• **Filename:** the name of this file indicates that this PDF contains the documentation information of environmental air data (EAD) captured at a CAPTOR node identified as Node-ID 22, deployed in the Spanish testbed in the 2018 campaign.

• **Project information:** CAPTOR project is funded by the European Union's Horizon 2020 Programme under the Grant Agreement No. 688110, www.captor-project.eu

• **Contact of the responsible of the dataset:**

- CSIC: Mar Viana, mar.viana@idaea.csic.es, tf:+34 93 400 61 26
- UPC: Jose M. Barcelo-Ordinas, joseb@ac.upc.edu, tf: +34 93 405 40 51

• **Description of the content of the data set:** the data set contents ozone calibrated data of a captor node located in a reference station during the 2018 summer campaign.

• **Testbed location description:** The testbed was found in the Barcelonès-Vallesos-Osona (Catalonia, Spain) a rural/suburban region in NE Spain located 60-70 km from a major urban area (Barcelona).

• **Node ID:** 18022

• **Location of the node:** Tona (Osona)

• **Technical information of the node:** specific information of the node:

- number and type of sensors in the CAPTOR node: 3 SGX Sensortech MICS 2614 metal-oxide O3 sensors, 1 Temperature/Relative humidity sensor (DHT1) Grove - Temp&Humi sensor,
- Information of calibration: the calibration of this node was performed applying a Multiple Linear Regression model to a training and validation set. The model consists of:

  $y \sim b0 + b1\,x1 + b2\,x2 + b3\,x3$, where y is the reference station data, x1 the values measured by the O3 sensor, x2 the values measured by the temperature sensor and x3 the values measured by the relative humidity sensor.

  The calibration coefficients {b0,b1,b2,b3} were calculated during a calibration period (14/06-01/07/2018) in the reference station of Tona. Additional information on the procedure to calibrate the node can be found in CAPTOR deliverable 2.3. "Software tool for Ozone Concentration Estimation Development",

- Root Mean Square Error (RMSE) obtained in the calibrated process defined as the square of the Mean Square Error of the testing data set with respect the reference values. RMSE = 9.69 ($\mu g/m^3$)
  R^2 (Coefficient of Determination) measures the proportion of variability in Y that can be explained using X and it is bound between 0 and 1. When R^2 is close to 1 it indicates that a large proportion of the variability in the response has been explained by the regression. R^2 = 0.93

• **Dates of the campaign:** this node was in a reference station from 14/06/2018 to 09/08/2018

• **Duration of the campaign:** 57 days

• **Length of the data size:** 2666 samples (rows) of the data set

• **Description of the data set file:** the data set file DataSet_Spain_EAD_Node-ID22_2018.csv contains (columns): Start_local (local date and time at which the sensors started monitoring), End_local (local date and time at which the sensors ended monitoring), and Ozone (averaged ozone concentration in $\mu g/m^3$); (rows): samples.

• **Additional information (optional):** none