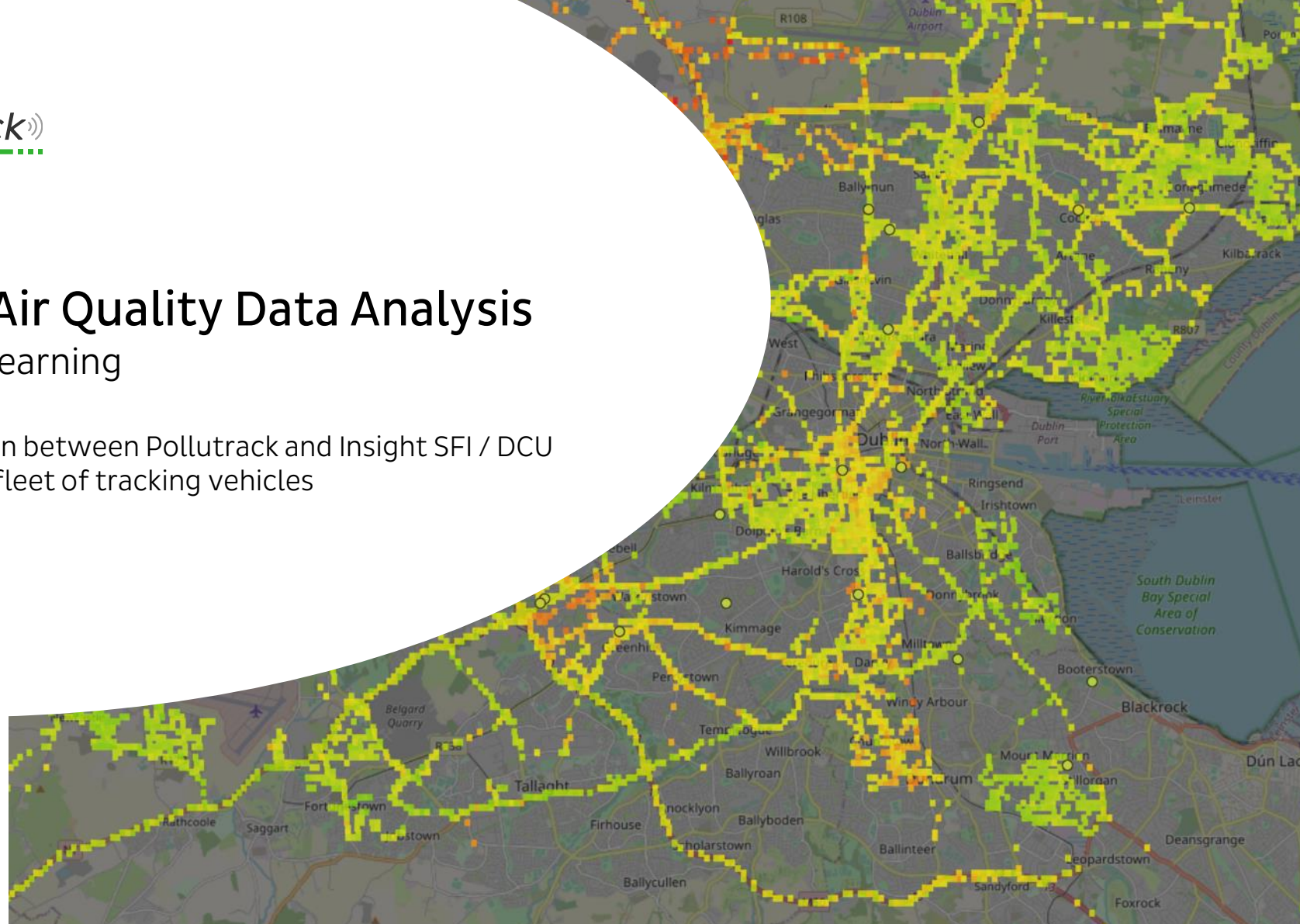




Hyperlocal Air Quality Data Analysis

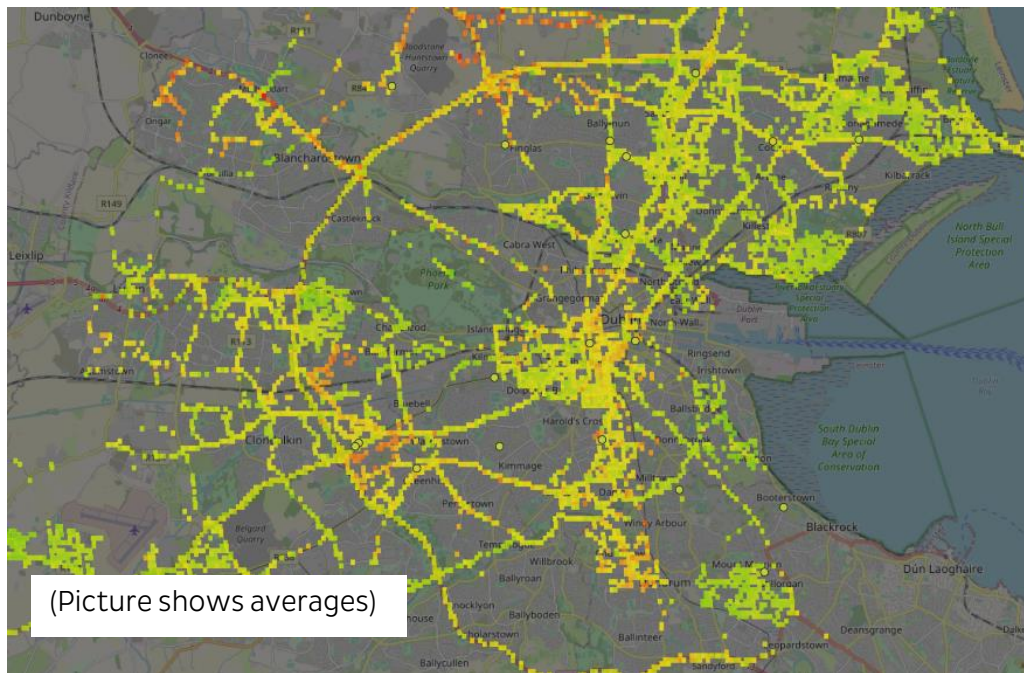
using Machine Learning

Research cooperation between Pollutrack and Insight SFI / DCU
supported by DPD's fleet of tracking vehicles





Raw Pollutrack Data -> individual datasets



Coordinates system: WGS84 (EPSG:4326)

1) Mobile sensors:

- PM2.5 data, every 12 seconds
- Moving across the city -> street coverage
- Daily sensor validation (cross-check with fixed stations)
- Measurement when vehicle is in use

Variables :

- Latitude, longitude
- Timestamp
- PM2.5
- (Sensor ID)

2) Fixed Stations:

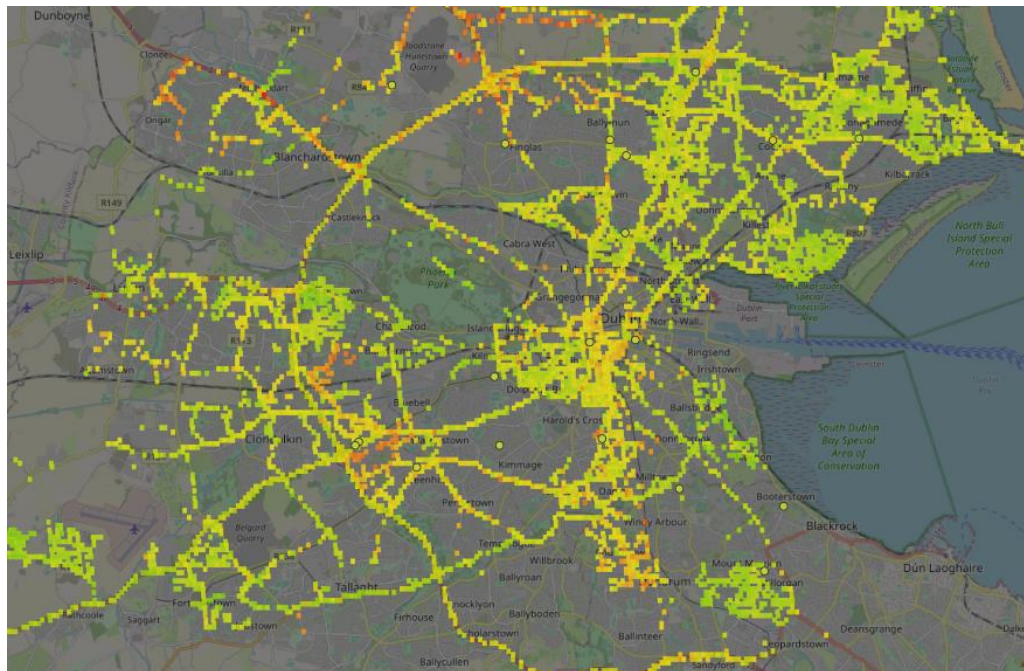
- PM2.5 data, every 30 seconds
- 10-20 locations un Dublin (Latitude, longitude -> constant for each sensor)
- Triple redundancy -> highly reliable
- 24/7 measurement

Variables :

- Sensor ID -> Latitude, longitude
- Timestamp
- PM2.5 ($\mu\text{g}/\text{m}^3$)



Spatio-temporal Data arrays
-> (1/2)hourly averages over 100x100m squares, sparse



Recommended grid:

100m x 100m grid (EPSG:3857)

½ hourly averages (48 values per day)

PM2.5 ($\mu\text{g}/\text{m}^3$)

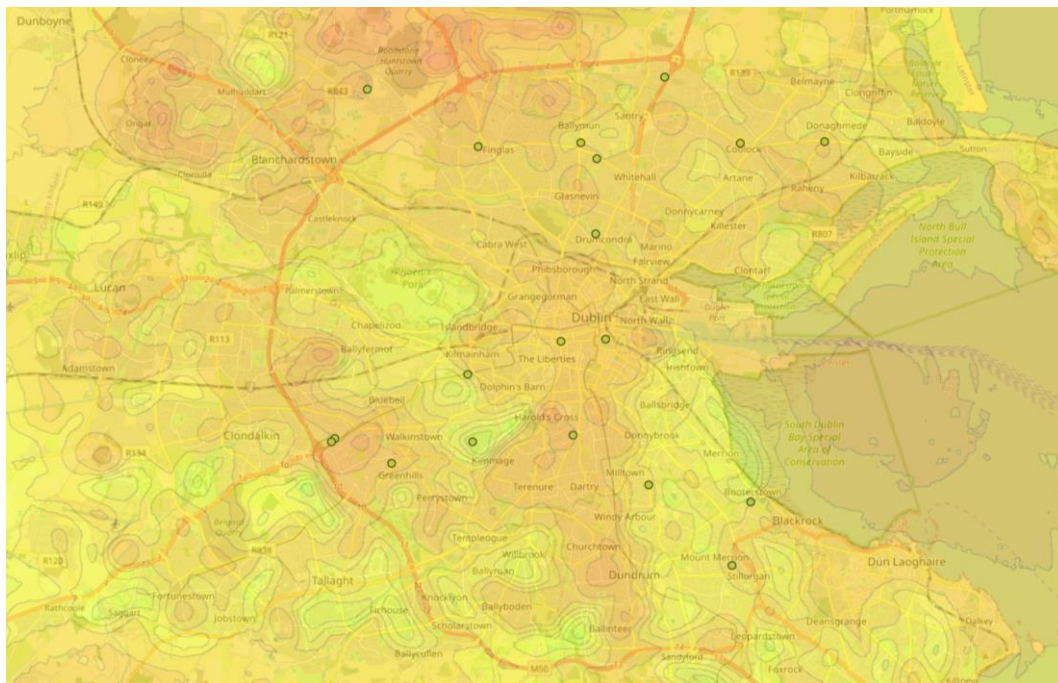
Note: due to the nature of mobile sensors,
the data is **spatially sparse**
(no complete coverage).
except in fixed station locations
(full temporal coverage)

100x100m is what we call « hyperlocal »
compared with other meteorological scales
(smallest scale usually about 10x10km).

Recommended projection system: WGS 84 / Pseudo-Mercator (EPSG:3857)



Spatio-temporal inter/extrapolation -> Heatmap



Today on Pollutrack's dashboard:

Simple spatial interpolation
(distance-weighted averages)
No machine learning, no temporal extrapolation
No influence of other variables
(streets, altitude, wind etc)

Main goal of the research:

Improve spatio-temporal inter/extrapolation
using machine learning

-> air quality patterns are somewhat consistent
from one day to the other, taking into account
hour and season (and other variables)

Machine learning should be able to « catch » this
consistency and provide advanced real-time
« best-of-knowledge » maps for decision-making
at the political level (city) and at the individual
level (personal behavior).

Prevision (temporal extrapolation) would be
great.