

## Background:

An Italian research group, De Vito et al. (2008), collaborated with Pirelli Labs to study the performance of low-cost sensors in monitoring urban pollution.

At the time, fixed stations mostly used industrial spectrometers to precisely estimate the concentration of air pollutants. Such stations were costly, difficult to deploy, and limited the density of measurement networks. Weather such as wind or rain could influence the movement of air. As such, fixed station measurements could misevaluate air pollutants present in a chaotic city environment.

The research team and Pirelli Labs developed a lower-cost, portable, multi-sensor device to improve the density of air pollutant measurement networks. The device weighed 2.5kg and was at the maximum 1 foot in height, width, and length. The device could sense, record, and communicate chemical concentrations through GSM. However, the device was limited by stability and selectivity issues. As such, a study was conducted to understand sensor performance over a 1-year period (March 2004 – April 2005) relative to fixed stations.

The UCI multivariate time-series dataset “Air Quality” () contains sensor readings of the low-cost sensor as well as readings from a reference certified analyzer.

The concentration measures for Carbon Monoxide and Non-Metanic HydroCarbons contain many negative values in the reference sensor. First, the data was been cleaned to replace negative numbers with NaN (Not a number). Then forward fill was applied to replace NaN with last known positive value. The replacement had a limit of 7 consecutive readings with NaN. As such there are large gaps in the graphs for Carbon Monoxide. Furthermore, the graph of Non-Metanic HydroCarbons only exists from March 2004 – May 2004 for the reference sensor.

Sensor behavior is seen to have contrasting changes after November 2004. The reference sensors indicate a rise in measure for the concentrations of Nitrogen Oxides and Nitrogen Dioxides. In contrast, low-cost sensors reflect a decrease in measure of concentration for the same chemicals. One possibility is the influence of an external factor upon the way in which concentration is measured for Nitrogen.

Overall, low-cost sensor data is seen to offer reliable readings with occasional outliers. The outliers can be factored out of data models or indicate the need to recalibrate a few times a year. The sensors would improve the density of urban pollution monitoring meshes but should be used with appropriate measures to mitigate the influence of external factors.

## Reference:

De Vito, S., Massera, E., Piga, M., Martinotto, L., & Di Francia, G. (2008). On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario. *Sensors and Actuators B: Chemical*, 129(2), 750-757. doi:<https://doi.org/10.1016/j.snb.2007.09.060>