Improving data quality in low-cost sensor networks for air quality

José M. Barcelo-Ordinas (joseb@ac.upc.edu)

SANS research group (UPC, http://sans.ac.upc.edu)

SANS research group

- **Faculty:** Jorge García-Vidal, Jose M. Barceló-Ordinas and Manel Guerrero-Zapata

- **PhD students:** Pau Ferrer-Cid, Juan A. Paredes-Ahumada, Xhensilda Allka

- **MsC/BsC students:** Aina Nadal-Main, Zhe Ye, Raul Povedano

Data quality in heterogeneous monitoring IoT networks

- **Long term vision:**
  - To design low-cost sensor networks with low maintenance cost and high-data quality
- **Challenges:**
  - How to obtain accurate monitoring data with low cost sensing nodes?
  - How to automatically identify nodes that do not operate correctly?
- **Tools:**
  - Advanced machine learning techniques
  - Build our own low-cost hardware
  - Access to real data: reference stations, testbeds, etc,

Data quality in heterogeneous monitoring networks
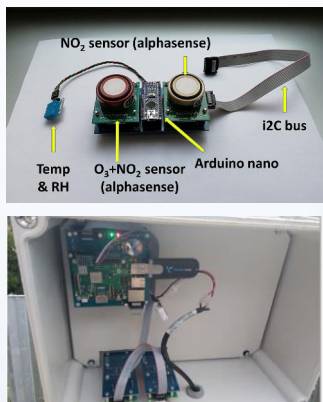
Air quality monitoring (CAPTOR nodes)

Mobility & activity monitoring

Cloud

Machine learning

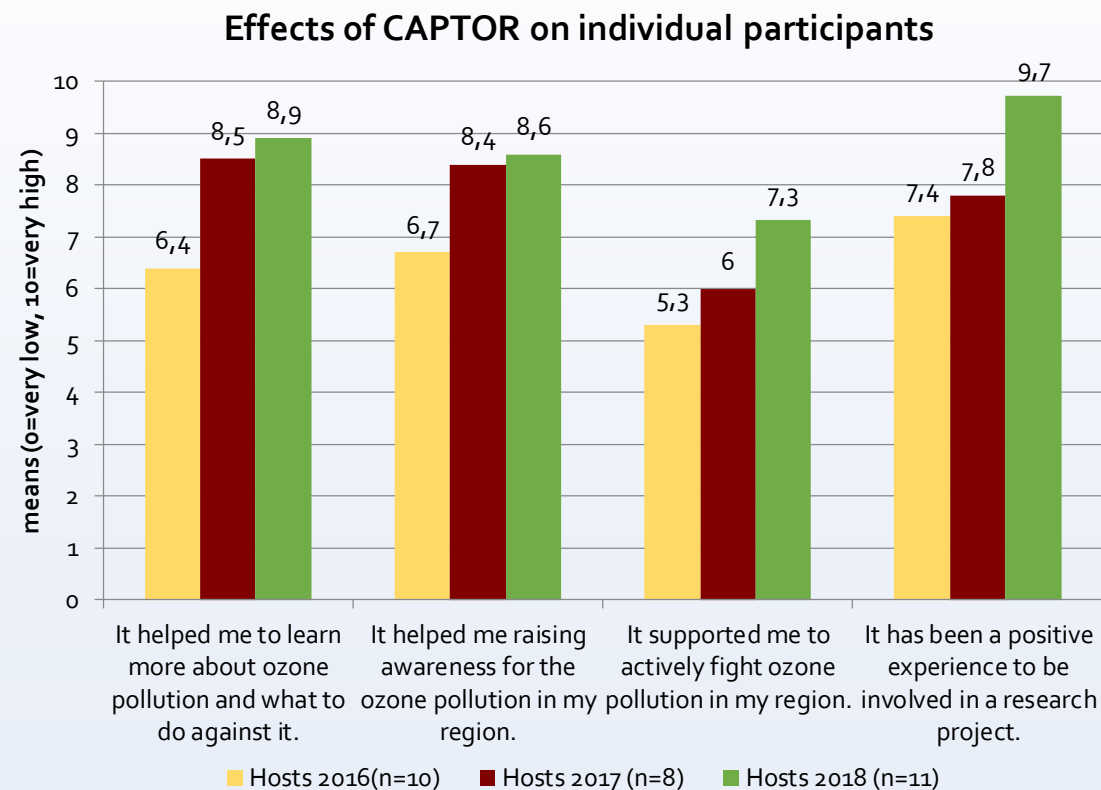Final Project Event – 29th and 30th June 2021

Citizen science H2020 project (2016-2018), UPC coordinator

- **OBJ: arise awareness among citizens:**
  - Testbeds in Catalonia (Vic, Manlleu and Tona), Italia and Austria,
  - Mainly O_3 as main pollutant (MOX and electro-chemical O_3 sensors)
 - Nodes deployed in schools and volunteer houses,
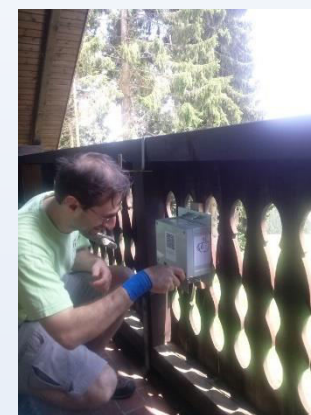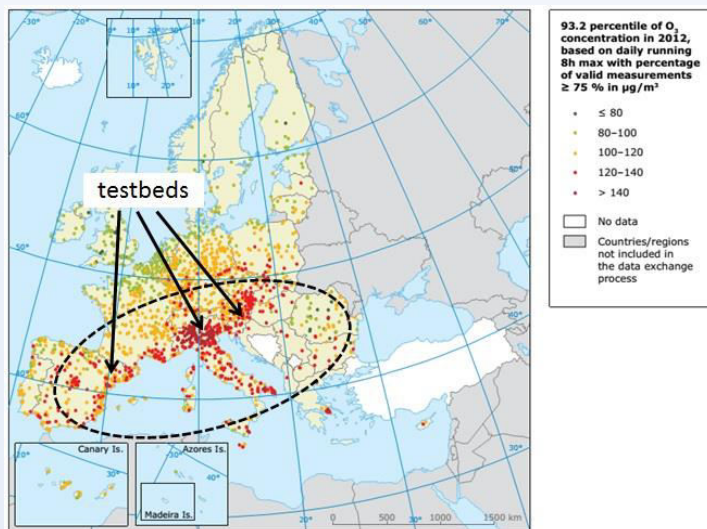
Citizen science H2020 project (2016-2018), UPC coordinator

- Captor considerably **helped to raise awareness for** O_3 in the whole Vic valley

- Dissemination activities let to an increasing interest in the topic and finally the uptake of it in the local television

- Municipality installed a public screen, where air quality data including ozone are shown

- Pressure for more transparency and strong political involvement of local organisations resulted in legal action to fight origins of O_3

Effects of CAPTOR on individual participants

Chart y-axis: means (o=very low, 10=very high)

Values:
- It helped me to learn more about ozone pollution and what to do against it: 6,4 / 8,5 / 8,9
- It helped me raising awareness for the ozone pollution in my region: 6,7 / 8,4 / 8,6
- It supported me to actively fight ozone pollution in my region: 5,3 / 6 / 7,3
- It has been a positive experience to be involved in a research project: 7,4 / 7,8 / 9,7

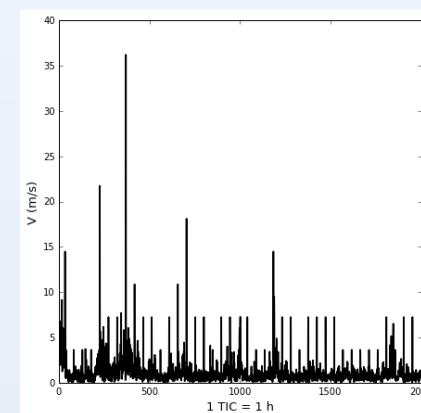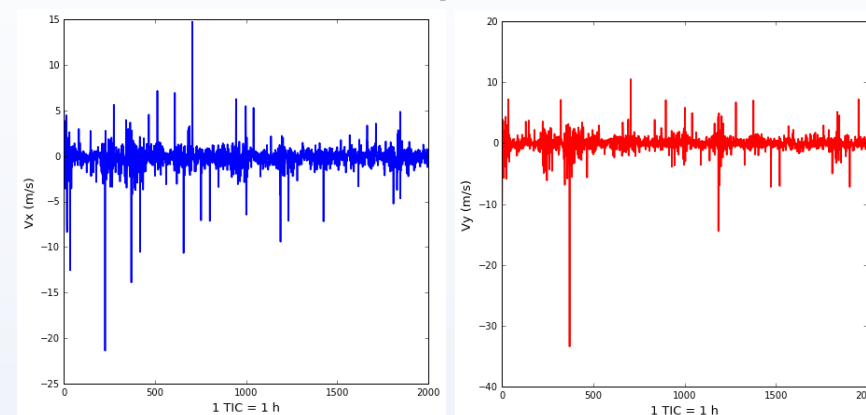Legend: ■ Hosts 2016(n=10) ■ Hosts 2017 (n=8) ■ Hosts 2018 (n=11)

Other initiatives: node installed at C/Castilla (BCN)

- In collaboration with municipality (framework of H2020 Growsmarter project)

- Node installed at C/Castilla (BCN)
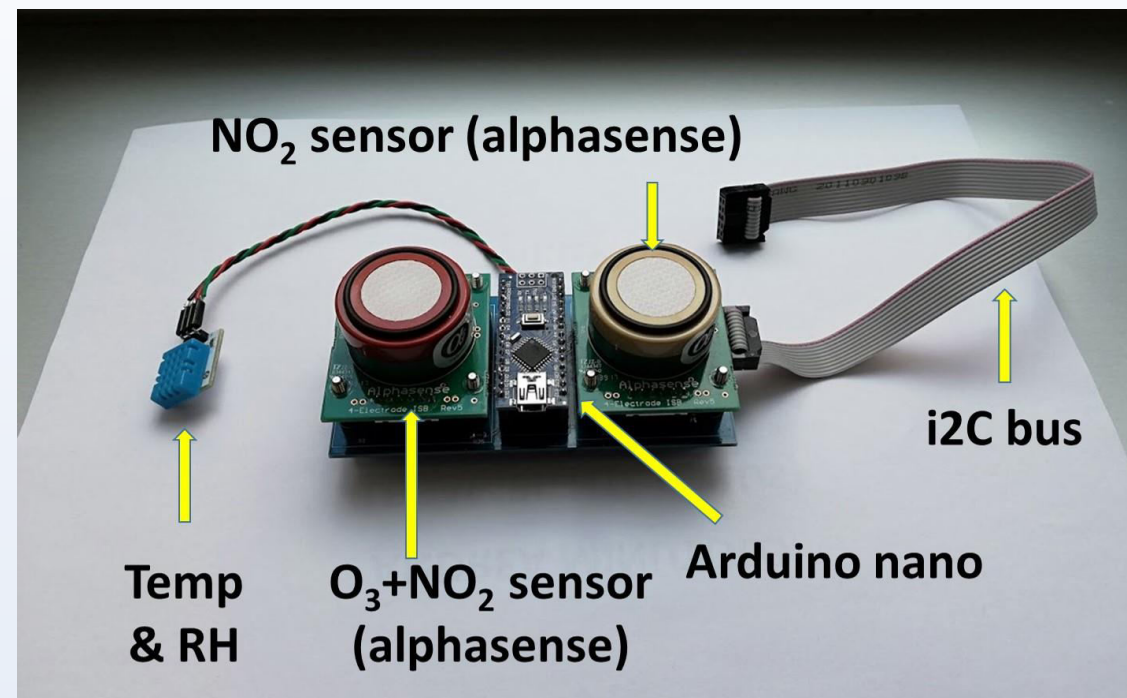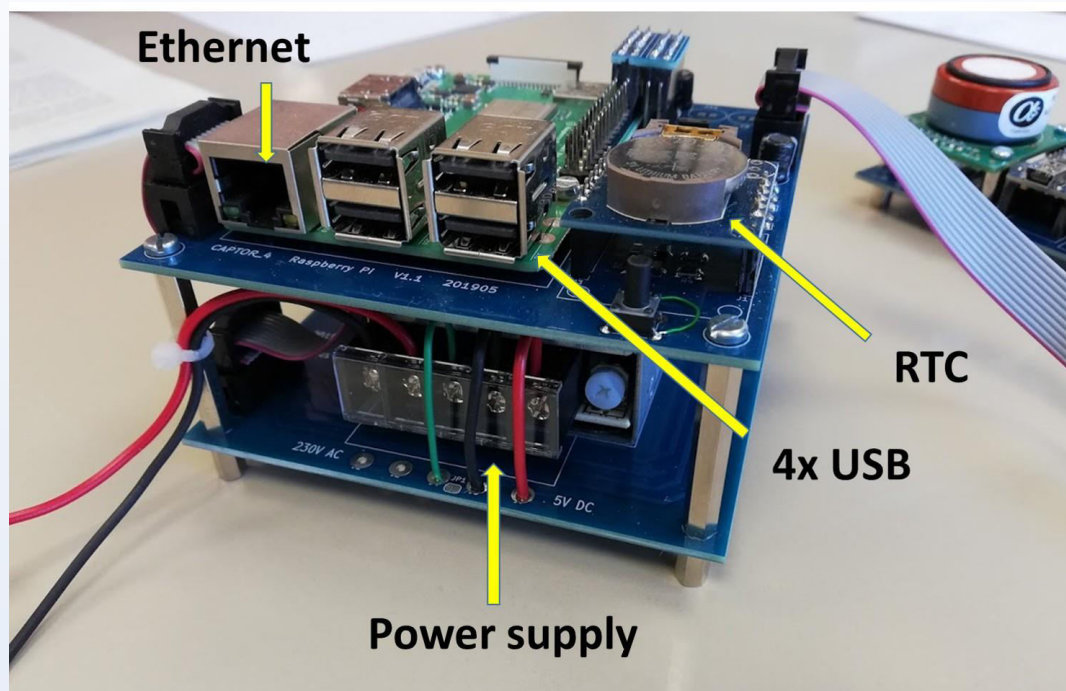
- Measures NO2, O3, wind (direction and speed)
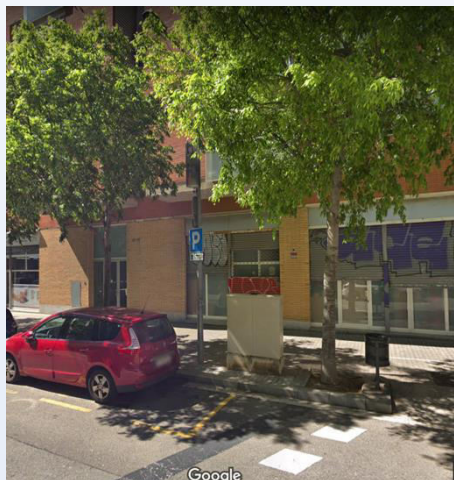
Vectorial components of wind

Wind speed

Current research and developments

- **New CAPTOR node:**
  - Alphasense electro-chemical (NO_2, O_3, NO), PMx, temp, RH,
  - Both fixed and mobile version

Current research and developments

- **State of art in-situ calibration against reference stations using machine learning techniques (MLR, SVR, RF, KNN)**

- *H2020 project CAPTOR dataset: Raw data collected by low-cost MOX ozone sensors in a real air pollution monitoring network"*, Data in Brief, 2021

- *"Multi-sensor data fusion calibration in IoT air pollution platforms"*, IEEE IoT J., 2020.

- *"A comparative Study of Calibration Methods for Low-Cost Ozone Sensors in IoT Platforms"*, IEEE IoT J., 2019.

- "Distributed Multi-Scale Calibration of Low-Cost Ozone Sensors in Wireless Sensor Networks", Sensors, 2019.

- "Self-calibration methods for uncontrolled environments in sensor networks: A reference survey", Ad Hoc Networks, 2019.

- *"Testing the performance of sensors for ozone pollution monitoring in a citizen science approach"*, STE, 2019.
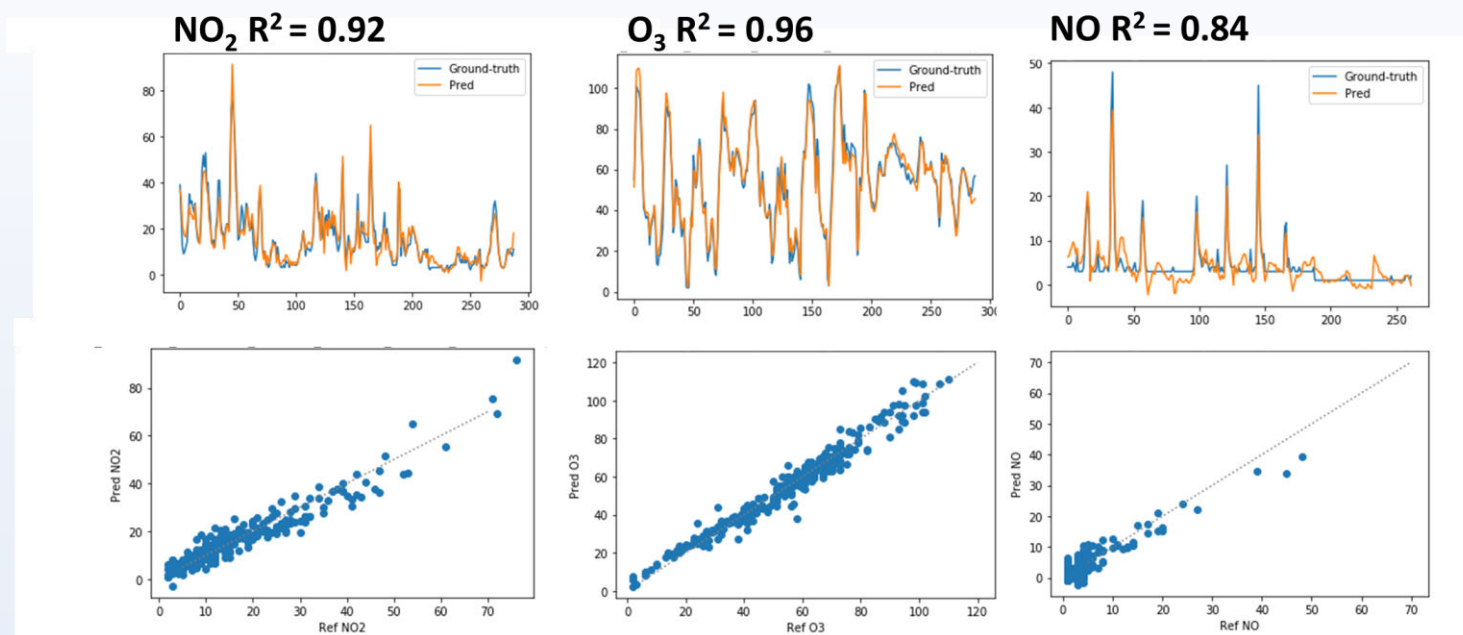
Current research and developments

- **Use of graph techniques:**
  - Graph Signal Processing (GSP) applied to heterogeneous networks of sensors and reference stations
  - This inferred graph topology provides a baseline for applying a wide range of graph-based tools: outlier detection, clustering, filtering, filling gaps of missing data, etc.

"Graph Learning Techniques Using Structured Data for IoT Air Pollution Monitoring Platforms", IEEE IoT J., 2021.
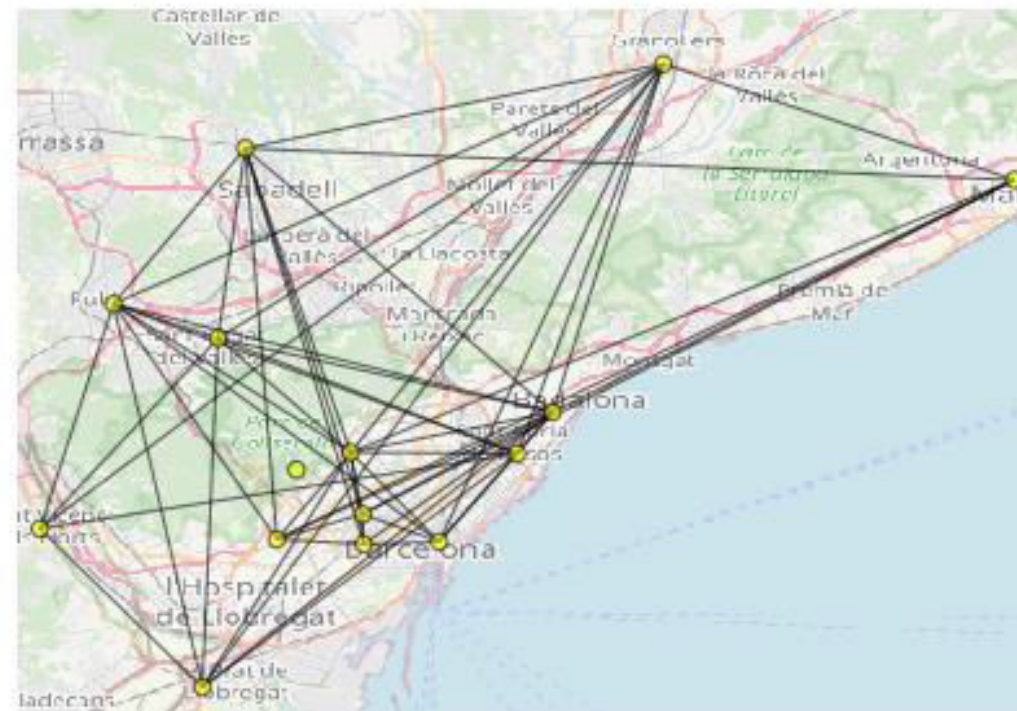
Automatically learned graph that interconnects air quality stations in Barcelona city and metropolitan area. These graphs can be used for applying a wide range of tools to improve data quality.

THANKS !!!

@nanosenaqm

www.nanosenaqm.eu