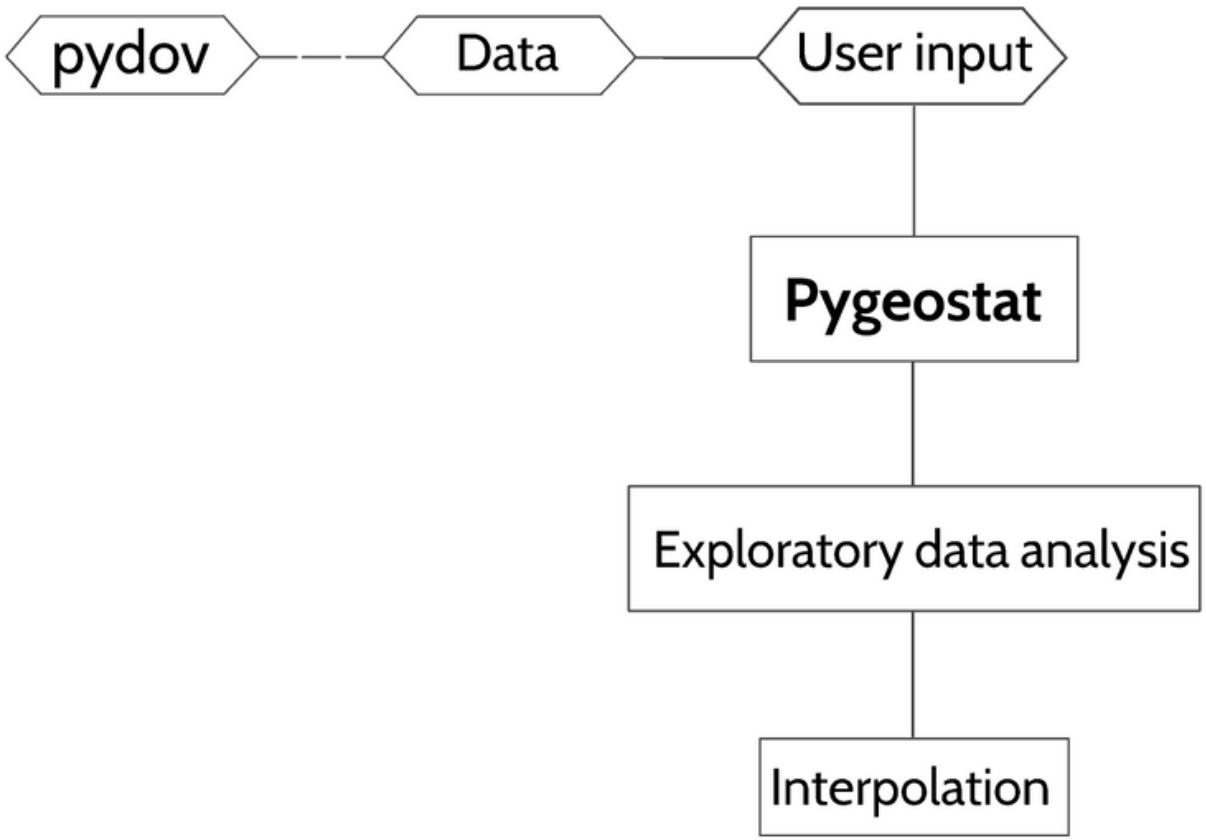
Masterthesis hydrogeology and mineral resources
Promotors: Thomas Hermans, Ellen Van De Vijver, Marleen Van Damme

Goal

Program
- Open-source and automated
- Available through DOV and github
- Will work for the data from DOV through pydov (not only PFAS or own dataset)
- Based on geostatistics

- Get familiar with the data (statistics, correlation, ...)
- Get an estimation of the data at unsampled locations
  - with a quantification of the accuracy of the estimation
- Get the probability of exceeding a threshold at unsampled locations

Results
- Part of the discussion of safety and remediation levels
- Used to prioritize the contaminated sites

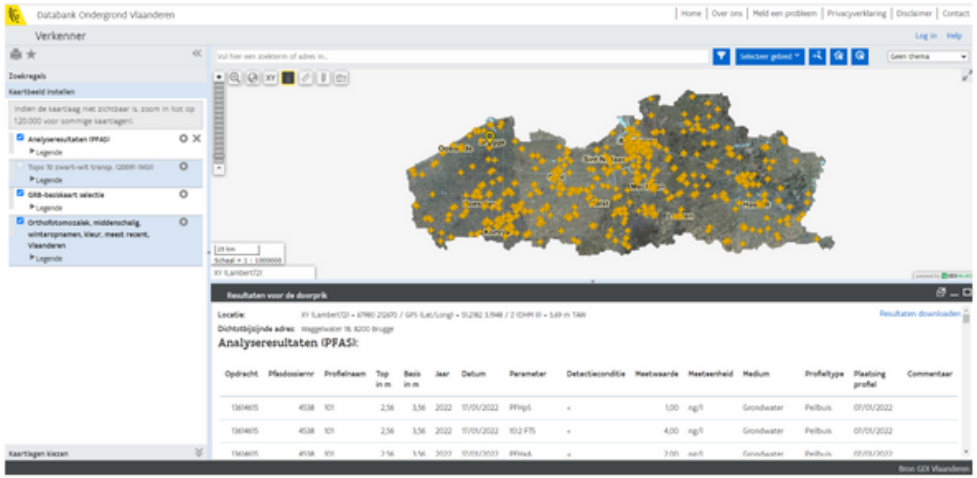

pydov
Data
User input
Pygeostat
Exploratory data analysis
Interpolation

Data

- x
- y
- top filter/borehole (m ground level)
- bottom filter/borehole (m ground level)
- date
- parameter(s) (PFOA, PFOS, etc.)
- value (ng/l or ng/kg ds)

- groundwater (ng/l) & soil (ng/kg ds) data

Data



DATABANK ONDERGROND VLAANDEREN

https://www.dov.vlaanderen.be/portaal/?module=verkenner

pydov

https://pydov.readthedocs.io/en/stable/
https://github.com/DOV-Vlaanderen/pydov.git

Data

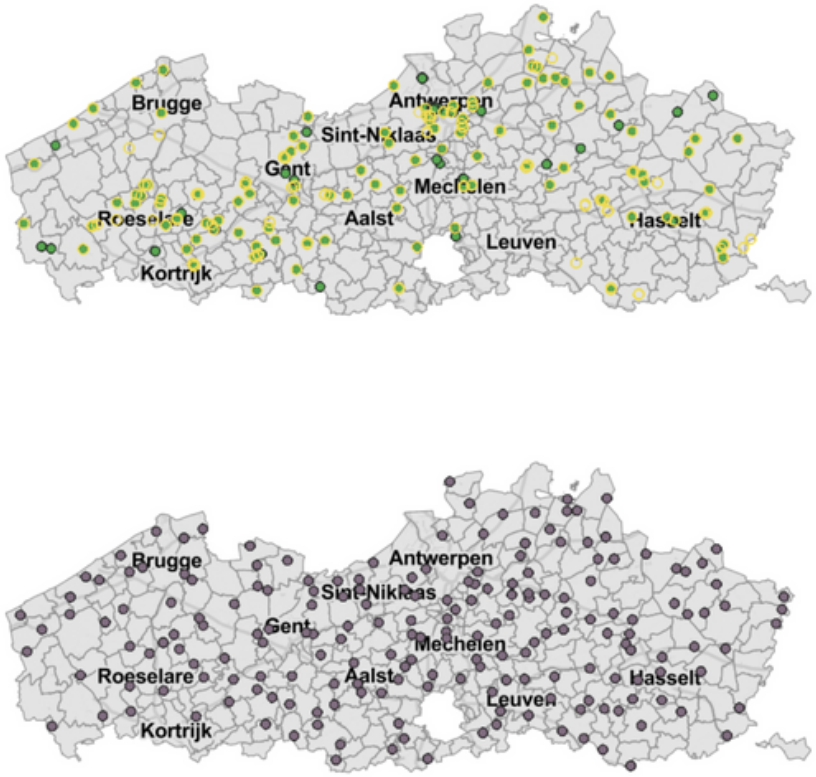
OVAM (Public Waste Agency of Flanders)
DOV (Database Flanders)

Soil
04/04/2022
769 locations

Groundwater
02/04/2022
478 locations

VMM (Flemish Evironment Agency)
pydov

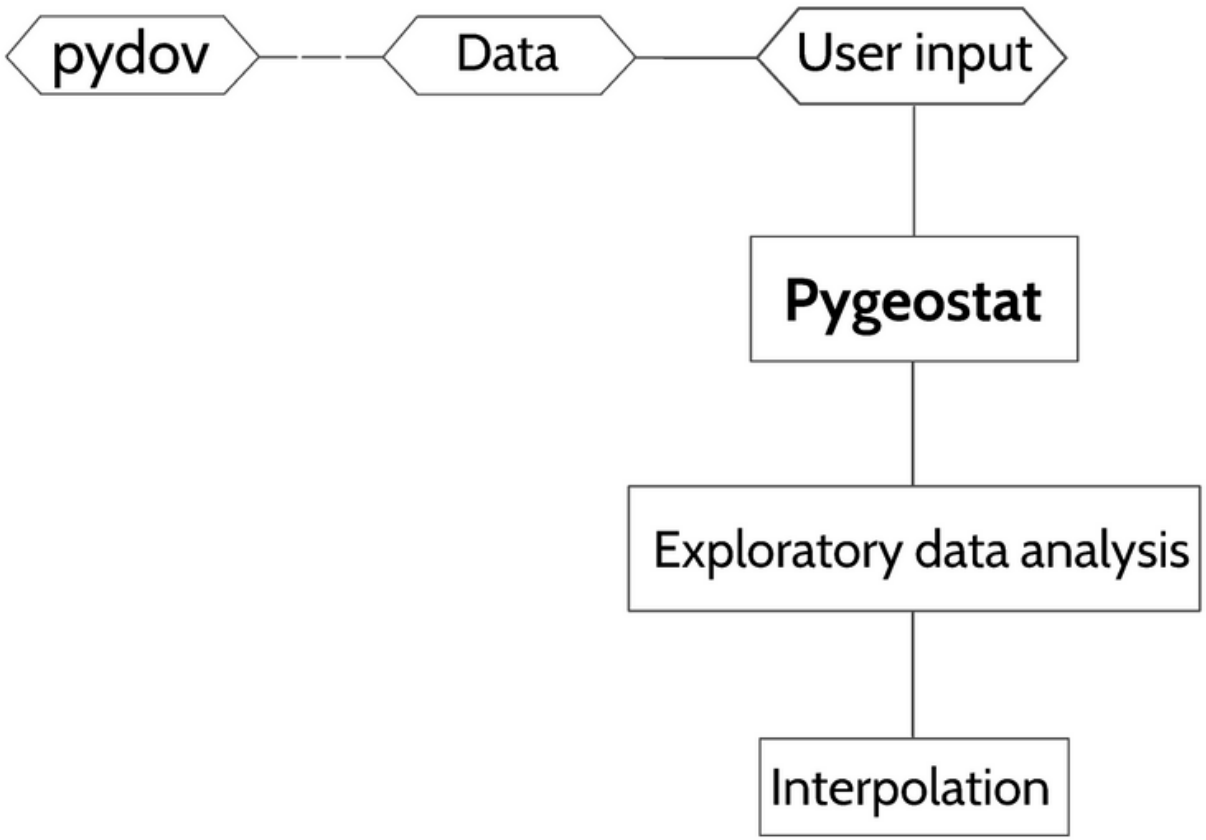
Groundwater
2022
194 locations



Data

- x
- y
- top filter/borehole (m ground level)
- bottom filter/borehole (m ground level)
- date
- parameter(s) (PFOA, PFOS, etc.)
- value (ng/l or ng/kg ds)

- groundwater (ng/l) & soil (ng/kg ds) data


pydov
Data
User input
Pygeostat
Exploratory data analysis
Interpolation

User input (selection)

General data info

- download data pydov (yes/no)
- inputfile (path to data/-)
- parameter(s) (PFOA-PFOS-EFSA4-Drinkwater20)
- log values (True/-)
- dimension (2D/3D)

Exploratory data analysis

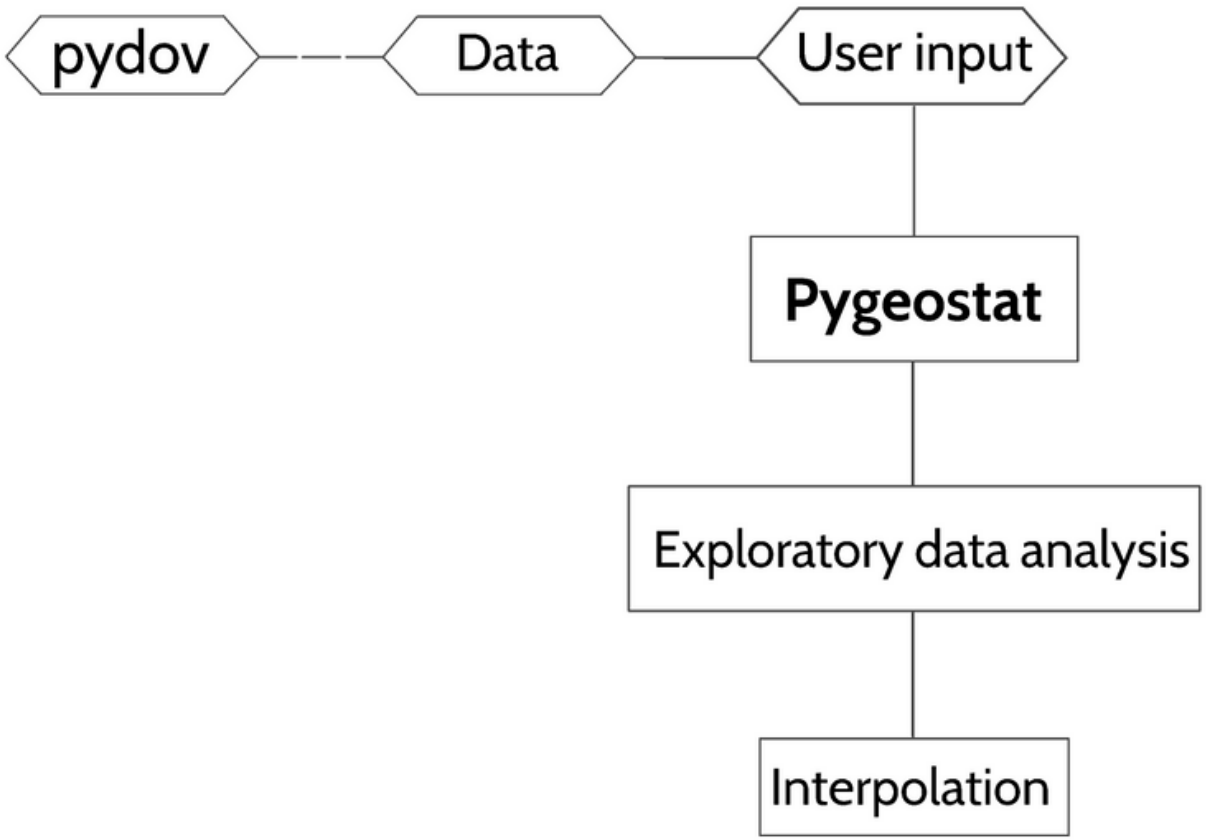
- exploration data analysis (yes/no)
- describe dataset (yes/no)
- correlation analysis (yes/no)

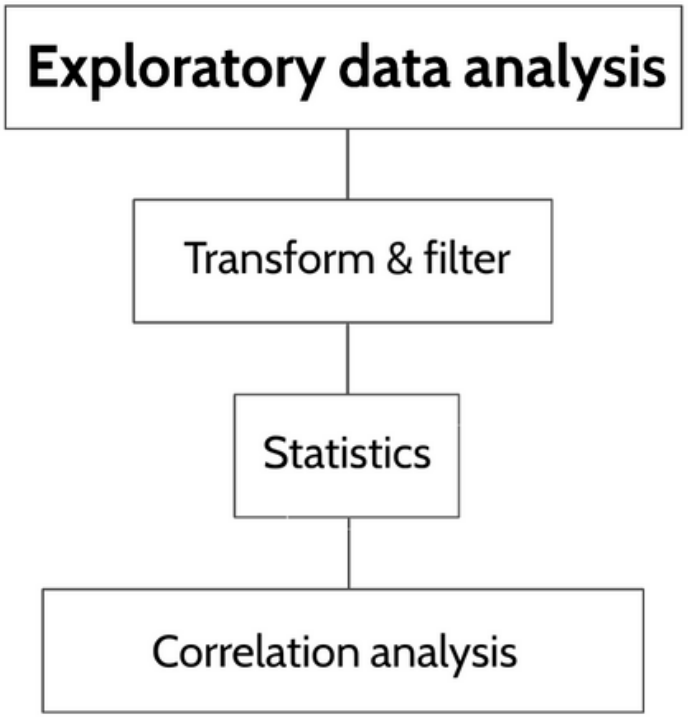
Interpolation

- interpolation (yes/no)
- ordinary or indicator kriging (OK/FIK)

PySgems

https://github.com/robinthibaut/pysgems.git


pydov
Data
User input
Pygeostat
Exploratory data analysis
Interpolation


Exploratory data analysis
Transform & filter
Statistics
Correlation analysis

Transform & filter

geometry	x	y	z_top	z_basis	date	PFOS	PFOA	EFSA4	Drinkwater20	xy-coordinates

geometry = x, y, z_top, z_basis

3D filtering (geometry):

- Check the parameter comments (f.e. PFOA = PFOA_L + PFOA_V)
- Take the most recent data
    - If multiple data points have the same date, take the **average**

-> **3D geometry is unique**

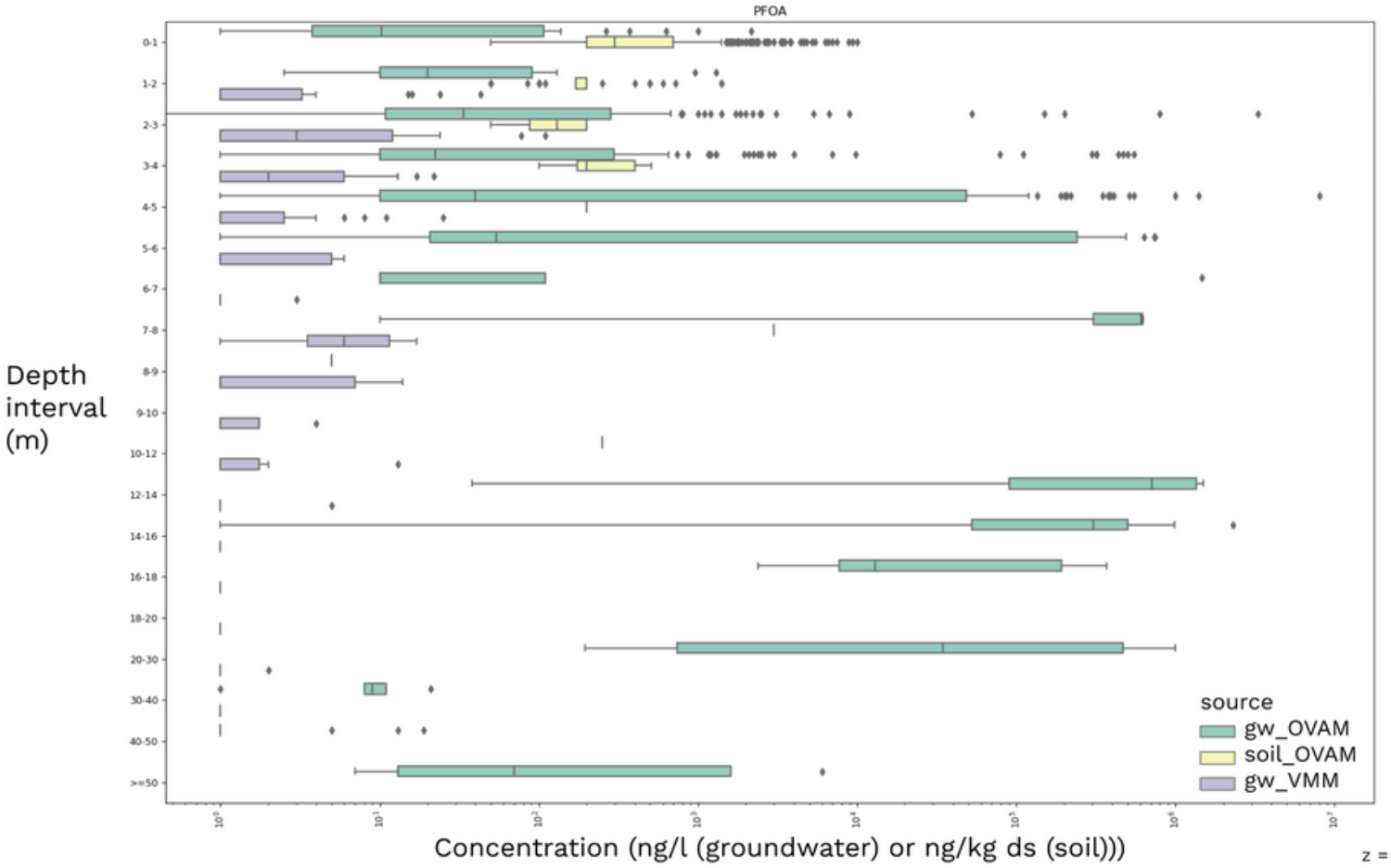
2D filtering (xy-coordinates):

- Take the most recent data
    - If multiple data points have the same date, take the **maximum** --> worst case scenario

-> **2D geometry is unique**

Statistics

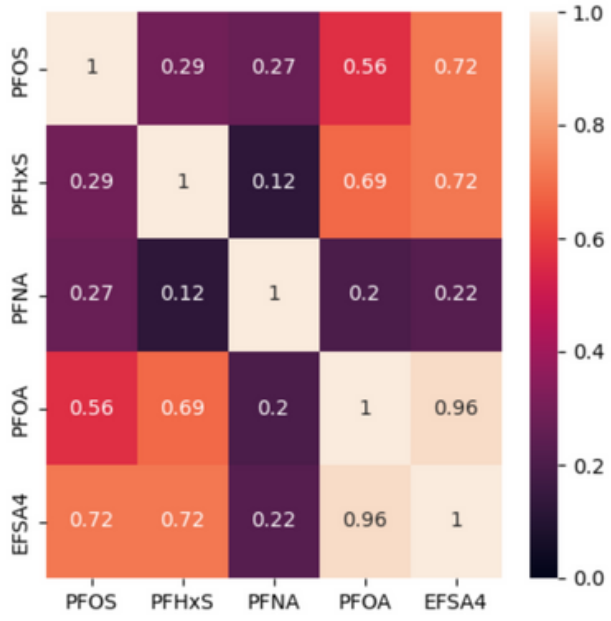
Combined 3D dataset (VMM, gw_OVAM, soil_OVAM)

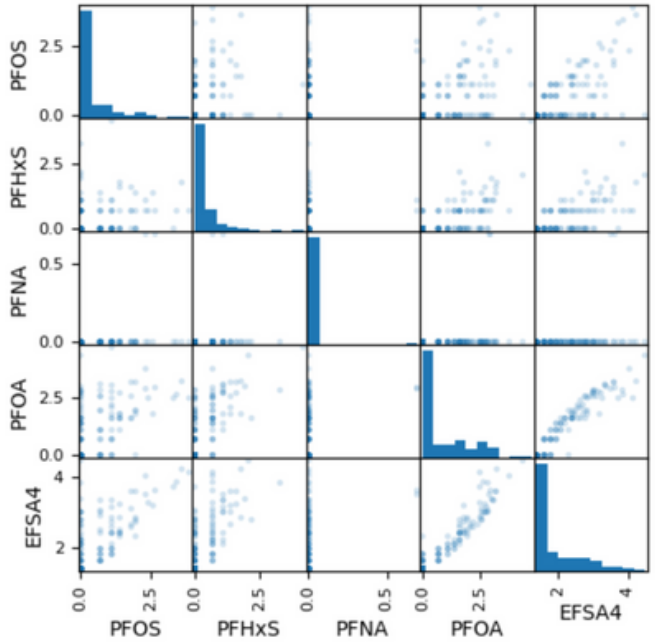


z = mean(basis filter, top filter)

Correlation analysis within same dataset

VMM data - log values

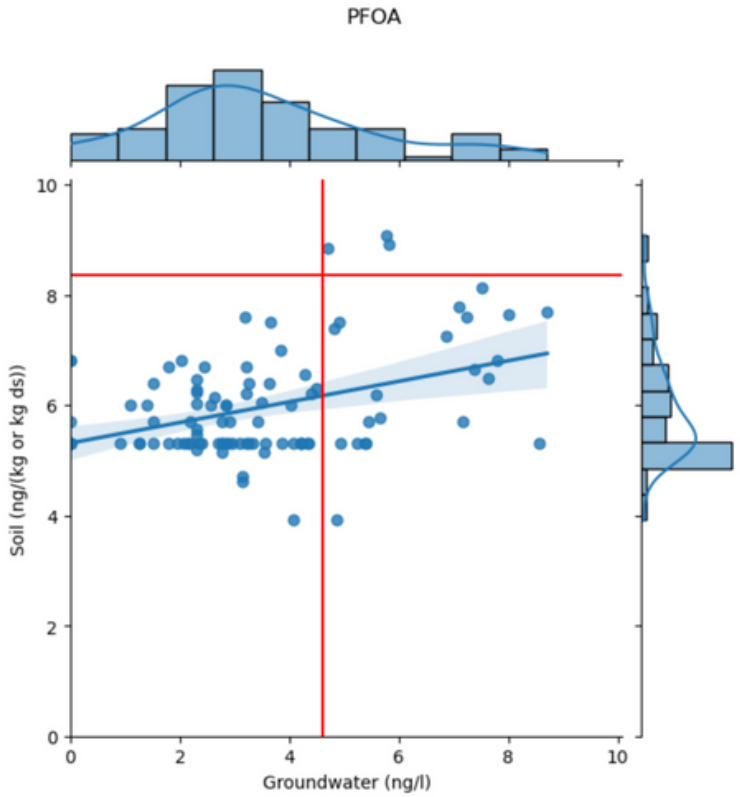




+ csv files

Correlation analysis between datasets

OVAM soil and groundwater 2D data - log values - PFOA



Legend

Best linear fit with uncertainty

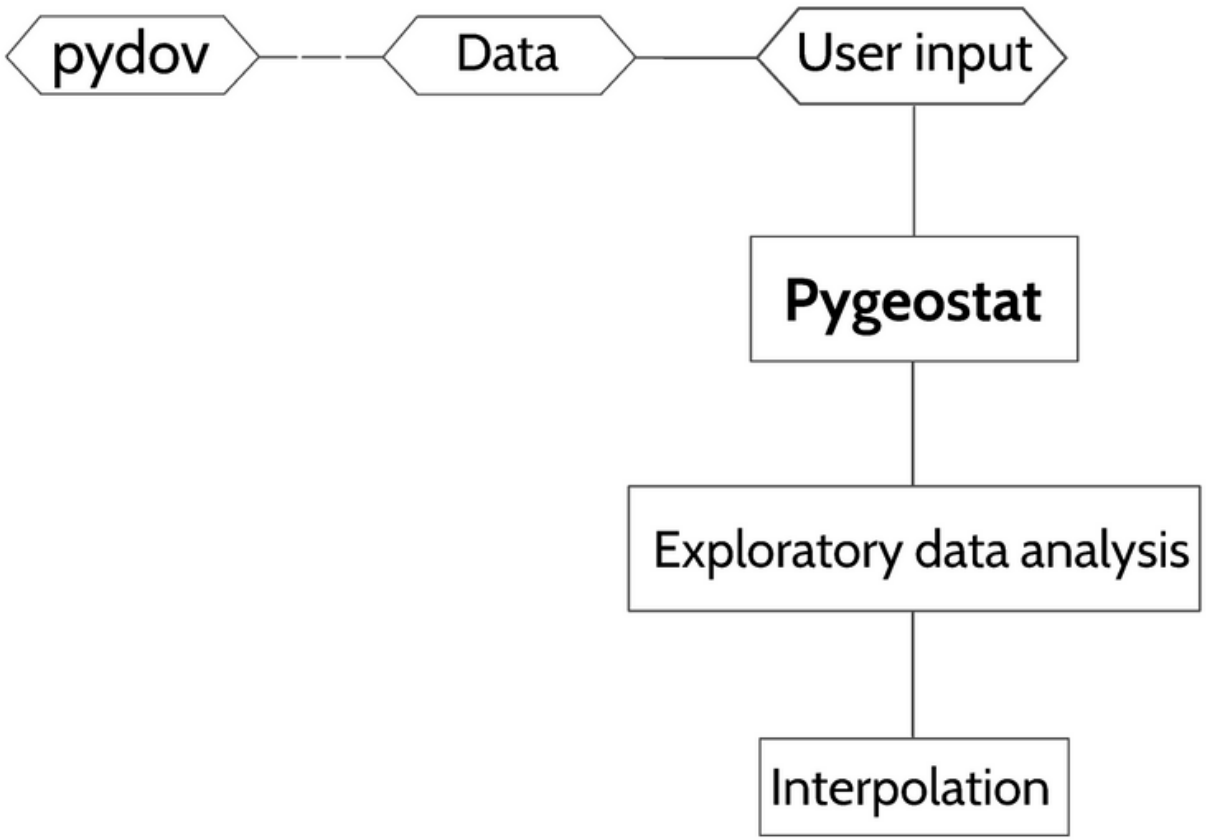
PFOA limit groundwater (100 ng/l), VITO (2020)

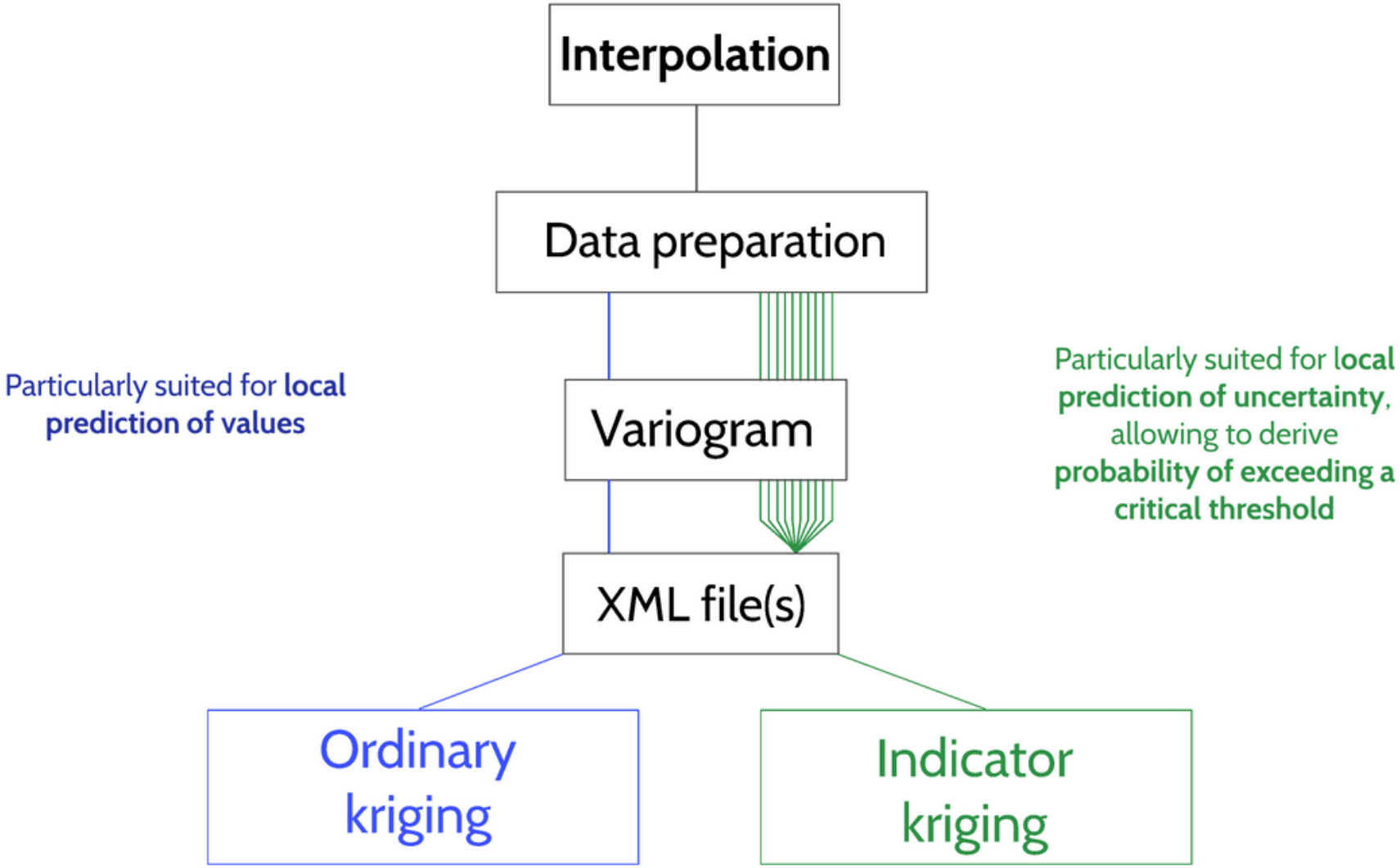
PFOA limit soil (4,3*10³ ng/kg ds), OVAM (2021)

VITO. (2020). Proposal for soil remediation values for Perfluorooctane sulfonic acid (PFOS) and perfluorooctanoic acid (PFOA). October, 189.

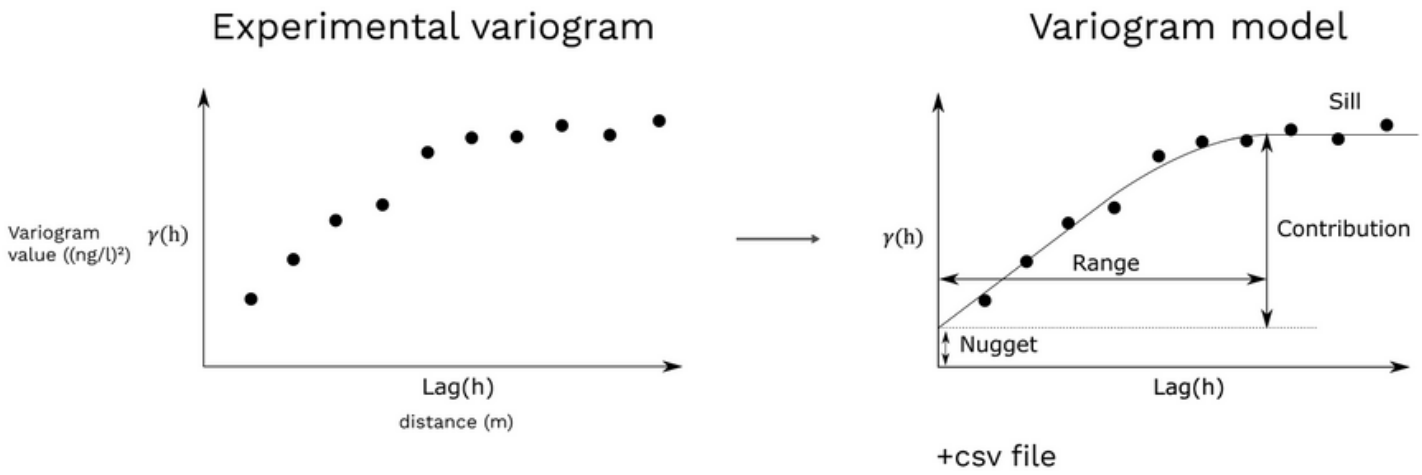
OVAM. (2021). Toetsingswaarden voor PFOS en PFOA in bodem en grondwater.

+ csv files


pydov
Data
User input
Pygeostat
Exploratory data analysis
Interpolation

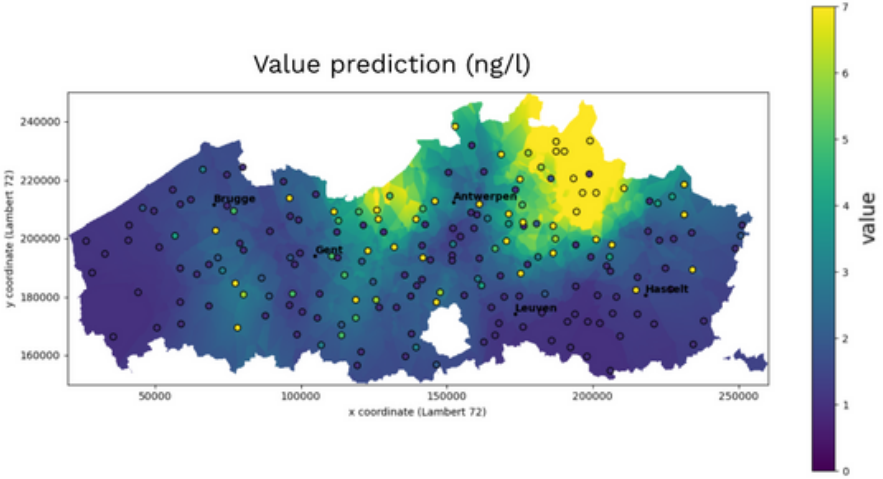

Interpolation
Data preparation
Variogram
XML file(s)
Ordinary kriging
Indicator kriging
Particularly suited for local prediction of values
Particularly suited for local prediction of uncertainty, allowing to derive probability of exceeding a critical threshold

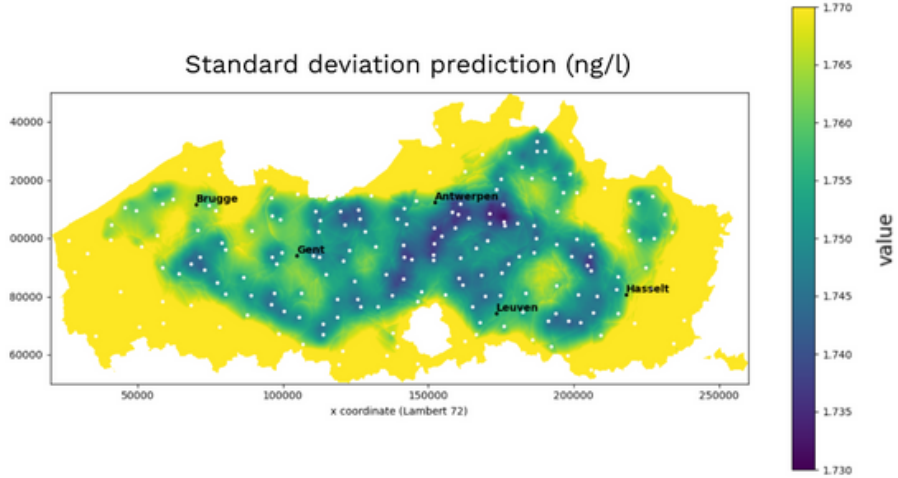
Variogram - general



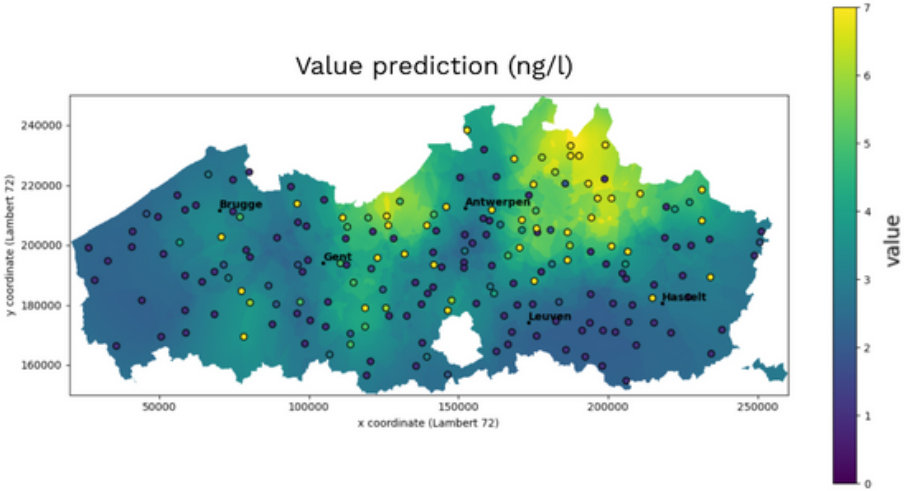
+csv file

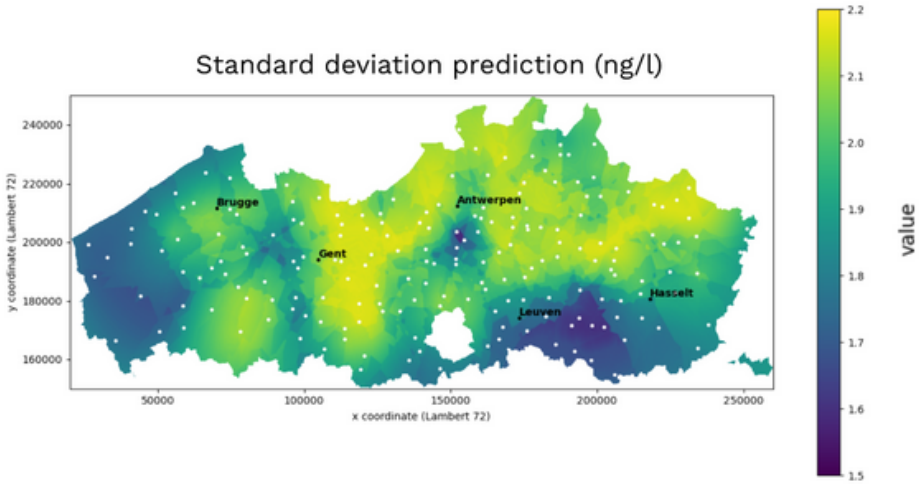
Ordinary kriging - VMM, PFOA





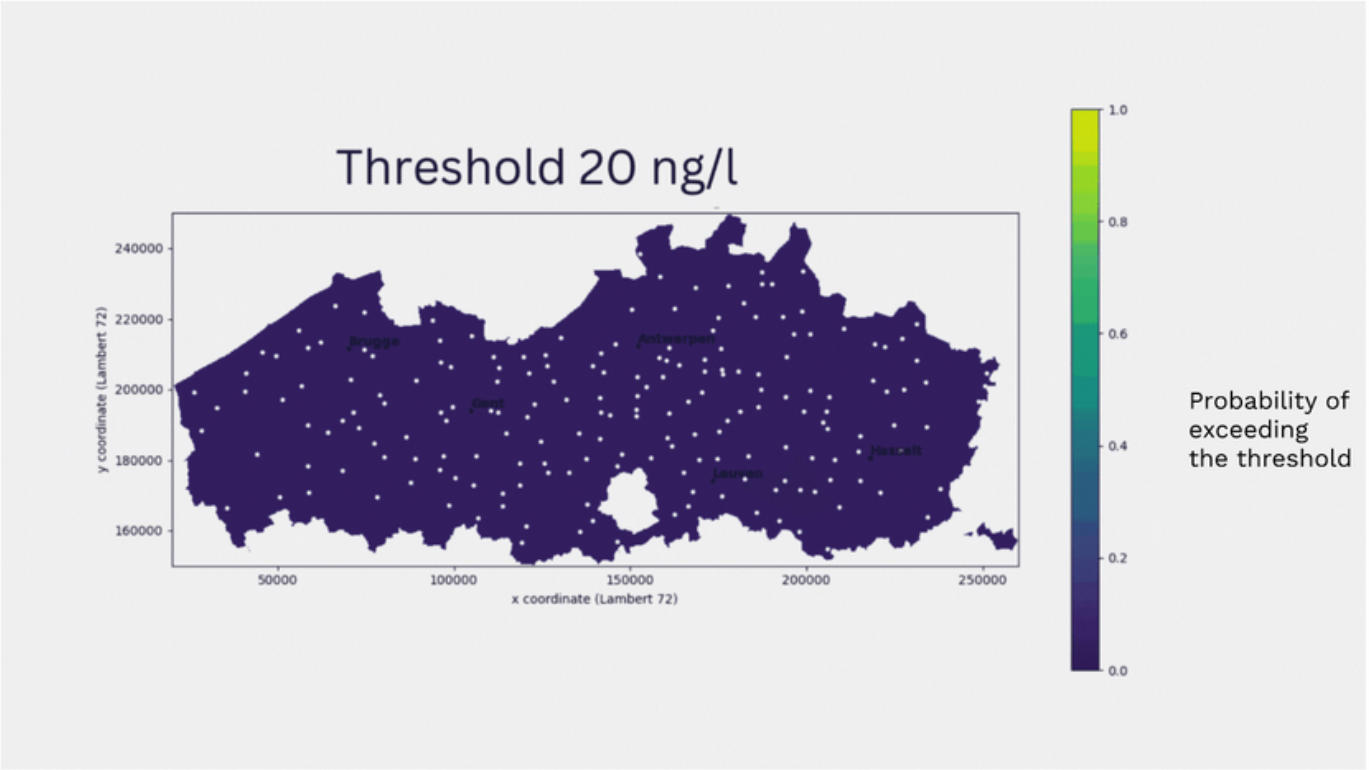
Indicator kriging - VMM, PFOA



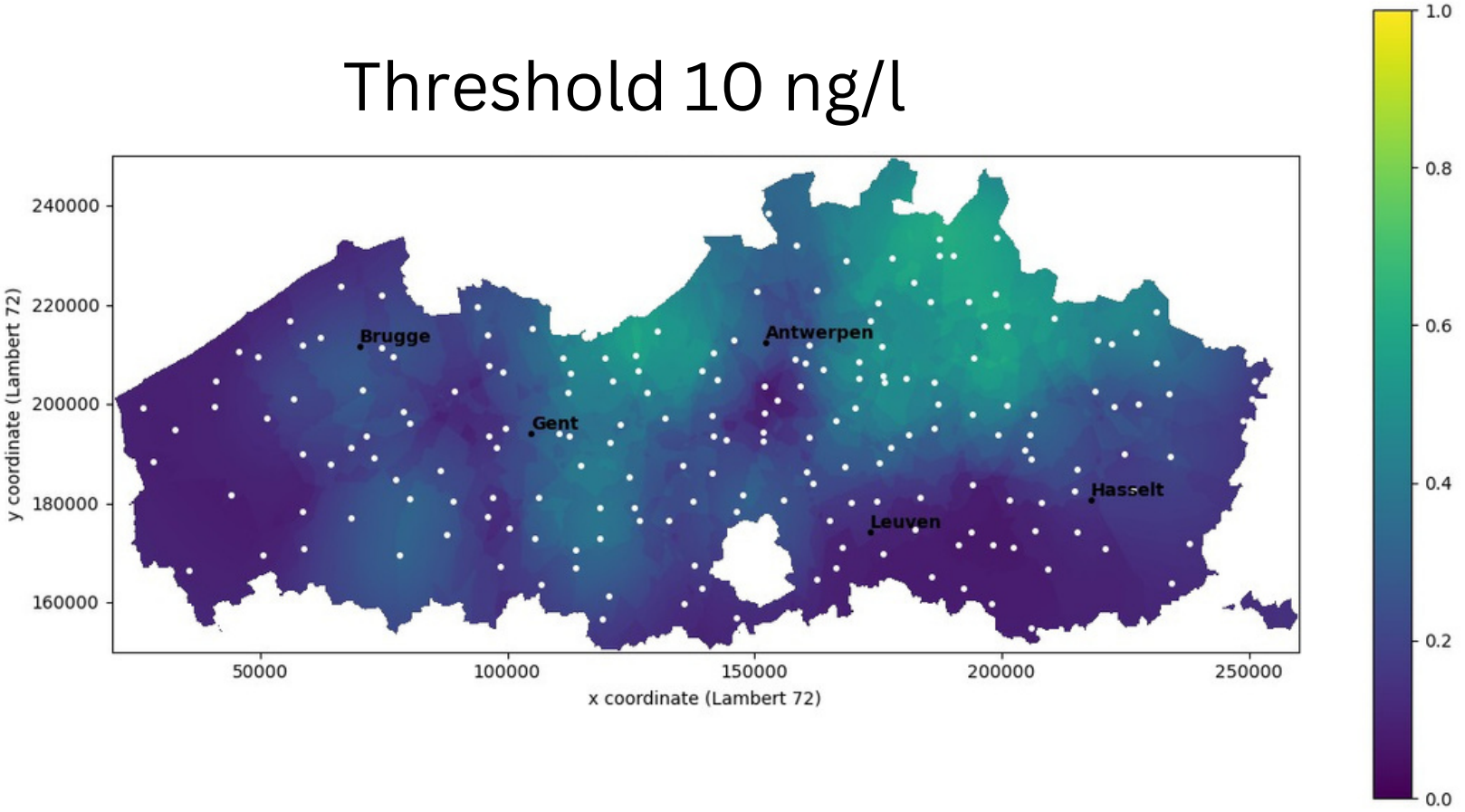


Indicator kriging - VMM, PFOA

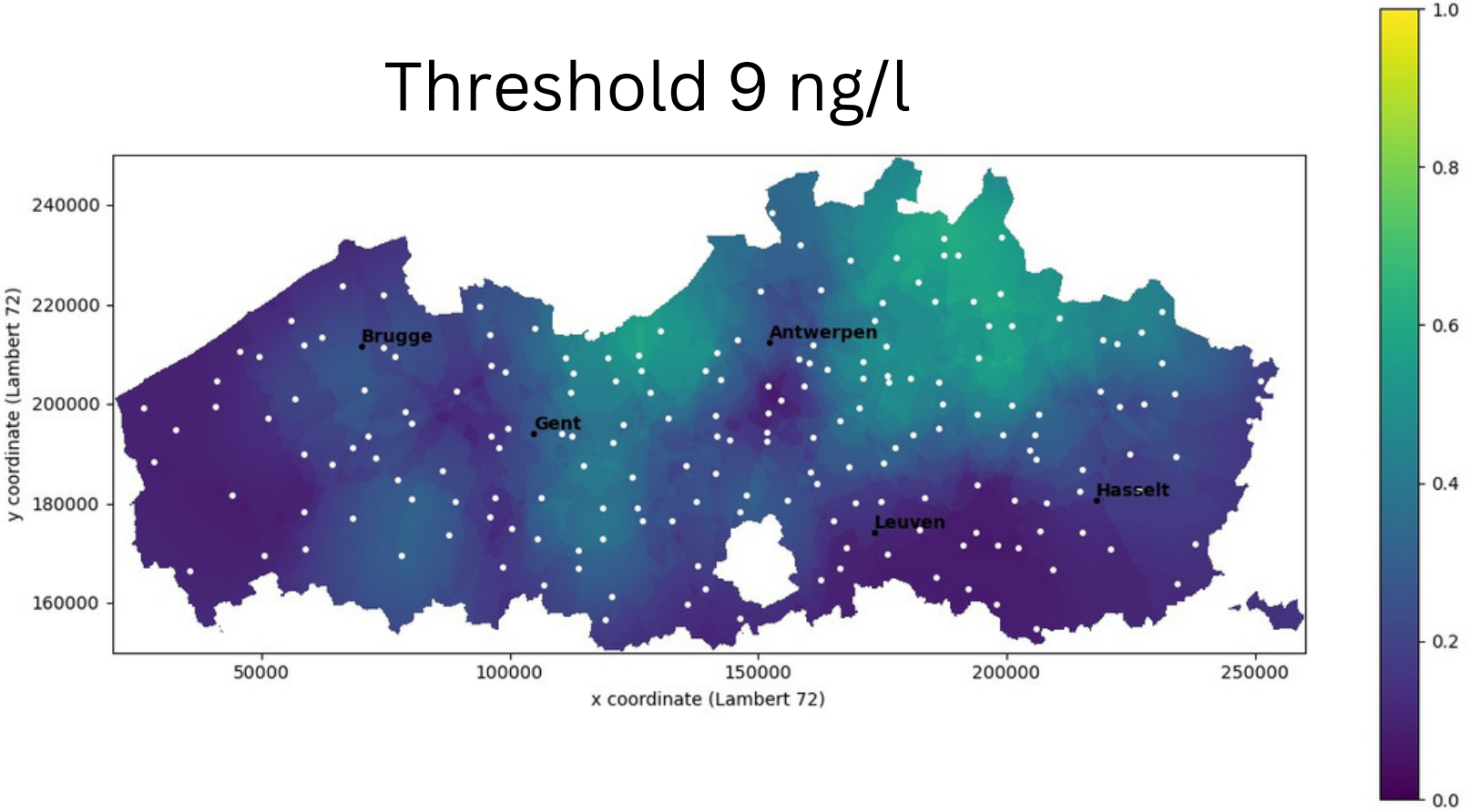
User-defined thresholds



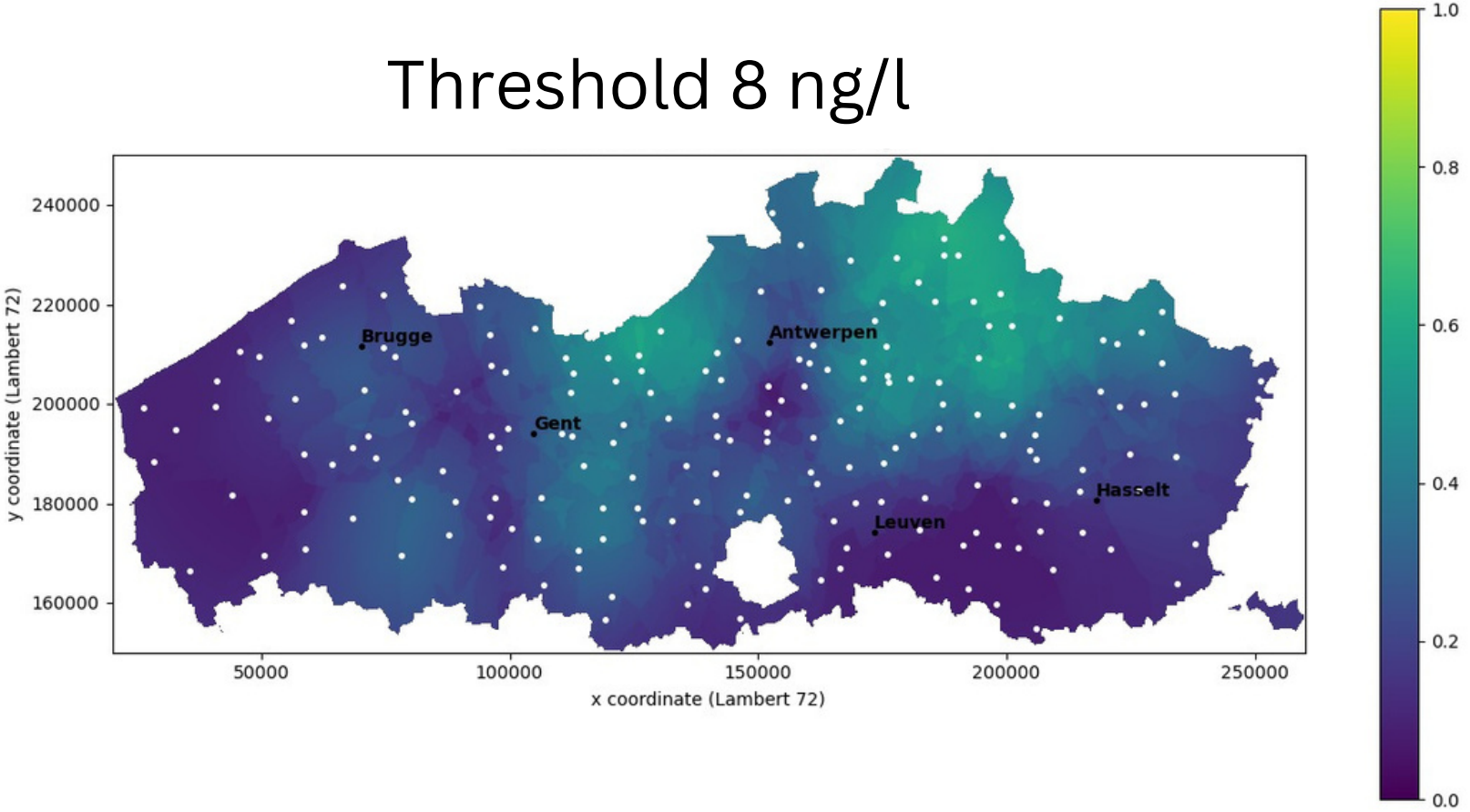
Threshold 10 ng/l



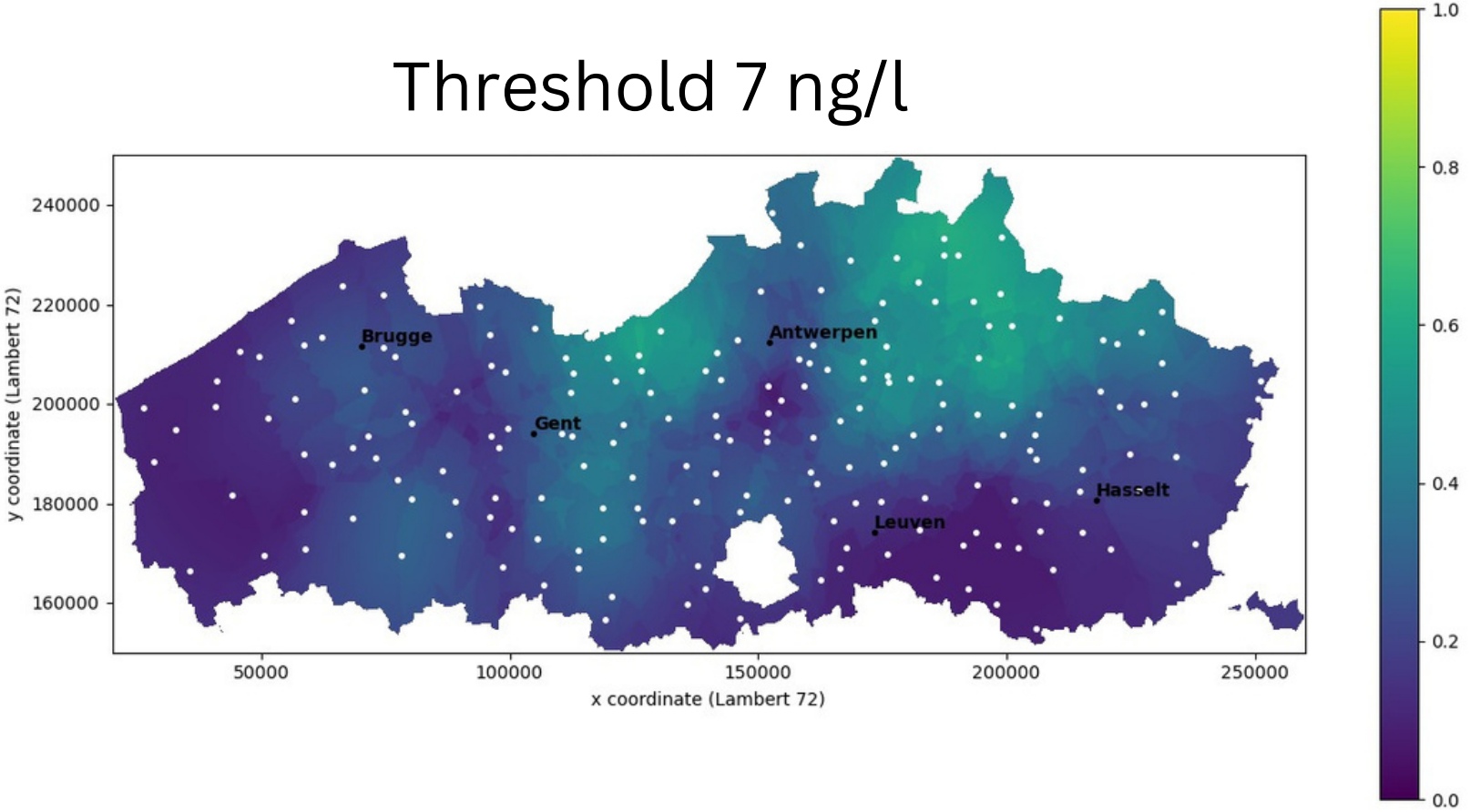
Threshold 9 ng/l



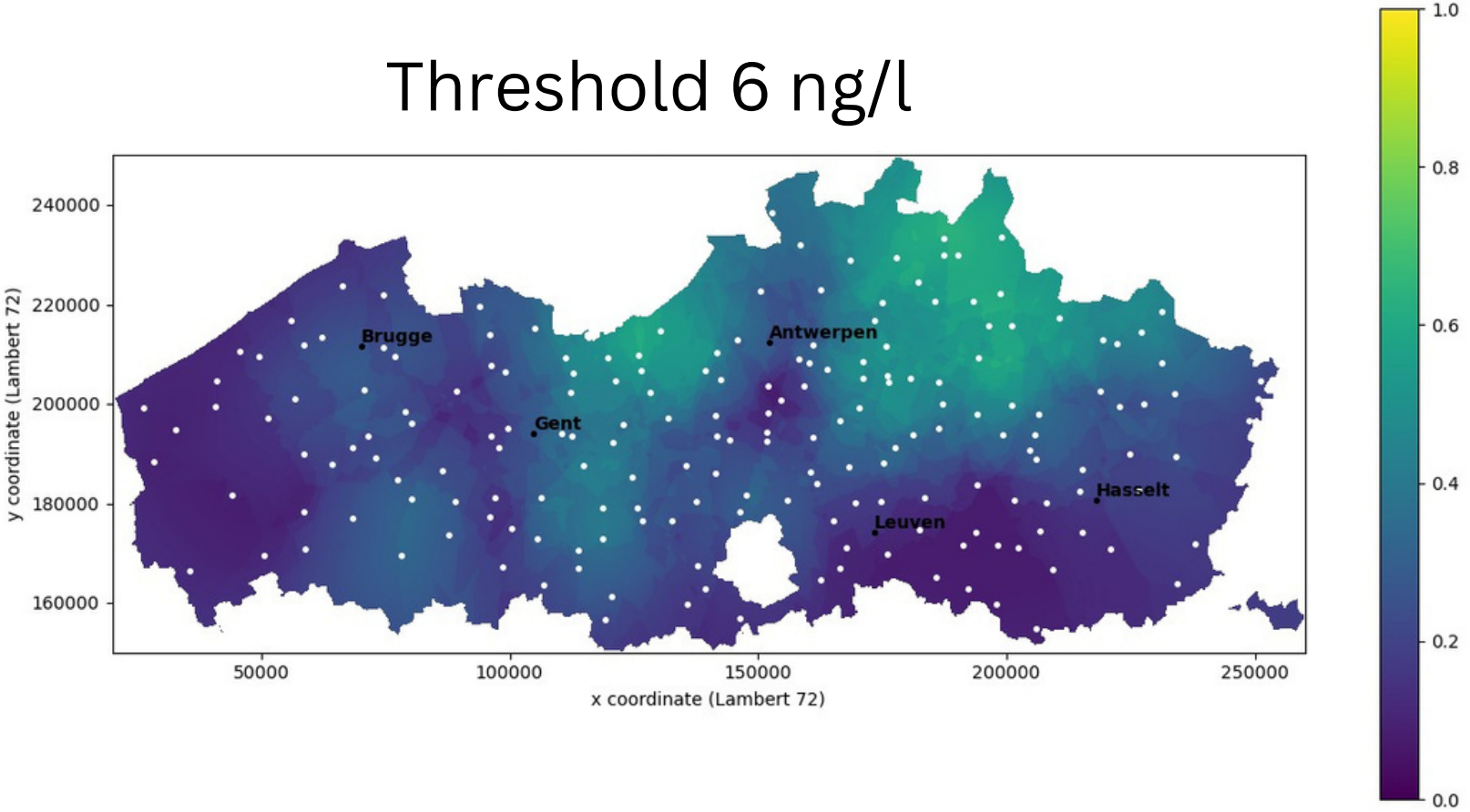
Threshold 8 ng/l



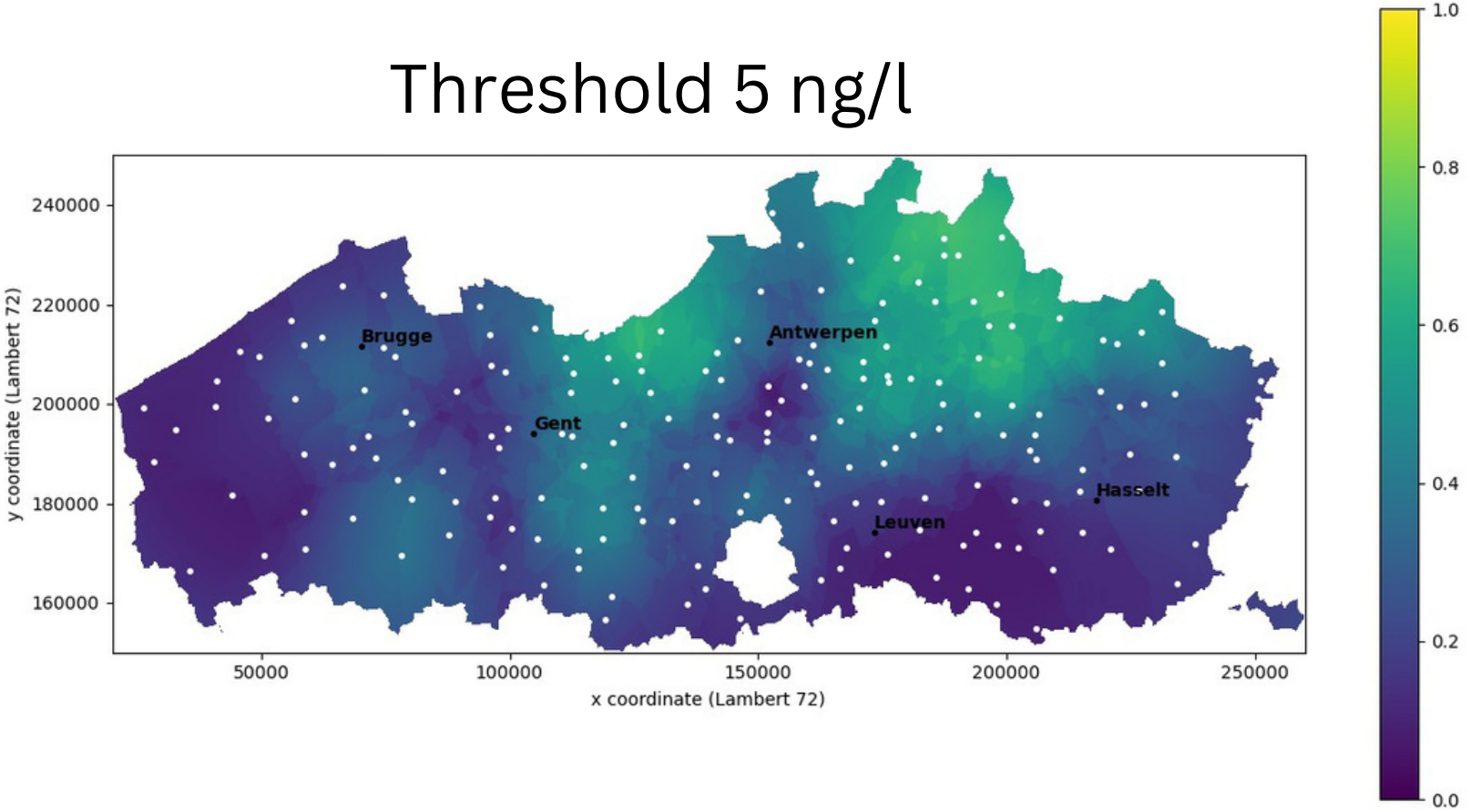
Threshold 7 ng/l



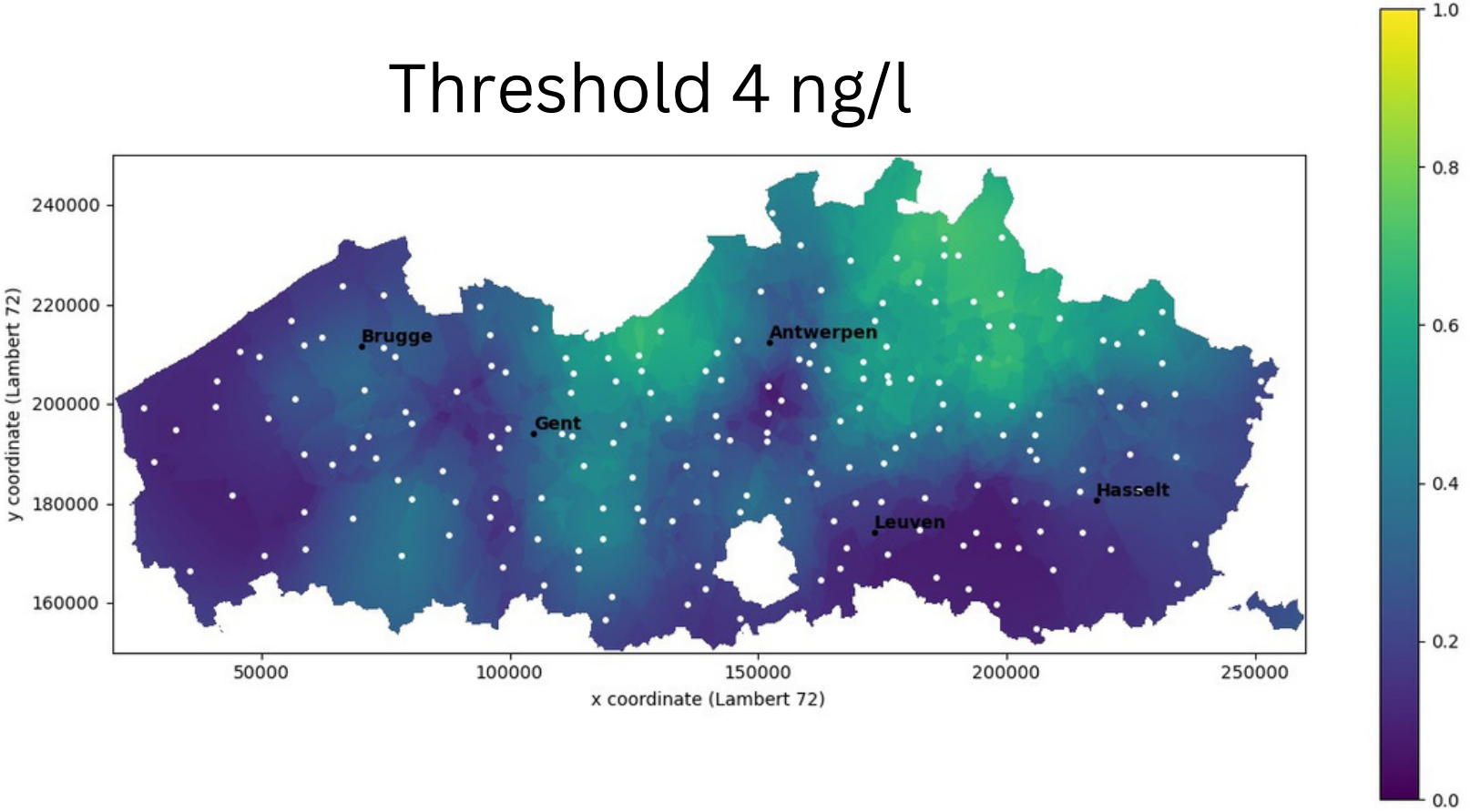
Threshold 6 ng/l



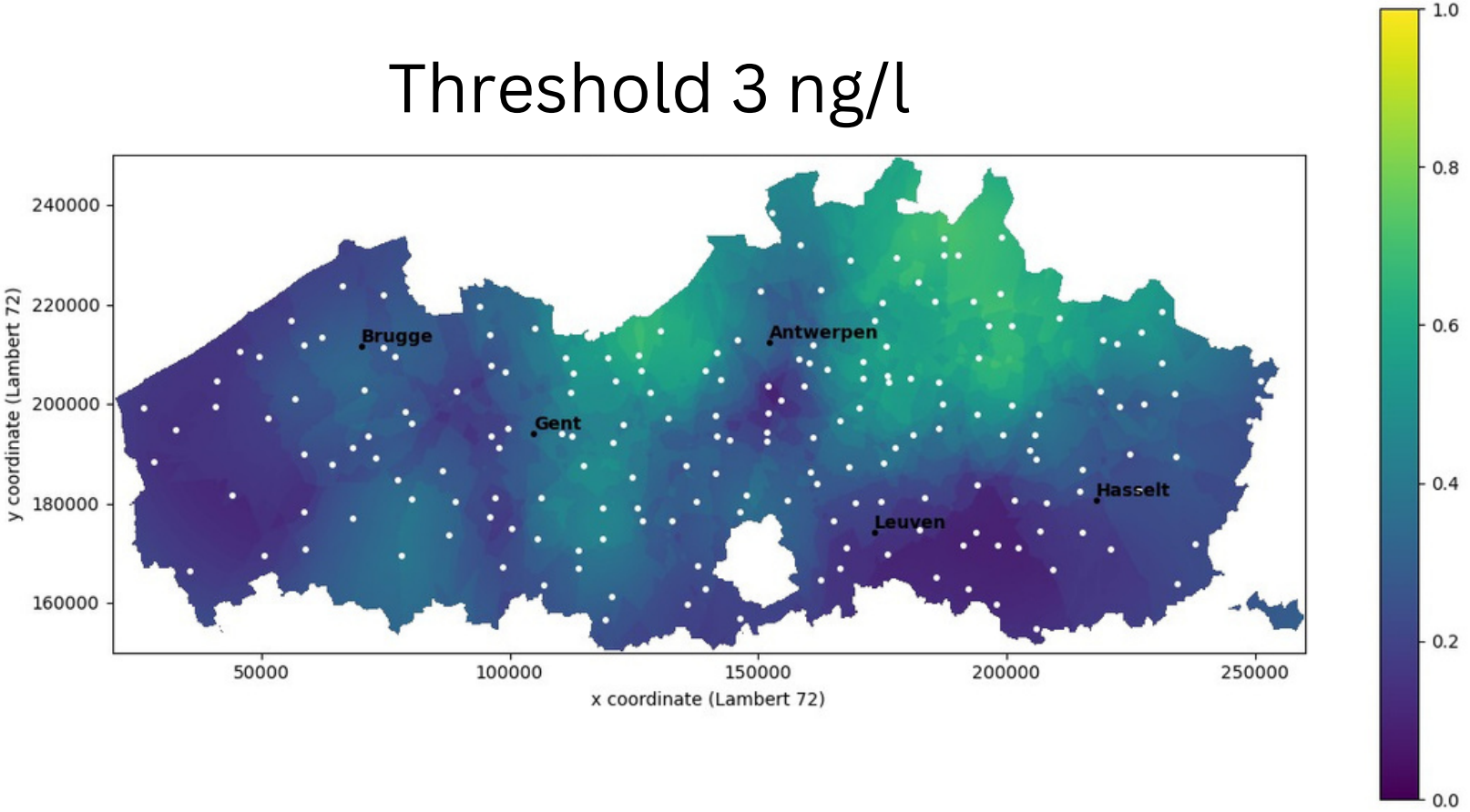
Threshold 5 ng/l



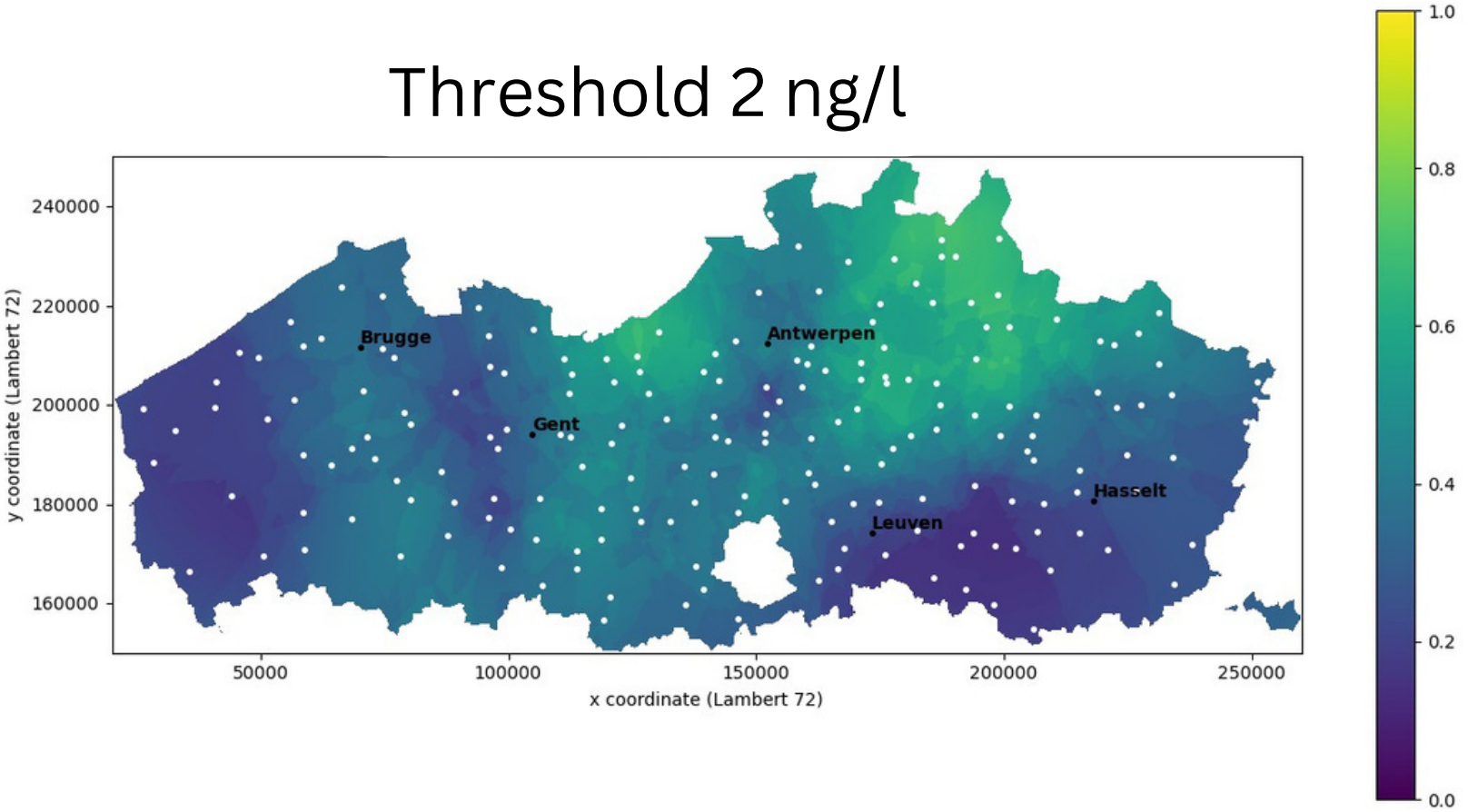
Threshold 4 ng/l



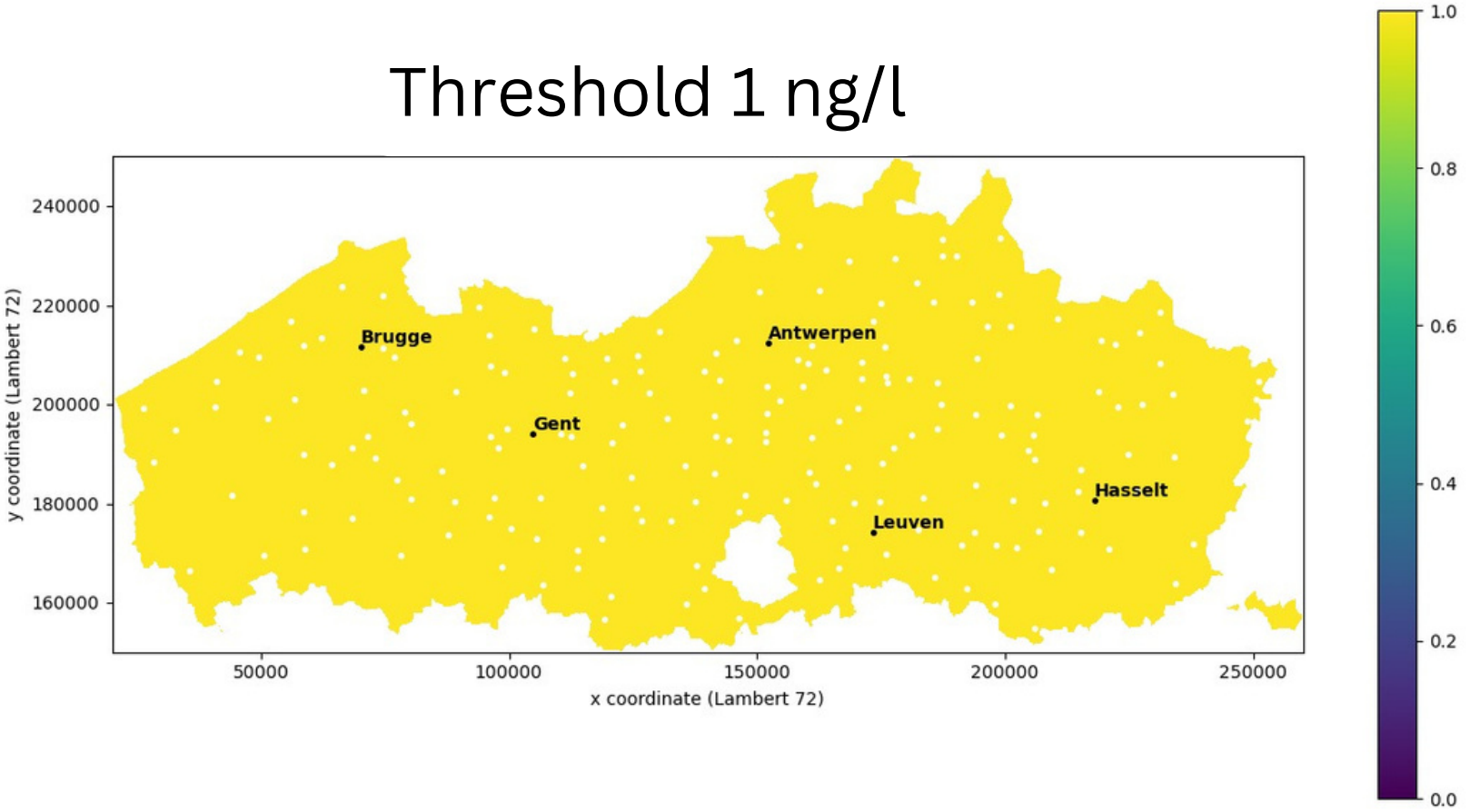
Threshold 3 ng/l



Threshold 2 ng/l



Threshold 1 ng/l



Contact

https://github.com/GuillaumeVandekerckhove

https://www.linkedin.com/in/guillaume-vandekerckhove/

guillaume.vandekerckhove@outlook.com