



CAPTOR

Collective Awareness Platform for Tropospheric Ozone Pollution

Work package	WP3
Deliverable number	D3.6
Deliverable title	Report on data validation (b).M24
Deliverable type	R (Document report)
Dissemination level	PU (Public)
Estimated delivery date	M24
Actual delivery date	18/01/2018
Actual delivery date after EC review	dd/mm/yyyy
Version	1.0
Comments	Final version

Authors

Mar Viana, Anna Ripoll, Marc Padrosa, Xavier Querol (CSIC)
Jorge García-Vidal, J.M. Barcelo-Ordinas (UPC)

Document History			
Version	Date	Contributor	Comments
V0	15/01/2018	Mar Viana, Anna Ripoll, Marc Padrosa (CSIC)	First draft
V0.1	15/01/2018	Kun Mean Hou	Peer review
V1.0	18/01/2018	Mar Viana	Final version
V2.0	dd/mm/yyyy		Final version after EC review

Table of Contents

LIST OF ABBREVIATIONS.....	5
EXECUTIVE SUMMARY	6
1. INTRODUCTION	7
2. RESULTS.....	7
2.1 DATA AVAILABILITY	7
2.2 SENSORS AND CALIBRATION OF NODES	7
2.3 LONG-TERM DEPLOYMENT OF CAPTOR AND RAPTOR NODES AT REFERENCE STATIONS	14
2.4 DATA PROCESSING AND CALCULATION OF BETA COEFFICIENTS TO OBTAIN THE FINAL OZONE CONCENTRATIONS 17	
2.5 QUANTIFICATION OF FINAL OZONE CONCENTRATIONS FOR CAPTOR AND RAPTOR NODES	24
3. SUMMARY AND CONCLUSIONS	34
4. ANNEX	36

List of Figures

Figure 1. Example: the best sensor of a CAPTOR node did not change.	10
Figure 2. Example: the best sensor of a CAPTOR changed.	11
Figure 3. Example: the scale of the signal of certain sensors changed.	12
Figure 4. Example: the signal of a RAPTOR sensor did not change.	14
Figure 5. Example: the signal of a RAPTOR sensor changed.	14
Figure 6. Comparison between CAPTOR raw data and ozone reference station data.	15
Figure 7. Comparison between RAPTOR raw data and ozone reference station data: Monte Cucco (Italy, left; July-September 2017), Palau Reial (Spain; centre; May-June 2017), and Tona (Spain; right; June-September 2017).	16
Figure 8. Comparison between calculated ozone concentrations from a CAPTOR node using different beta coefficients, for a node where the best performing sensor was stable over time.	20
Figure 9. Comparison between calculated ozone concentrations from a CAPTOR node using different beta coefficients, for a node where the best performing sensor was not stable over time.	21
Figure 10. Comparison between reference ozone concentrations and those calculated from a RAPTOR node using beta coefficients from the Calibration 1 and Calibration 2 periods.	23
Figure 11. Comparison between reference ozone concentrations and those calculated from a RAPTOR node using beta coefficients from the Calibration 1 and Calibration 2 periods. The RAPTOR node was deployed at a reference station.	24
Figure 12. Comparison between reference ozone concentrations and those calculated with CAPTOR nodes at reference stations during the entire summer campaign.	25

List of Tables

Table 1. Number of CAPTOR and RAPTOR nodes deployed in each testbed in 2017.	7
Table 2. Root mean square errors (RMSEs, in $\mu\text{g}/\text{m}^3$) calculated for each CAPTOR sensor and node for the 2017 calibration periods (Cal1 and Cal2). Numpt: number of datapoints. The lowest RMSE/node and calibration is highlighted in red.	19
Table 3. Root mean square errors (RMSEs, in $\mu\text{g}/\text{m}^3$) calculated for each RAPTOR node for the 2017 calibration periods (Cal1 and Cal2). Numpt: number of datapoints.	22
Table 4. Environmental assessment of ozone concentrations monitored with the CAPTOR and RAPTOR nodes in Spain. Comparison with EU limit and WHO guideline values.....	30
Table 5. Environmental assessment of ozone concentrations monitored with the CAPTOR and RAPTOR nodes in Italy. Comparison with EU limit and WHO guideline values.....	31
Table 6. Environmental assessment of ozone concentrations monitored with the CAPTOR and RAPTOR nodes in Austria. Comparison with EU limit and WHO guideline values.	33

List of Abbreviations

RMSE Root Mean Square Error

Executive Summary

Description of the work

The aim of deliverable D3.3b is to report on data validation during the summer 2017 monitoring campaign. This deliverable presents a review of the CAPTOR and RAPTOR nodes deployed during the campaign and the data they reported, focusing on data quality and validation. It includes also the first quantitative results on ozone concentrations of the project for the Austria and Italy testbeds, and the second campaign results for the Spanish testbed (after the 2016 summer campaign). D3.2b reports on data generation.

Objectives

The main objective of the deliverable is to present ozone concentration time series for all of the locations possible in the Spanish, Austrian and Italian testbeds. To this end, the calibration procedure and the methodology applied for the selection of the most adequate calibration coefficients are discussed.

1. Introduction

The previous deliverable D3.3a, submitted to the European Commission, presented the results from the summer ozone monitoring campaign in 2016, which was the first to be carried out in the framework of CAPTOR. As shown in D3.3a, the results from the 2016 campaign were not as positive as expected and this was mainly due to technical failures linked to the hardware and software of the CAPTOR nodes. In addition, during the 2016 campaign only CAPTOR (and not RAPTOR) nodes were systematically deployed, and data were only collected in Spain (and not in Italy and Austria).

The present deliverable summarises the results obtained during the 2017 campaign, which improved significantly with regard to 2016. Because data availability was improved (see D3.2b), new tools for data processing and node calibration were implemented and are described in the sections below.

2. Results

2.1 Data availability

Data availability is discussed in detail in D3.2b. A brief summary is provided in Table 1 to aid in the interpretation of the results in the sections below:

Table 1. Number of CAPTOR and RAPTOR nodes deployed in each testbed in 2017.

Nr. nodes	CAPTOR	RAPTOR
Spain	25	1
Italy	10	10
Austria	0	15
Total	35	26

The performance of the CAPTOR and RAPTOR nodes with regard to data generation may be described as follows:

- 22/35 CAPTORS and 16/26 RAPTORS reported full datasets for the entire period (calibrations and campaign) and the data are reliable.
- 10/35 CAPTORS reported data during calibrations and the campaign, but at some point the scale of the sensor signal changed (what we described as “re-basing” in D3.3a) and data from this moment on are not reliable.
- 1/35 CAPTOR and 2/26 RAPTORS reported data during calibrations and the campaign but the data are not reliable.
- 1/35 CAPTOR reported data during calibrations and the campaign but the temperature and humidity failed and it could not be calibrated.
- 1/35 CAPTOR and 1/26 RAPTOR did not submit any data.
- 7/26 RAPTORS reported data, but the datasets are incomplete.

2.2 Sensors and calibration of nodes

CAPTOR nodes

In 2016 the CAPTOR nodes were built with temperature and relative humidity sensors and 3 ozone

sensors (for the nodes at hosts' homes) or 5 ozone sensors (for nodes at reference stations) (all of them are metal-oxide sensors). In 2017 this setup was modified to include 5 ozone sensors in all nodes (hosts and reference stations), in addition to temperature and humidity:

- 2 MIC sensors reused from the 2016 campaign (referred to as s1 and s2)
- 2 brand new MIC sensors (s3 and s4)
- 1 brand new MQ sensor (s5)

MIC sensors from the previous campaign had to be reused due to the discontinuation in their production by the manufacturer. This is also the reason for the need to test a new type of sensor (MQ). The data from the MQ sensors in the 2017 campaign require additional processing and are not reported in this deliverable.

The calibration was also improved in 2017 with regard to 2016, by using a 2-step calibration procedure:

- Step 1 – Calibration at a central reference station: an initial calibration was carried out at the Palau Reial reference station, located in Barcelona at the CSIC premises. This step was used as an initial test for the sensors, where not-working units were identified and exchanged for new ones. It also allowed us to obtain a first comparison with reference ozone concentrations, which favoured the detection of under-performing sensors, at a location which was easily accessible.
- Step 2 – Calibration at a local reference station: after validation in Step 1, the nodes were split into groups and transferred to reference stations in the vicinity of their assigned host location. The nodes were calibrated again at these reference stations for 2 weeks. This second step allowed us to calibrate the nodes while exposed to ambient ozone concentrations closer to those they were intended to monitor (as opposed to the lower concentrations registered in Barcelona). Unfortunately, it was only possible to carry out this second step in Spain and Italy, as in Austria reference stations in the host areas were not available.

As in 2016, the CAPTOR nodes were calibrated before and after the monitoring campaign at the local reference station (Step 2), and the results are referred to as Calibration 1 and Calibration 2, respectively.

Different behaviours were identified for the different nodes resulting from the variability of the individual sensors. These behaviours are described in the following Figures, where sensor data are always presented in arbitrary electrical resistance units (raw data) and not yet in terms of mass concentration ($\mu\text{g}/\text{m}^3$). It is essential to understand this, given that the subsequent conversion to mass concentrations is described after page 24.

In 7 of the CAPTOR nodes (**Figure 1**) the best sensor did not change between Calibrations 1 and 2. This implies that the performance of the best sensor and also of the remaining sensors did not change significantly, and that the best performing sensor prior to the campaign was also the best sensor at the end of the summer. As shown in **Figure 1**, when plotting the individual sensor vs. reference concentrations the highest R^2 value was consistently obtained for s3 (highlighted in red in the Figure). It is interesting to note the different order of magnitude of the signal generated by the different sensors, e.g., s3 reported values 1 order of magnitude higher than the rest of the sensors, despite sourcing from the same batch as them (s1, s2, s4) and being previously unused just like s4

(s1 and s2 had been used in the 2016 campaign).

A similar behaviour was observed in 15 of the CAPTOR nodes, where best sensors were identified with high R^2 values (>0.75) but where the best sensor changed between the two calibration periods. This is the case of the example shown in **Figure 2**. As in the case above, the different sensors had different responses (see for example s2 and s3 vs. s4 and even s1, during Calibration 1), and while s2 was the best performing sensor during Calibration 1 its performance clearly decreased towards Calibration 2, when s4 was the best sensor. This behaviour was unexpected given that the decreasing performance of a sensor (e.g., s2) could be reasonable, but not so much an improvement in performance (e.g., s4). This kind of response raised the question of drifts in sensor performance, and has implications with regard to the selection of the sensor to be used to obtain the ozone time series during the campaign, when the nodes were at hosts' homes and no reference data was available for comparison. In these cases, the final selection of the beta coefficient to be applied was made by the combined assessment of the RMSE, the R^2 value and the expert's judgement, and assuming a certain degree of uncertainty.

Finally, **Figure 3** shows an example of the behaviour observed for 10 of the CAPTOR nodes, where a change in scale was detected at a given point during the monitoring campaign (in this example, after 26/07/2017). This kind of changes in scale, or re-basing of the time series (with a new lower limit), was detected mainly after a break in the time series, as in the example, but also in absence of such a break. In addition, on some occasions it was quite evident, as in the example, but occasionally it was only very subtle and difficult to identify. It might even be related to ageing of the sensors. Because of the change in the scale or the lower limit of the signal generated by the sensors it became unreasonable to use a single beta coefficient to correct the time series before and after the change. Therefore, for the CAPTORS where this incidence occurred the final ozone time series calculated do not cover the entire monitoring period, and instead are only presented for the period for which a reliable beta coefficient can be applied.

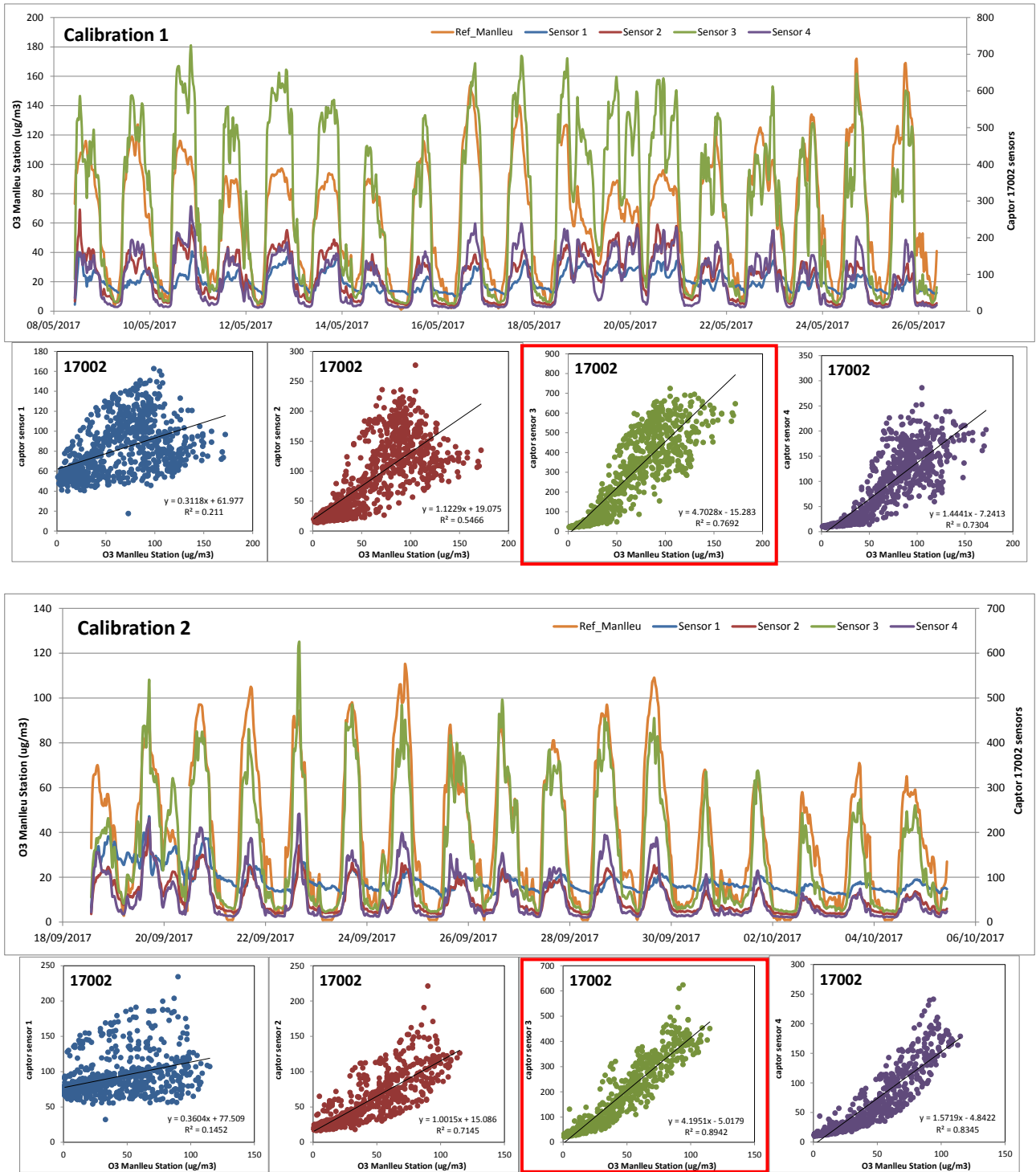


Figure 1. Example: the best sensor of a CAPTOR node did not change.

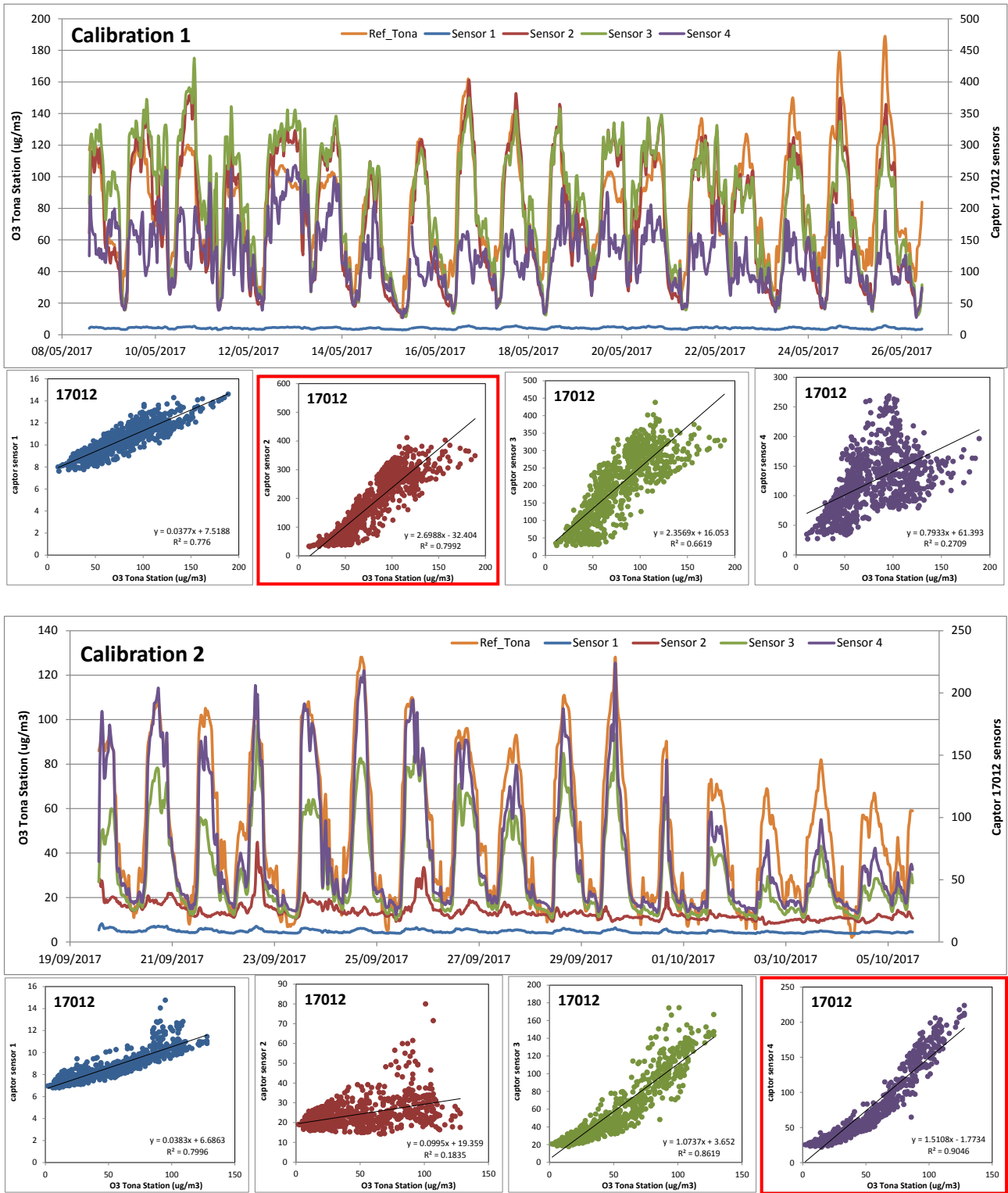


Figure 2. Example: the best sensor of a CAPTOR node changed.

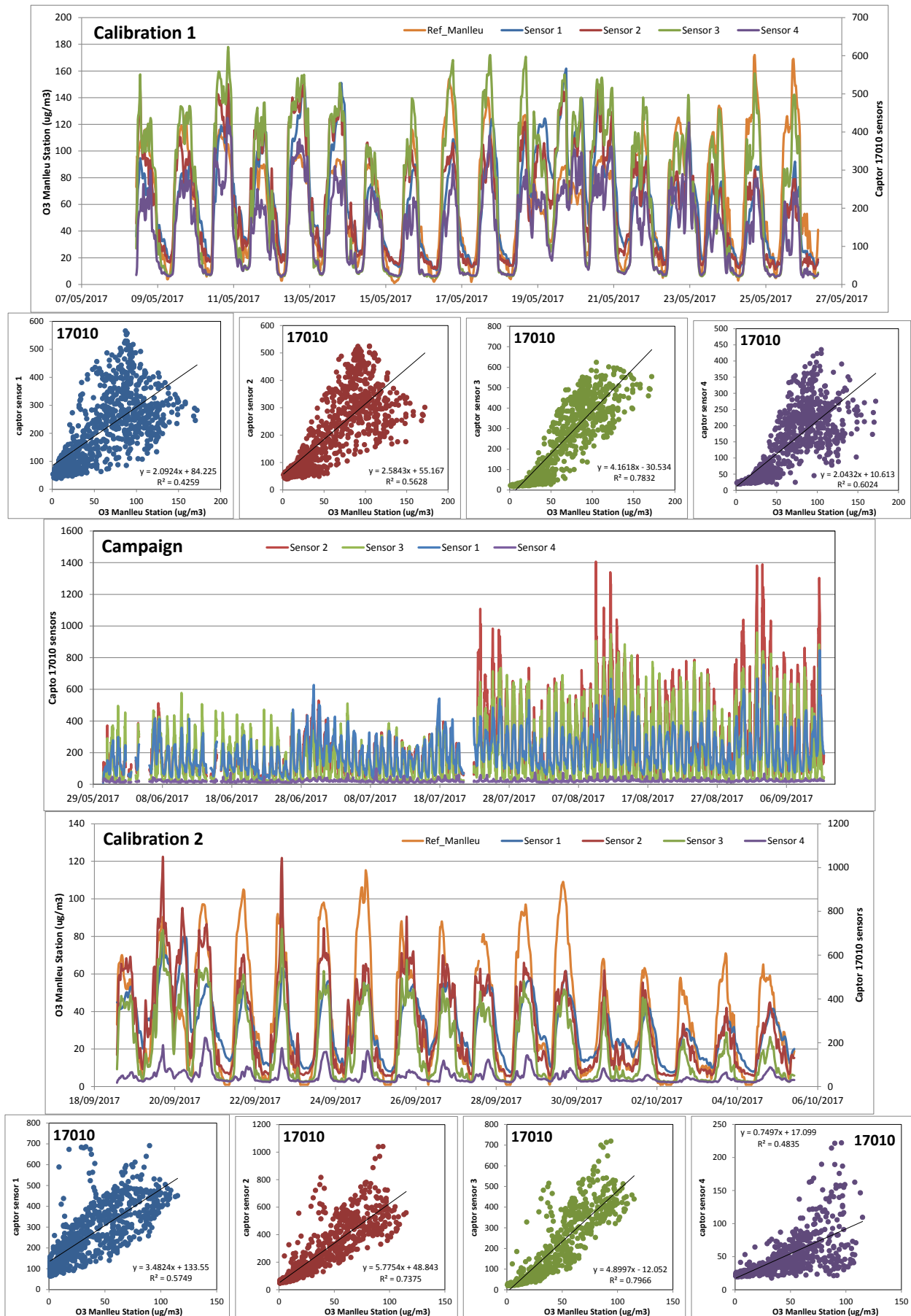


Figure 3. Example: the scale of the signal of certain sensors changed.

RAPTOR nodes

The RAPTOR nodes included one ozone, one NO₂, one temperature and one relative humidity sensor (all of them, electrochemical). The calibration of RAPTORS followed the same procedure as with the CAPTORS, i.e., the 2-step process described above. However, not all RAPTORS could undergo calibration after the campaign (Calibration 2) due to difficulties in access to the reference stations especially in Austria. Because the RAPTORS only have 1 ozone sensor there was no need to search for the best performing sensor, and instead the data processing focused on assessing whether the performance had improved/decreased throughout the monitoring period (before and after the campaign).

In total 12 of the RAPTORS seemed to show a similar performance before and after the campaign (**Figure 4**). As shown in the Figure, in these cases the comparison between the sensor and reference data provided similar slopes and R² values, suggesting the absence of a drift during the period of time assessed.

Conversely, on 4 occasions the RAPTORS did show a different performance before and after the monitoring campaign, as shown in **Figure 5**. It becomes evident from the time series that after 08/09/2017 the comparability between the sensor and reference data decreased, aside from the expected decrease in ambient ozone concentrations reported by the reference instrument. As in the case of the CAPTORS, it is worth noting that this change in the sensor's signal followed a brief break in the time series, suggesting that a technical issue (e.g., power shortage) may have occurred and caused the performance to decrease. However, also as in the case of the CAPTORS, these changes in response at times do not originate from technical issues and they may be subtle and difficult to identify. They could even be linked to ageing of the sensors. Further research is necessary to fully understand this behaviour.

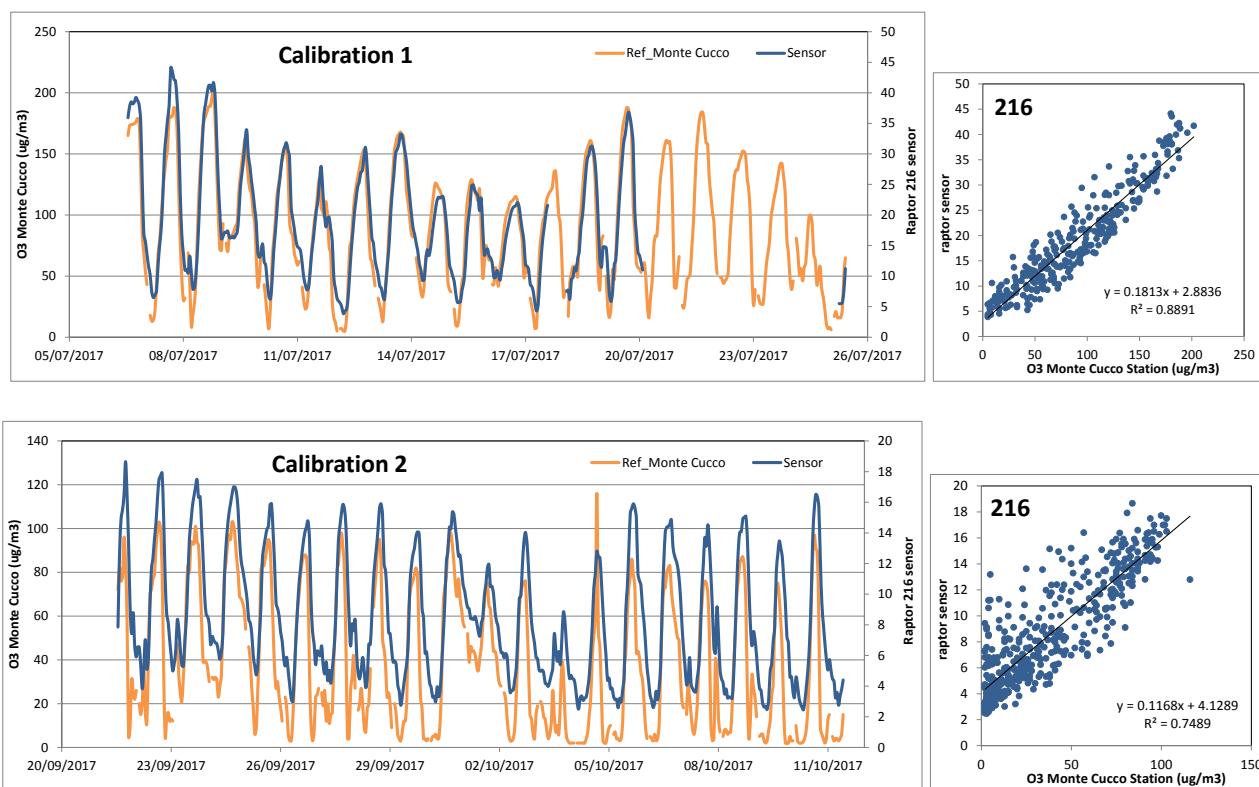


Figure 4. Example: the signal of a RAPTOR sensor did not change.

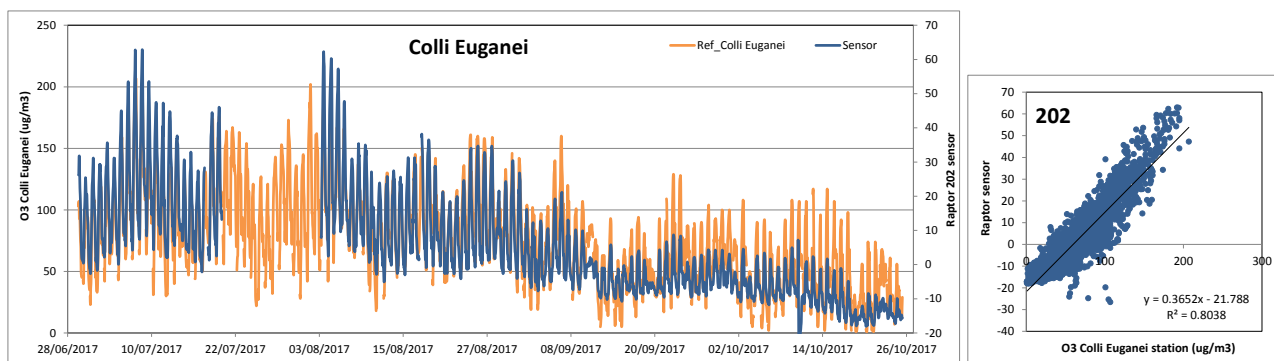


Figure 5. Example: the signal of a RAPTOR sensor changed.

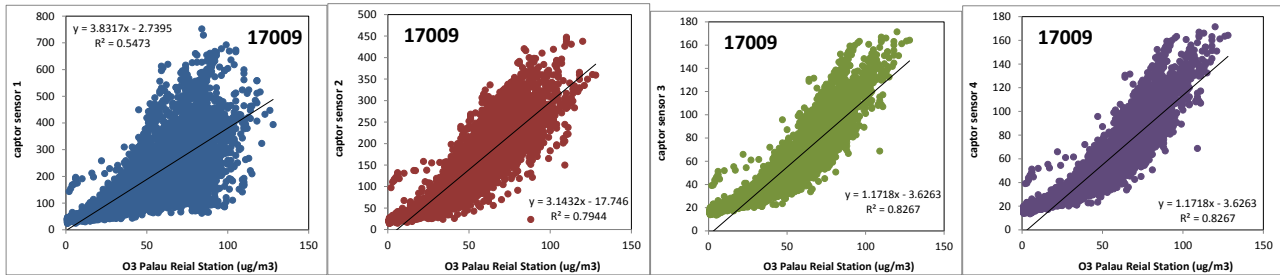
2.3 Long-term deployment of CAPTOR and RAPTOR nodes at reference stations

With the aim to validate the long-term performance of both types of nodes, 8 CAPTORS and 3 RAPTORS were deployed at reference stations in Spain and Italy during the entire duration of the monitoring campaign (in addition to during the calibration periods). The results are shown in **Figure 6** and **Figure 7**, in terms of electrical resistance and not of mass concentrations ($\mu\text{g}/\text{m}^3$), as discussed above. Different stations and even countries were selected in order to challenge the nodes with different ozone concentration levels and meteorological conditions.

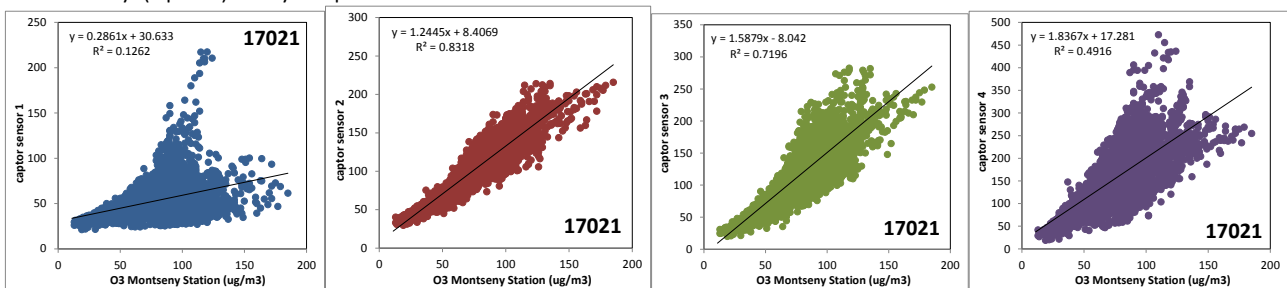
In general, CAPTOR and RAPTOR nodes were able to produce relatively long time series, with relatively few breaks due to technical failures. As shown in the Figures, the comparison with reference data varied as a function of the individual sensors. It should be noted that, as will be discussed below, final ozone concentrations for the CAPTOR nodes were calculated taking into account only the best performing sensor, and therefore for each node in **Figure 6** only the sensor

with the highest R^2 coefficient should be considered. Results evidenced that, while some of the sensors showed a relatively high data dispersion (e.g., Tona, $R^2 < 0.69$; Vic, $R^2 < 0.72$; Cuneo, $R^2 < 0.41$), higher R^2 coefficients were obtained for 5 of the 8 CAPTOR nodes tested ($R^2 > 0.75$ for Palau Reial, Montseny, Manlleu, Monte Cucco, and Osio Sotto). R^2 coefficients were always > 0.80 for the RAPTOR nodes (3 in total), reaching $R^2 = 0.95$ in the case of Tona. This kind of comparison is especially useful as it allows us to assess the comparability between sensor and reference data for longer periods of time, as well as between different sensors in a given node.

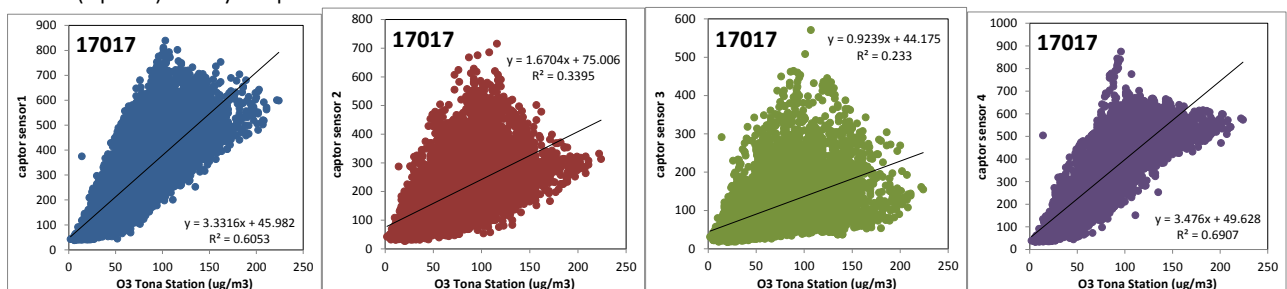
Palau Reial (Barcelona, Spain): June-October 2017



Montseny (Spain): July-September 2017



Tona (Spain): May-September 2017



Vic (Spain): May-October 2017

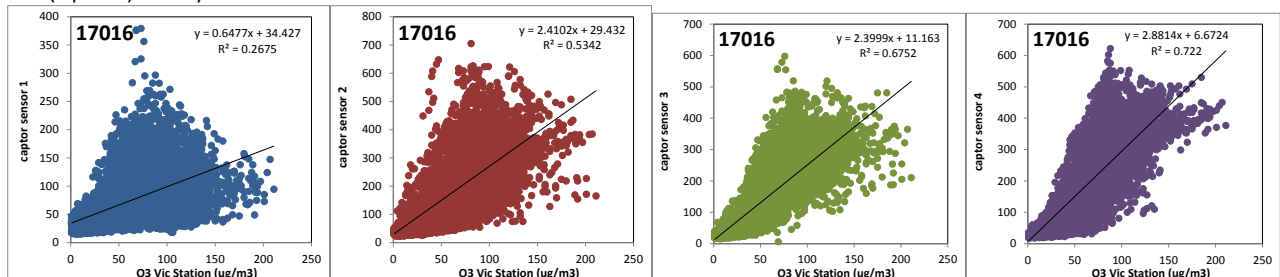
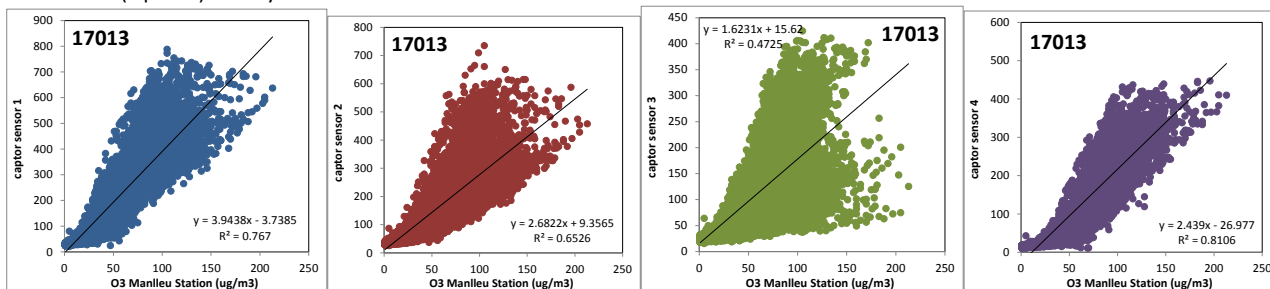
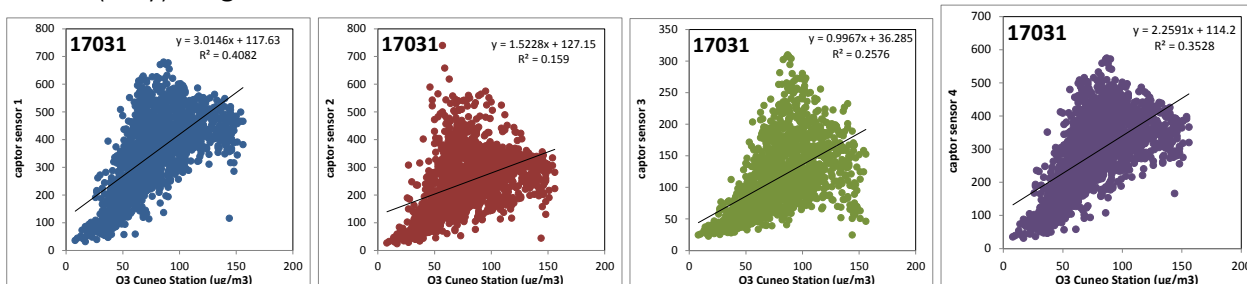


Figure 6. Comparison between CAPTOR raw data and ozone reference station data.

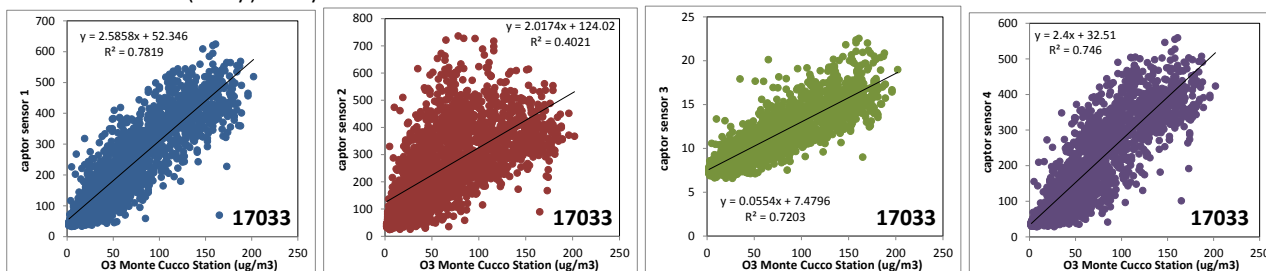
Manlleu (Spain): May-October 2017



Cuneo (Italy): August-October 2017



Monte Cucco (Italy): July-October 2017



Osio Sotto (Italy): August-October 2017

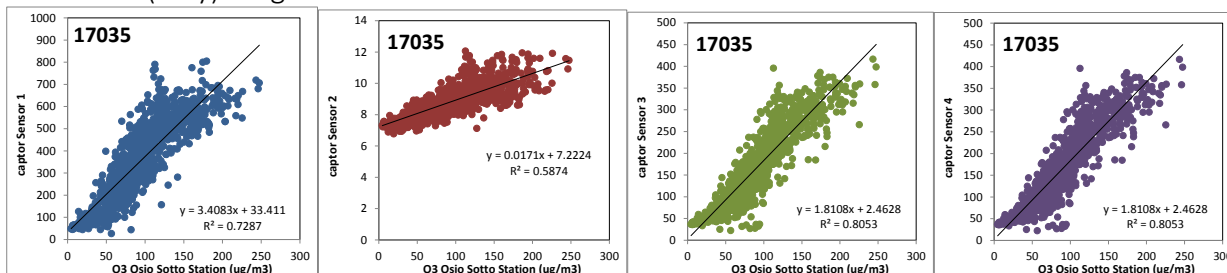


Figure 6. Continued.

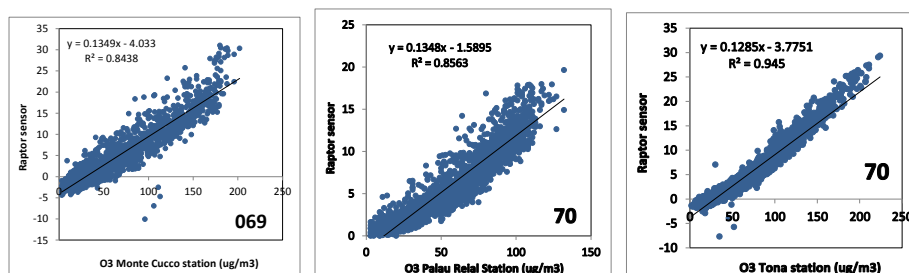


Figure 7. Comparison between RAPTOR raw data and ozone reference data: Monte Cucco (Italy, left; July-September 2017), Palau Reial (Spain; centre; May-June 2017), and Tona (Spain; right; June-September 2017).

2.4 Data processing and calculation of beta coefficients to obtain the final ozone concentrations

CAPTOR nodes

As described in the previous sections, the CAPTOR nodes were calibrated before and after the monitoring campaign at the closest reference station. After Calibration 2, the full dataset was processed for each node, i.e., the Calibration 1, campaign and Calibration 2 datasets. For each calibration dataset, the individual sensor raw data were regressed against the simultaneous reference ozone concentrations and regression coefficients (beta) were calculated for each individual sensor. This was done individually for each node in such a way that, for each node, 4 beta coefficients were obtained for Calibration 1, and 4 for Calibration 2. **The calibration algorithm is described in detail in deliverables D2.3a and D2.3b.**

These beta coefficients were then applied to each individual sensor dataset for the calibration periods, and the root mean square error (RMSE) was calculated. The RMSE is the mean difference, in absolute concentration values ($\mu\text{g}/\text{m}^3$), between the reference and sensor data for each calibration period. The 8 RMSEs calculated per node are shown in

Table 2. The sensor with the lowest RMSE per node and per calibration period was selected as the best performing sensor, for a given node.

Initially, our intention was to use a combination of the different sensor signals within each node to calculate the resulting ozone concentration time series. However, it was observed that the errors of the different sensors were correlated, and therefore applying a fusion algorithm to combine the different sensor outputs did not reduce the RMSE. As a result, a single-sensor approach was used instead.

Once the best sensor had been selected, a single beta coefficient (obtained from regressing the raw sensor data against the reference data) had to be selected to be used for the calculation of the final ozone time series for each node. Two options were then available: the beta coefficient from the Calibration 1 or the Calibration 2 period. In addition, shuffling of the input data was included as another variable in the analysis (discussed in detail in deliverable D2.9 pp 13). As a result, time series of ozone concentrations were computed applying the following beta coefficients to the raw data of the best sensor (see examples in **Figure 7** and **Figure 8**):

- Cal1A used the beta coefficients from the Calibration 1 period to the Calibration 1 dataset, without shuffle.
- Cal1B used betas from the Calibration 1 period to the Calibration 1 dataset, with shuffle.
- Cal1C used betas from the Calibration 2 period to the Calibration 1 dataset, with shuffle.
- Cal2A used the beta coefficients from the Calibration 2 period to the Calibration 2 dataset, without shuffle.
- Cal2B used betas from the Calibration 2 period to the Calibration 2 dataset, with shuffle.
- Cal2C used betas from the Calibration 1 period to the Calibration 2 dataset, with shuffle.

Finally, the time series obtained in this way were inter-compared and compared with the reference dataset in each case, and the one with the highest similarity to the reference data (during calibration) was selected as final. **The beta coefficient of this time series was then used to calculate the ozone mass concentration during the campaign period, i.e., the final ozone time series for each node (whether that node was deployed with a host or at a reference station).**

As described in previous sections, the CAPTOR nodes included 2 reused and 2 brand new MIC sensors. It is interesting to note that, in general, the best performing sensor was one of the new sensors (s3 or s4), after applying the criteria described above (lowest RMSE and highest R^2). It is probable that this is related to ageing of the older sensors (s1 and s2), although with the current data it is not possible to clearly detect or quantify drifts due to ageing. It should be noted, however, that the differences between the reused and the new sensors were in some cases only small.

Examples of the results of this methodology are shown in the Figures below, as a function of the different node behaviours described above (**Figure 1**, **Figure 2**, **Figure 3**). In the cases where the best sensor was stable over time (significantly lowest RMSE, e.g., **Figure 7**; 7 CAPTOR nodes in total), the selection between the beta coefficient from Calibration 1 or 2 was relatively arbitrary, as results usually showed only relatively minor differences. **Figure 7** shows an example of such a node, which evidences two main issues:

1. Differences between ozone concentrations calculated with the three types of beta coefficients (with and without shuffle, and calculated in either of the calibration periods) were minor, with $R^2 > 0.90$ in most of the cases. This was especially true for the high concentrations, as differences increased slightly for the lower-end concentrations. This is relevant from a mathematical perspective but not from the point of view of environmental pollution or population exposure.
2. Shuffling the data proved to be an improvement, as the similarities between the calculated time series and the reference were usually larger when shuffling had been applied.

Whenever the best performing sensor was not stable over time, but the CAPTOR nodes provided continuous and reliable data throughout the study period (which was the case for 15 of the nodes), deciding on the beta coefficient to be applied resulted to be a more complex task. In these cases, two options were possible:

- Selecting the second better performing sensor, if it was more stable than the best one, or
- Selecting the beta coefficient from one of the calibration periods, but assuming a larger uncertainty in the data. The result from this approach was usually an underestimation or overestimation of the ozone concentrations during one of the calibration periods.

An example of this behaviour is shown in **Figure 8**. In this case the best sensor for Calibration 1 was s2, while it was s4 for Calibration 2. For their respective calibrations the sensors are able to reproduce adequately the reference concentrations, but when they are applied to the other calibration dataset (e.g., Cal1Cs4) the concentrations were underestimated. Consequently, it was decided that the best approach would be to select a third sensor (s1) which produced good correlations between the calculated ozone and reference data ($R^2 > 0.91$) for both calibration periods, despite not being the best sensor in either of them. This strategy was applied to all the CAPTOR nodes where the best sensor was not stable over time.

Finally, the 10 CAPTOR nodes where changes in the scale of the signal was detected were processed as in the cases above, depending on the stability of the best sensor during the period for which the data were considered reliable.

RAPTOR nodes

Calibration was somewhat easier for the RAPTOR nodes given that they only had one ozone sensor. The datasets from both calibration periods were assessed, and the beta coefficient to be used for

the final time series were selected on the basis of their RMSEs (**Table 2**). It should be noted that, in addition to the calibration applied following the methodology described for the CAPTOR nodes, the LIMOS team produced an additional calibration based on a proprietary algorithm. For this reason, the results presented in this deliverable refer only to the calibrations carried out with the methodology described above (not using the proprietary algorithm).

Table 2. Root mean square errors (RMSEs, in $\mu\text{g}/\text{m}^3$) calculated for each CAPTOR sensor and node for the 2017 calibration periods (Cal1 and Cal2). Numpt: number of datapoints. The lowest RMSE/node and calibration is highlighted in red.

CAPTOR

Testing RMSE's (ug/m3)						
Captor ID	Calibration	Numpt	Sensor 1	Sensor 2	Sensor 3	Sensor 4
17001	Cal1	378	14.41	12.80	11.00	10.11
	Cal2	122	7.68	7.22	7.70	7.24
17002	Cal1	377	14.81	13.50	11.17	10.92
	Cal2	356	12.74	8.90	7.20	8.70
17003	Cal1	378	13.96	14.58	10.18	11.69
	Cal2	355	8.06	13.64	7.04	7.37
17004	Cal1	400	14.33	17.55	9.24	9.30
	Cal2	275	10.69	10.60	7.50	7.55
17005	Cal1	378	11.79	14.32	10.89	10.49
	Cal2	355	11.02	12.77	8.45	9.33
17006	Cal1	377	16.85	14.50	12.04	12.53
	Cal2	337	12.51	11.89	11.04	10.61
17007	Cal1	328	11.87	13.57	10.54	12.60
	Cal2	337	10.00	10.92	9.90	8.46
17009	Cal1	472	13.35	10.88	10.60	12.33
	Cal2	425	11.19	10.75	10.98	12.02
17010	Cal1	378	14.02	14.21	11.38	12.56
	Cal2	355	10.19	8.30	8.44	12.47
17011	Cal1	376	12.99	12.31	11.20	13.05
	Cal2	174	11.38	10.26	12.55	11.31
17012	Cal1	376	9.62	10.97	11.03	16.58
	Cal2	337	8.52	13.98	9.28	8.97
17013	Cal1	473	11.28	11.67	11.55	12.22
	Cal2	391	7.90	8.96	11.01	10.08
17014	Cal1	377	16.90	11.24	16.35	10.59
	Cal2	323	12.80	9.10	15.02	12.74
17015	Cal1	503	19.05	19.99	12.55	25.57
	Cal2	410	9.61	12.31	12.02	15.00
17016	Cal1	433	15.93	12.20	10.53	11.00
	Cal2	410	10.53	11.70	10.47	10.64
17017	Cal1	346	10.00	10.70	15.21	9.98
	Cal2	357	10.90	12.19	14.51	10.88
17018	Cal1	325	15.26	12.89	14.07	13.14
	Cal2	278	12.67	11.39	10.95	10.95
17019	Cal1	225	20.31	20.73	21.73	11.93
	Cal2	159	11.21	12.96	13.32	10.29
17020	Cal1	217	10.60	9.45	12.90	10.80
	Cal2	264	11.88	10.75	11.62	10.36
17021	Cal1	367	17.00	10.31	12.86	15.85
	Cal2	251	11.47	7.10	8.21	7.64
17022	Cal1	459	13.61	12.68	12.96	12.70
	Cal2	275	17.43	10.91	10.32	8.37
17023	Cal1	393	15.30	12.39	16.65	10.97
	Cal2	275	10.68	8.36	8.46	10.17
17024	Cal1	95	17.03	14.96	13.83	11.34
	Cal2	136	13.42	11.70	12.16	11.75
17025	Cal1	399	10.80	12.50	10.57	9.67
	Cal2	275	14.13	9.69	10.14	9.15
17026	Cal1	312	10.16	9.44	12.24	14.43
	Cal2	284	10.20	10.11	10.93	18.60
17027	Cal1	377	18.67	15.08	12.81	11.62
	Cal2	337	9.99	9.15	11.17	8.74
17028	Cal1	192	22.13	14.49	16.00	16.75
	Cal2	197	13.57	10.08	10.07	11.49
17029	Cal1	91	12.10	9.69	11.49	11.80
	Cal2	125	9.25	9.39	9.07	9.07
17030	Cal1	150	19.70	16.70	18.30	16.61
	Cal2	192	12.55	11.09	10.23	11.79
17031	Cal1	163	12.14	11.44	13.50	12.04
	Cal2	136	8.21	10.96	8.65	8.42
17033	Cal1	159	18.69	18.89	18.40	20.85
	Cal2	165	9.60	12.40	14.41	12.02
17034	Cal1	188	19.16	17.75	17.99	16.07
	Cal2	57	7.80	7.54	7.03	8.19
17035	Cal1	175	14.41	16.86	13.83	15.18
	Cal2	238	9.36	9.83	9.61	10.33

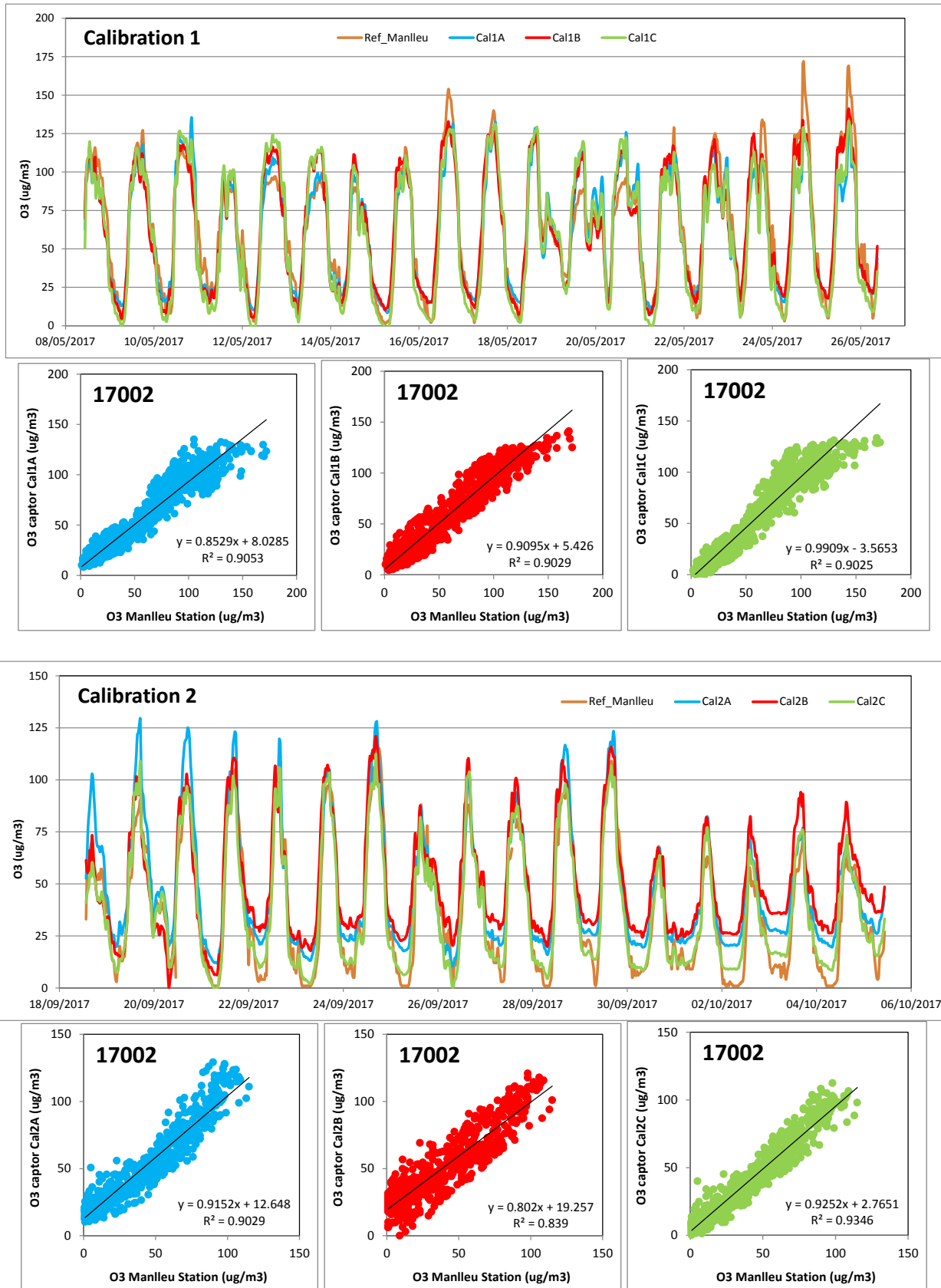


Figure 8. Comparison between calculated ozone concentrations from a CAPTOR node using different beta coefficients, for a node where the best performing sensor was stable over time.

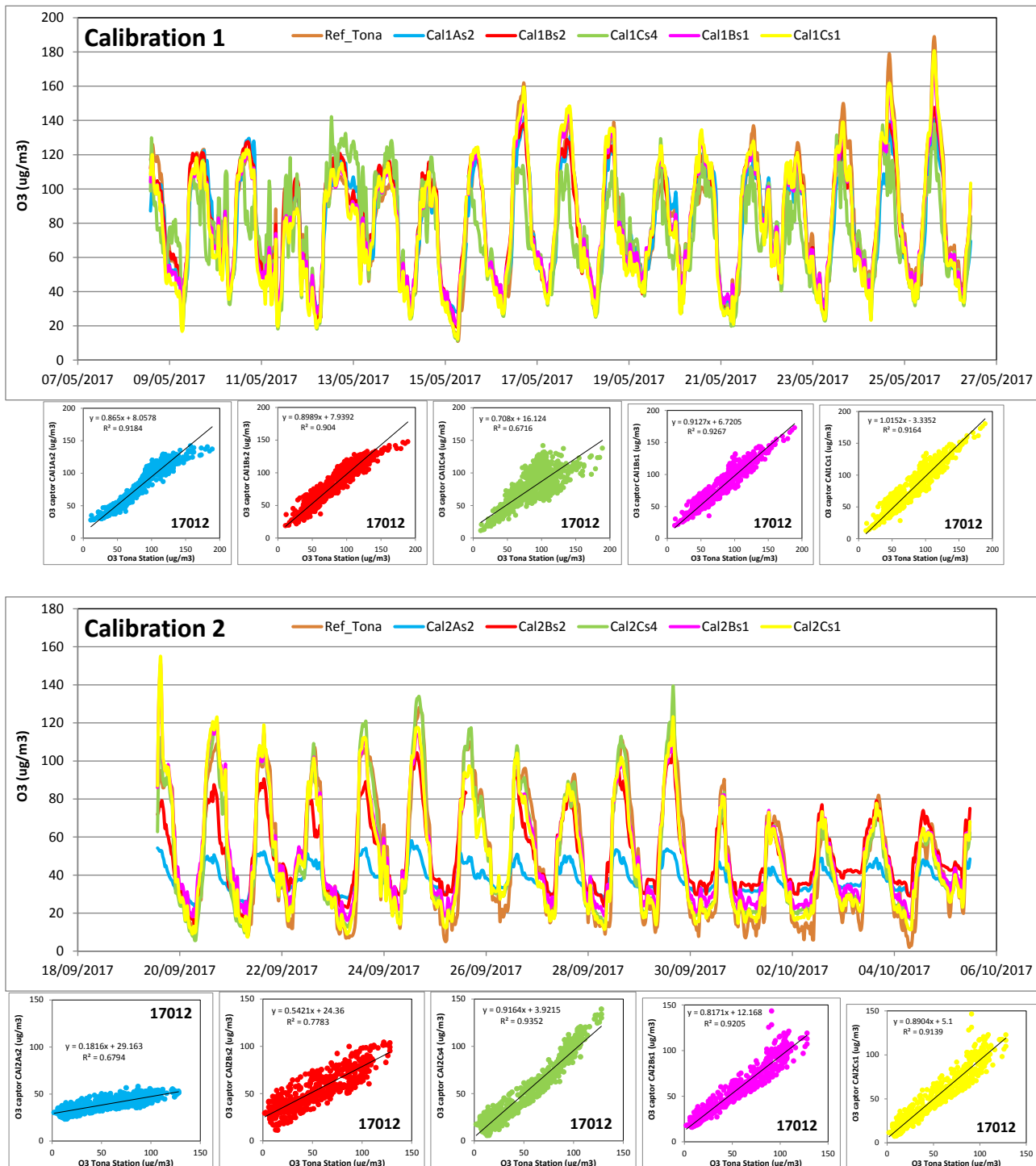


Figure 9. Comparison between calculated ozone concentrations from a CAPTOR node using different beta coefficients, for a node where the best performing sensor was not stable over time.

Table 3. Root mean square errors (RMSEs, in $\mu\text{g}/\text{m}^3$) calculated for each RAPTOR node for the 2017 calibration periods (Cal1 and Cal2). Numpt: number of datapoints.

Raptor ID	Calibration	Numpt	Testing RMSE's ($\mu\text{g}/\text{m}^3$)
69	Cal1	106	17.37
	Cal2	140	12.86
70_PR	Cal1	243	6.50
	Cal2	242	9.91
70_Tona	Cal1	229	5.82
	Cal2	260	4.65
71	Cal1	286	21.82
	Cal2	153	6.04
72	Cal1	293	21.61
	Cal2	210	15.65
73	Cal1	293	18.84
	Cal2	153	8.32
75	Cal1		
	Cal2	210	16.76
89	Cal1	292	23.98
	Cal2	210	16.11
90	Cal1	293	20.38
	Cal2	197	12.66
202	Cal1	153	12.14
	Cal2	186	13.29
204	Cal1	67	15.71
	Cal2	108	12.23
206	Cal1	114	9.39
	Cal2	145	10.52
208	Cal1	114	9.92
	Cal2	145	10.44
210	Cal1	51	13.95
	Cal2		
212	Cal1	93	11.97
	Cal2	159	10.51
214	Cal1	61	9.60
	Cal2		
216	Cal1	103	17.00
	Cal2	153	13.14
218	Cal1	61	9.33
	Cal2	160	7.56
302	Cal1	293	19.25
	Cal2	210	16.69
304	Cal1		
	Cal2	153	9.20
306	Cal1	175	19.39
	Cal2	210	10.64
308	Cal1	293	19.25
	Cal2		
310	Cal1	293	20.31
	Cal2		
312	Cal1	293	24.15
	Cal2	210	16.11
314	Cal1	293	22.59
	Cal2	210	13.77
316	Cal1	290	18.43
	Cal2		

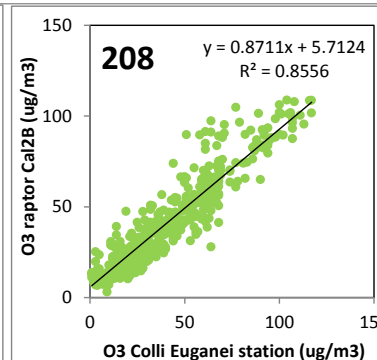
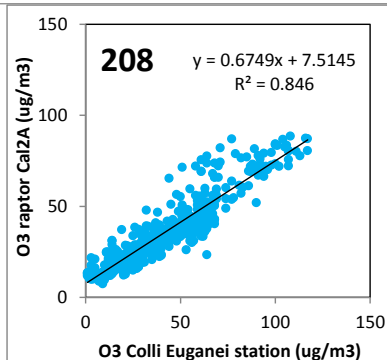
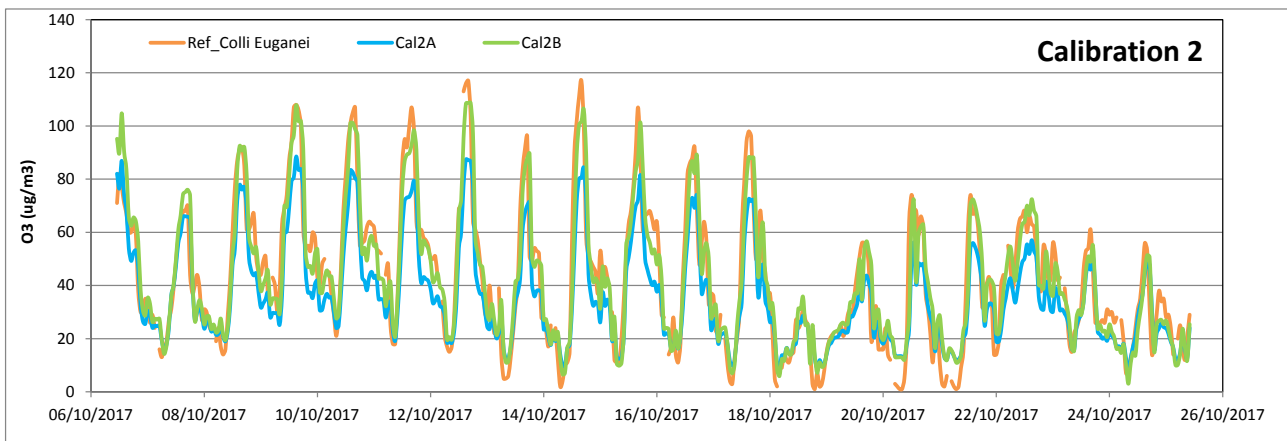
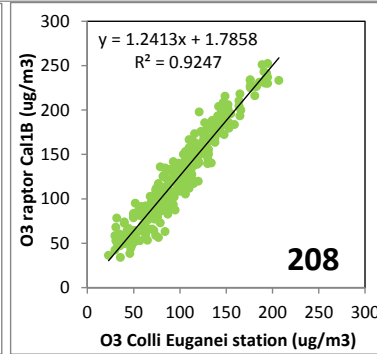
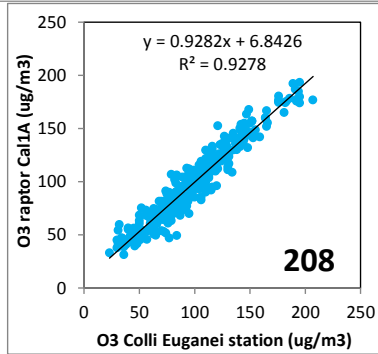
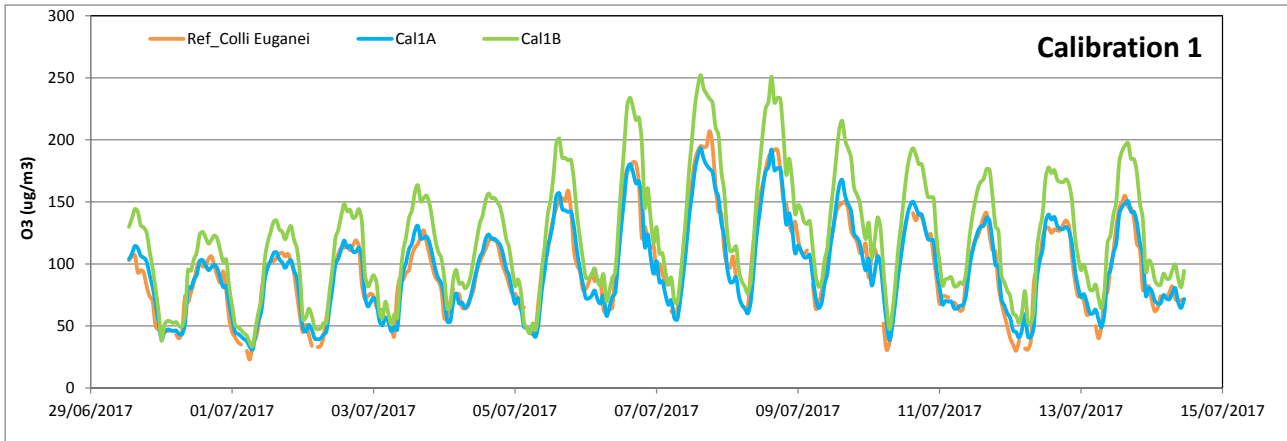


Figure 10. Comparison between reference ozone concentrations and those calculated from a RAPTOR node using beta coefficients from the Calibration 1 and Calibration 2 periods.

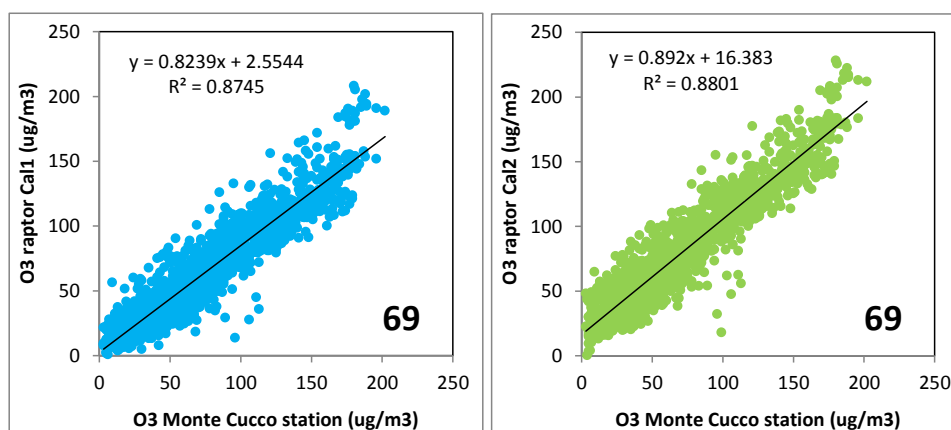
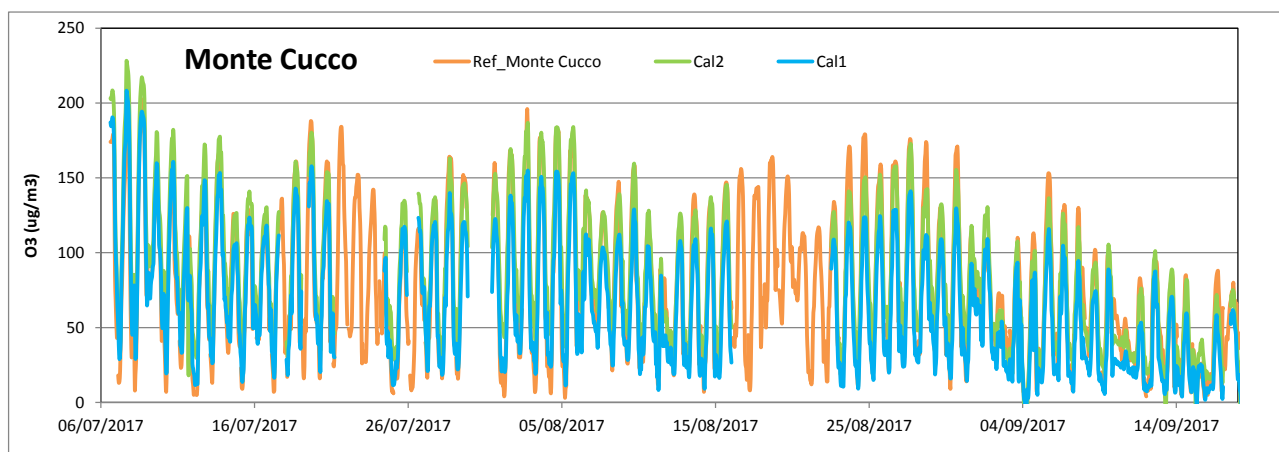


Figure 11. Comparison between reference ozone concentrations and those calculated from a RAPTOR node using beta coefficients from the Calibration 1 and Calibration 2 periods. The RAPTOR node was deployed at a reference station.

Figure 9 shows an example of the calibration results for one of the RAPTOR nodes. As in Figure 8, the sensor was not stable over time: the beta from the first calibration (CalA) was able to reproduce accurately the reference ozone concentrations during Calibration 1 (Cal1A, $R^2=0.936$; $y=0.9282x+6.8426$), but when it was applied to the second calibration period this sensor underestimated the reference concentrations (Cal2A, $R^2=0.846$, $y=0.6749x+7.5145$). The opposite was true for Cal1B and Cal2B.

Finally, Figure 10 shows an example of a RAPTOR node deployed at a reference station in Italy. Irrespective of the beta coefficient used, results evidence a relatively high degree of correlation between the RAPTOR and reference data ($R^2>0.87$) for a 2-month period.

2.5 Quantification of final ozone concentrations for CAPTOR and RAPTOR nodes

The final ozone concentration time series were calculated for each node following the methodology described in the previous sections, and shared with the hosts as the result of the summer monitoring campaign. The concentration time series for the campaign, together with the time series and correlation plot for the calibration period and sensor, which were selected as final, are shown for each individual node in the Annex. For the sake of brevity, only the results from the 9 CAPTOR and 9 RAPTOR nodes which were deployed at reference stations are shown in the main

text (Figure 12). The main conclusions extracted from this comparison are:

- CAPTOR: R^2 coefficients were relatively high, exceeding 0.80 (range: 0.80-0.95) for all of the 35 CAPTOR nodes except for one ($R^2=0.78$). Correlation coefficients reached >0.90 for 14 nodes. For the 9 CAPTOR nodes deployed at reference stations, where the datasets were longer than 2 months (and this more prone to technical failures and higher data dispersion), R^2 values ranged between 0.82-0.91.
- RAPTOR: R^2 values ranged between 0.70 and 0.95 for the 9 nodes at reference stations, similarly to the CAPTOR nodes.
- CAPTOR: The nodes seem to present an upper detection limit at approximately 150-170 $\mu\text{g}/\text{m}^3$. This is a limitation of the sensing technology, which cannot be overcome with the calibration algorithm. This upper limit may be clearly observed in nodes Manlleu, Vic and Tona (Figure 12). However, the 2 nodes in Monte Cucco were able to reproduce higher concentrations ($>170 \mu\text{g}/\text{m}^3$). These different behaviours are yet unexplained, and could be related to the individual variability of the sensors. It suggests that a non-linear regression approach may be more appropriate for this kind of sensor, as opposed to the linear approach used so far. Further research is necessary to clarify this issue.
- RAPTOR: the RAPTOR nodes did not show an upper limit. This would imply that the linear regression approach applied could be adequate for this kind of sensing devices.
- CAPTOR & RAPTOR: Over all, the nodes at reference stations were able to produce relatively long (>2 months) time series, with few technical issues and working autonomously (without frequent interaction being required).

CAPTORS:

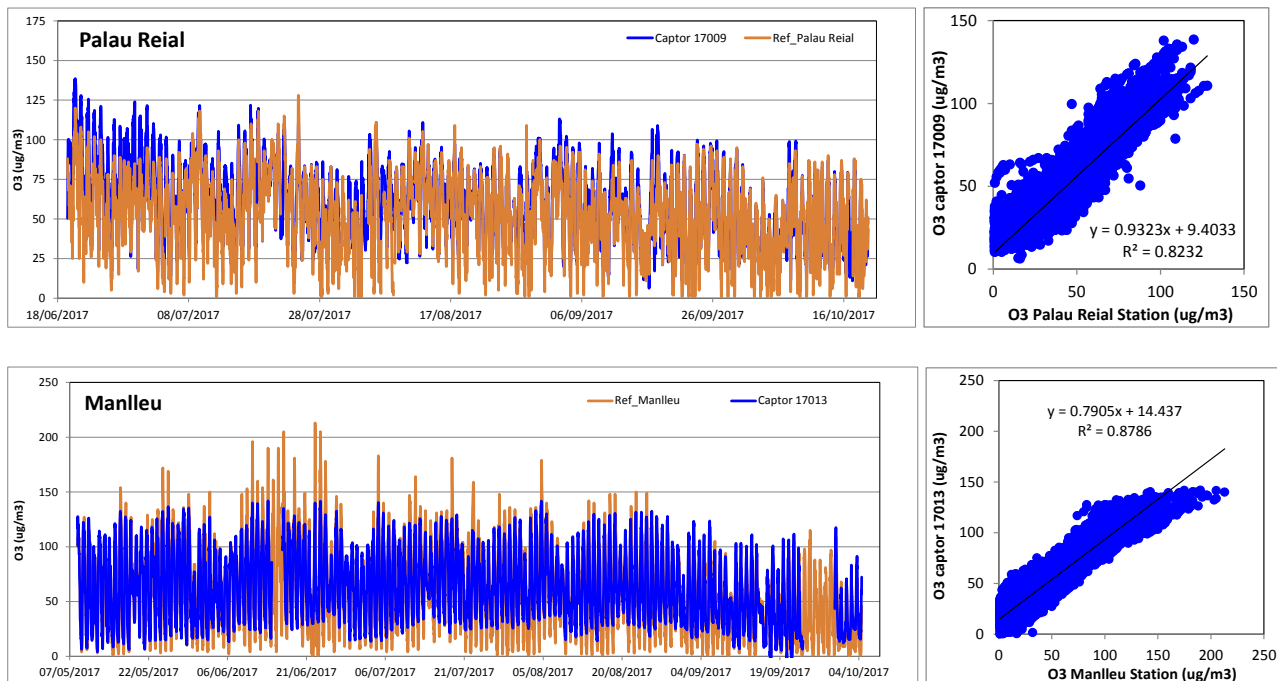


Figure 12. Comparison between reference ozone concentrations and those calculated with CAPTOR nodes at reference stations during the entire summer campaign.

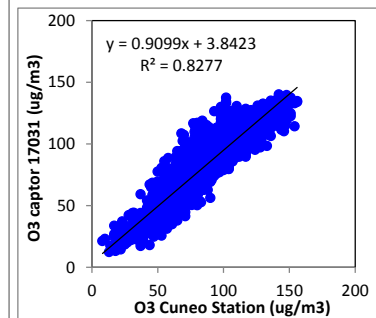
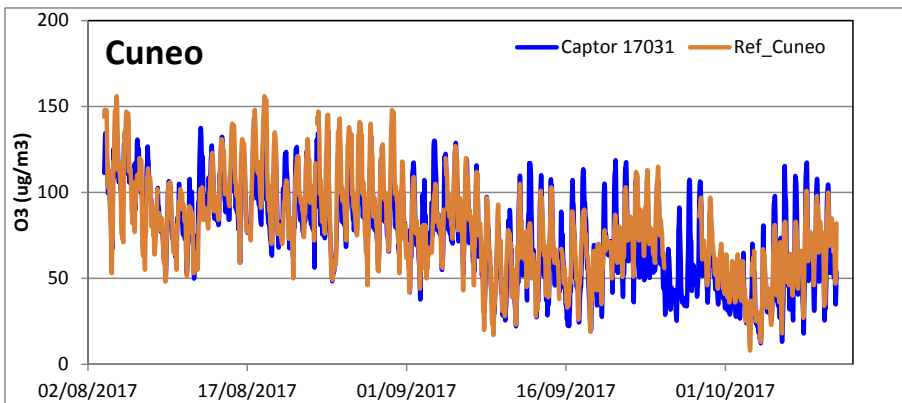
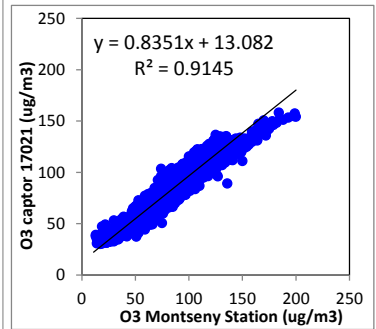
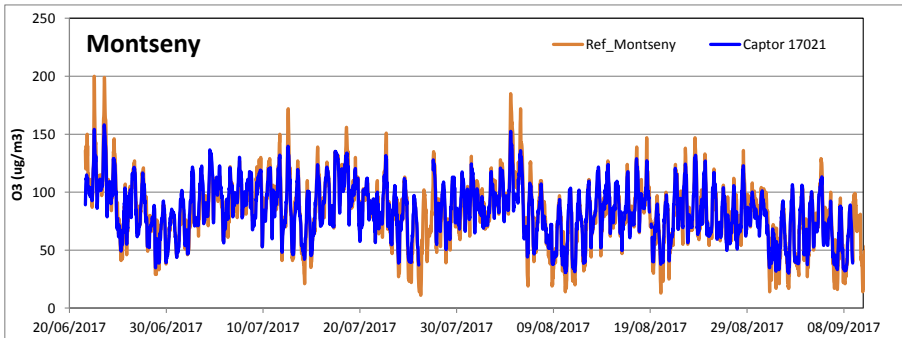
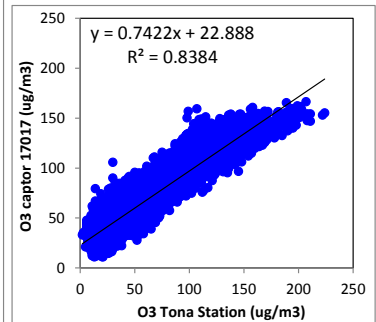
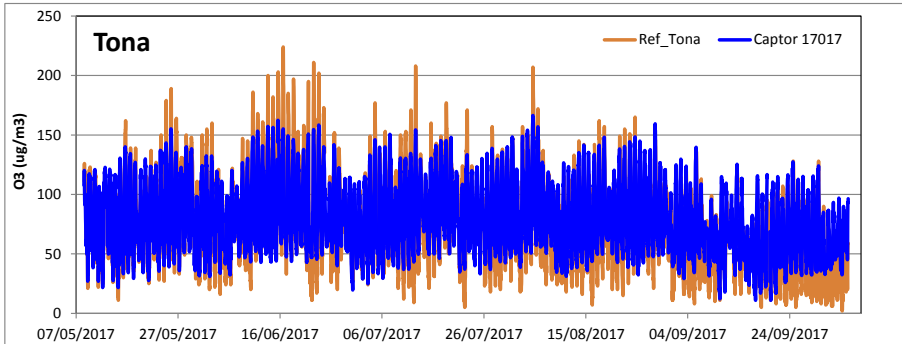
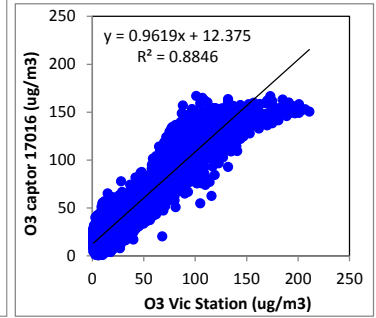
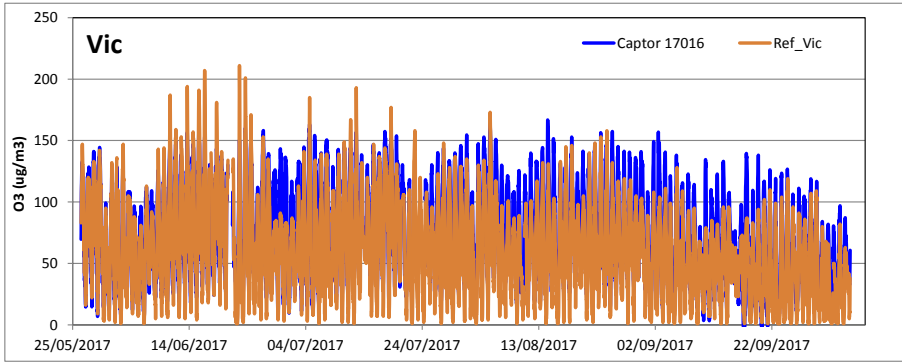
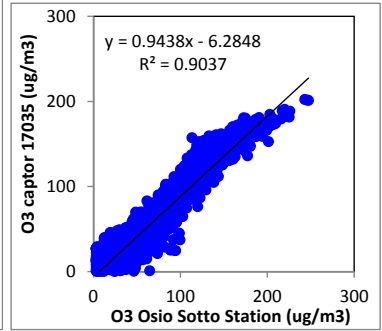
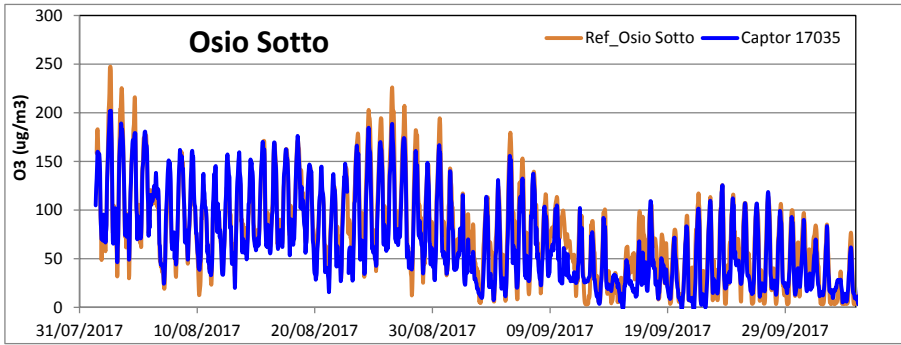
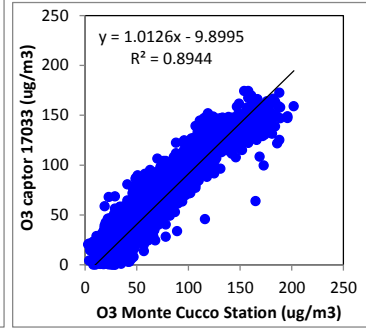
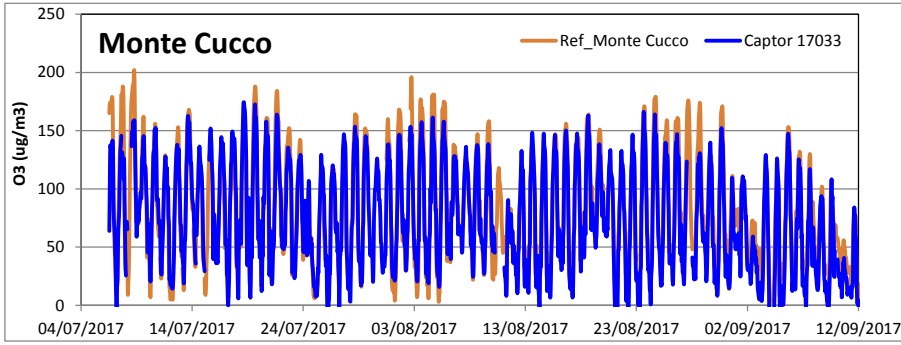


Figure 12. Continued.



RAPTORS:

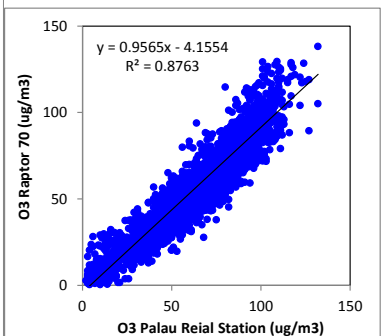
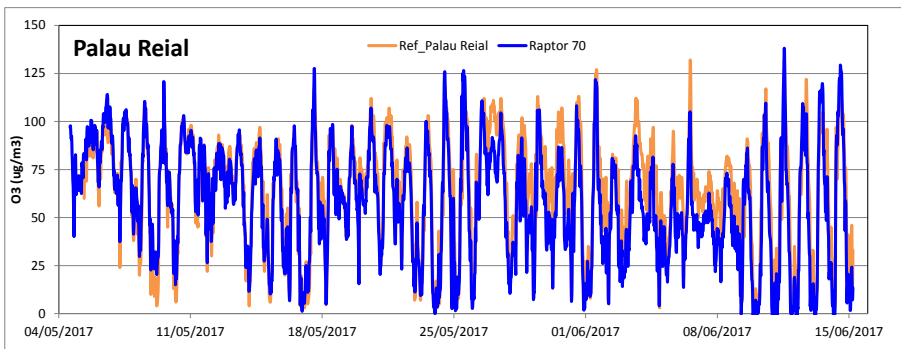
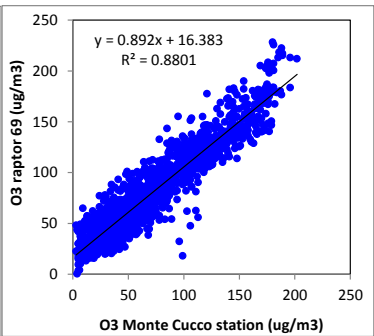
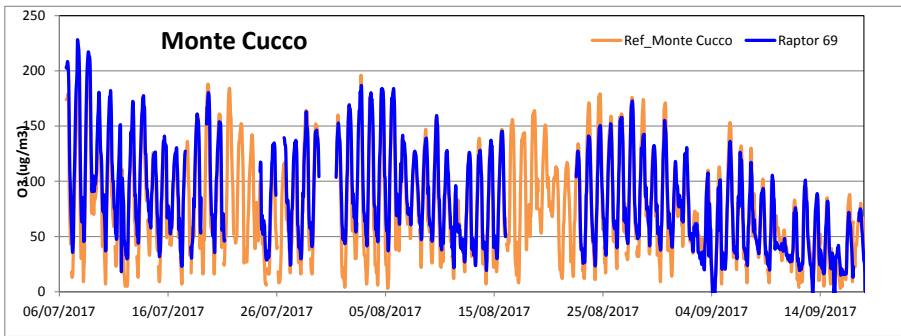


Figure 12. Continued.

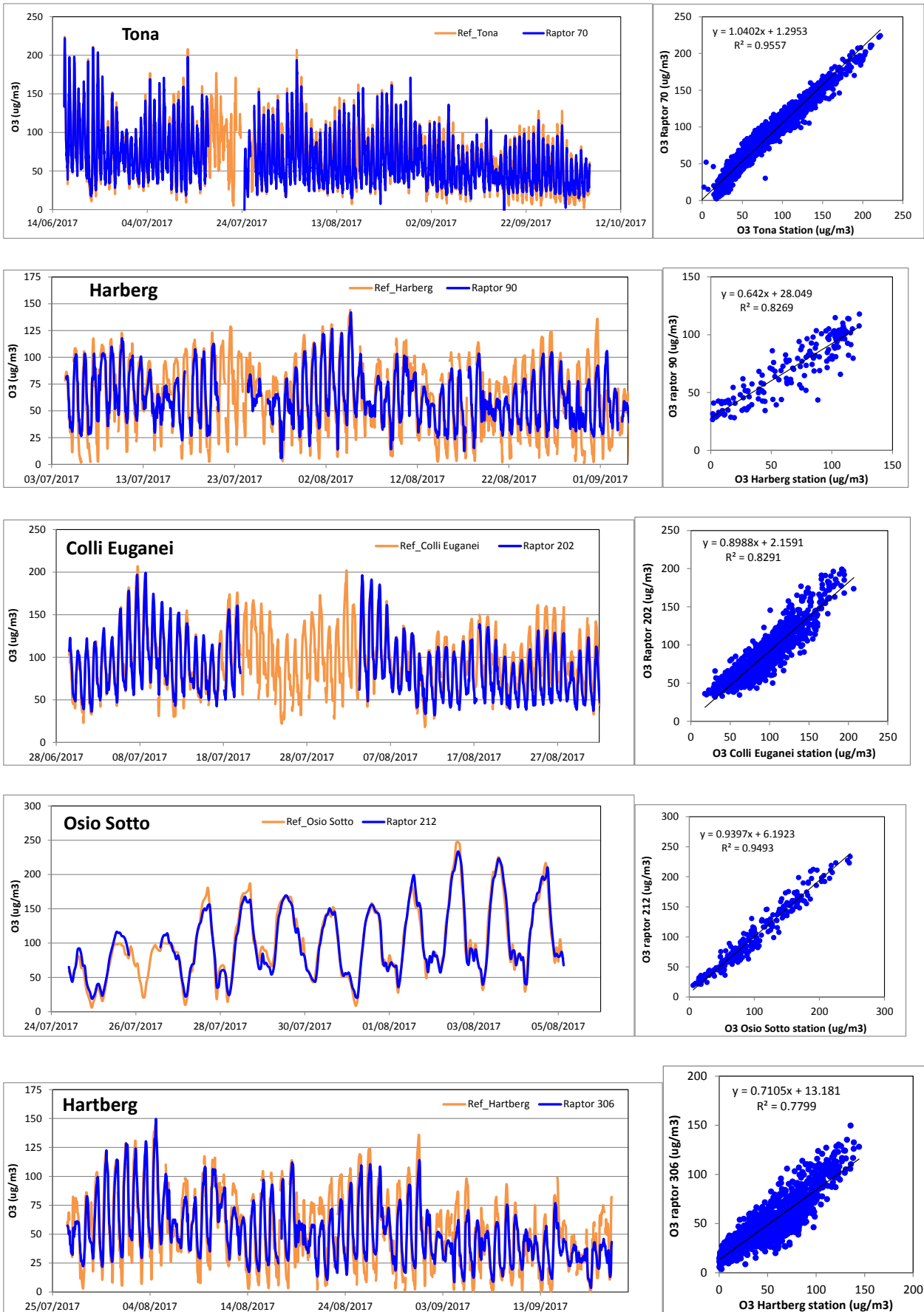


Figure 12. Continued.

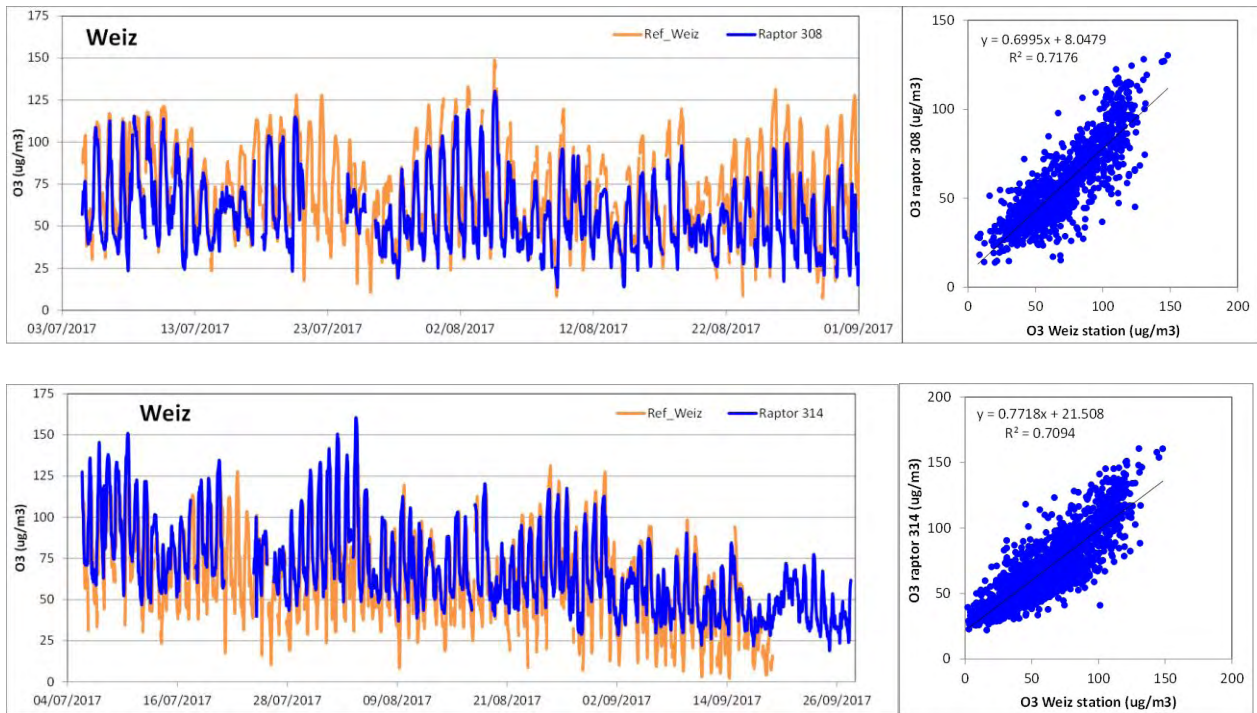


Figure 12. Continued.

The ultimate goal of CAPTOR as a project is to produce scientifically robust ozone concentration time series to monitor ozone pollution with a citizen science approach. Thus, once the time series in the section above were obtained and validated, an environmental assessment was carried out by comparing the concentrations obtained with legislative and guideline limit values from EU and the World Health Organisation (WHO). The results from this assessment are summarised in the Tables below, for the Spain, Italy and Austria nodes. These Tables were provided to the NGO partners of CAPTOR, who in turn distributed them among the CAPTOR and RAPTOR hosts, thus fulfilling the dissemination and communication goals of the project.

Table 4. Environmental assessment of ozone concentrations monitored with the CAPTOR and RAPTOR nodes in Spain. Comparison with EU limit and WHO guideline values.

Air quality standards	Averaging period	Concentration	Number of exceedances									
			Manlleu (reference station)	Manlleu (CAP17013)	Vic (reference station)	Vic (CAP17016)	Cantonigròs (CAP17001)	St. Vicenç de Torello (CAP17002)	St. Pere de Torello (CAP17003)	Gurb (CAP17005)	Sta. Cecília de Voltregà (CAP17010)	St Martí Sescorts (CAP17011)
Human health target value	Maximum daily 8-hour mean	100 µg/m ³ (WHO)	77	74	70	58	48	42	67	36	77	80
		120 µg/m ³ , not to be exceeded on more than 25 days per year averaged over 3 years	31	12	29	26	0	2	20	1	41	33
Information threshold	1 hour	180 µg/m ³	7	0	12	0	0	0	0	0	0	0
Alert threshold	1 hour	240 µg/m ³	0	0	0	0	0	0	0	0	0	0

Air quality standards	Averaging period	Concentration	Number of exceedances									
			Tona (reference station)	Tona (CAP17017)	Tona (RAP070)	Centelles (CAP17004)	Sta. Eulalia de Riuprimer (CAP17006)	Hostalets de Balenyà (CAP17007)	Taradell (CAP17012)	Calldetenes (CAP17014)	Tona (CAP17023)	Sta. Eugènia de Berga (CAP17027)
Human health target value	Maximum daily 8-hour mean	100 µg/m ³ (WHO)	98	94	51	52	43	52	86	68	38	49
		120 µg/m ³ , not to be exceeded on more than 25 days per year averaged over 3 years	59	55	26	7	0	10	54	17	2	0
Information threshold	1 hour	180 µg/m ³	28	0	19	0	0	0	10	0	0	0
Alert threshold	1 hour	240 µg/m ³	0	0	0	0	0	0	0	0	0	0

Table 4. Continued.

Air quality standards	Averaging period	Concentration	Number of exceedances										
			Palau Reial (reference station)	Palau Reial (CAP17009)	Palau Reial (RAP070)	Montseny (reference station)	Montseny (CAP17021)	Llinars del Vallès (CAP17022)	Cànoves (CAP17025)	Badalona (CAP17018)	Barcelona (CAP17019)	St Cebrià de Vallalta (CAP17020)	Vilanova del Vallès (CAP17026)
Human health target value	Maximum daily 8-hour mean	100 µg/m ³ (WHO)	5	2	4	49	46	28	36	3	0	2	5
		120 µg/m ³ , not to be exceeded on more than 25 days per year averaged over 3 years	0	0	0	15	8	9	4	0	0	0	0
Information threshold	1 hour	180 µg/m ³	0	0	0	5	0	0	0	0	0	0	0
Alert threshold	1 hour	240 µg/m ³	0	0	0	0	0	0	0	0	0	0	0

Table 5. Environmental assessment of ozone concentrations monitored with the CAPTOR and RAPTOR nodes in Italy. Comparison with EU limit and WHO guideline values.

Air quality standards	Averaging period	Concentration	Number of exceedances				
			Osio Sotto (reference station)	Osio Sotto (CAP17035)	Osio Sotto (RAP212)	Pontida (RAP214)	Stezzano (RAP218)
Human health target value	Maximum daily 8-hour mean	100 µg/m ³ (WHO)	47	34	12	4	27
		120 µg/m ³ , not to be exceeded on more than 25 days per year averaged over 3 years	34	29	10	4	20
Information threshold	1 hour	180 µg/m ³	45	13	21	23	26
Alert threshold	1 hour	240 µg/m ³	3	0	0	0	0

Table 5. Continued.

Air quality standards	Averaging period	Concentration	Number of exceedances					
			Cuneo (reference station)	Cuneo (CAP17031)	Boves (CAP17024)	Carmagnola (CAP17029)	Cuneo (RAP204)	Peveragno (RAP210)
Human health target value	Maximum daily 8-hour mean	100 µg/m ³ (WHO)	32	32	50	31	15	24
		120 µg/m ³ , not to be exceeded on more than 25 days per year averaged over 3 years	19	12	38	0	3	3
Information threshold	1 hour	180 µg/m ³	0	0	0	0	0	0
Alert threshold	1 hour	240 µg/m ³	0	0	0	0	0	0

Air quality standards	Averaging period	Concentration	Number of exceedances					
			Monte Cucco (reference station)	Monte Cucco (CAP17033)	Monte Cucco (RAP069)	Gossolengo (CAP17028)	Pontenure (CAP17034)	Cortemaggiore (RAP216)
Human health target value	Maximum daily 8-hour mean	100 µg/m ³ (WHO)	57	58	51	37	32	21
		120 µg/m ³ , not to be exceeded on more than 25 days per year averaged over 3 years	45	43	41	26	6	5
Information threshold	1 hour	180 µg/m ³	21	0	37	0	0	0
Alert threshold	1 hour	240 µg/m ³	0	0	0	0	0	0

Table 5. Continued.

Air quality standards	Averaging period	Concentration	Number of exceedances				
			Veneto Colli Euganei (reference station)	Veneto Colli Euganei (RAP202)	Barbarano Vicentino (RAP206)	Isola Vicentina (RAP208)	Ponte di Barbarano (CAP17030)
Human health target value	Maximum daily 8-hour mean	100 µg/m ³ (WHO)	65	43	31	30	25
		120 µg/m ³ , not to be exceeded on more than 25 days per year averaged over 3 years	43	22	16	16	10
Information threshold	1 hour	180 µg/m ³	21	17	3	15	0
Alert threshold	1 hour	240 µg/m ³	0	0	0	0	0

Table 6. Environmental assessment of ozone concentrations monitored with the CAPTOR and RAPTOR nodes in Austria. Comparison with EU limit and WHO guideline values.

Air quality standards	Averaging period	Concentration	Number of exceedances												
			Hartberg reference station	Weiz reference station	RAP-72 Unteraichen	RAP-73 Hainersdorf Wickie	RAP-90 Hartberg Hauptplatz	RAP-302 Winzendorf	RAP-304 St.Katherin	RAP-306 Hartberg HERZ	RAP-308 Weiz bahnhof	RAP-310 NMS Gleisdorf	RAP-312 Pöllauberg	RAP-314 Weiz Hauptplatz	RAP-316 W.E.I.Z. Innov.
Human health target value	Maximum daily 8-hour mean	100 µ/m ³ (WHO)	26	32	3	4	13	13	0	15	13	0	7	29	10
		120 µ/m ³ , not to be exceeded on more than 25 days per year averaged over 3 years	3	3	0	0	2	1	0	2	1	0	0	14	0
Information threshold	1 hour	180 µ/m ³	0	0	0	0	0	0	0	0	0	0	0	0	0
Alert threshold	1 hour	240 µ/m ³	0	0	0	0	0	0	0	0	0	0	0	0	0



3. Summary and conclusions

The main conclusions extracted from this deliverable are summarised in this section.

Operational:

- **The 2017 summer ozone campaign provided significantly improved results with regards to the 2016 campaign in terms of data availability.**
- This was mainly due to improvements in terms of hardware (internal design of the nodes) and software (data transmission, storage and access).
- RAPTOR nodes were deployed systematically in 2017, as opposed to 2016 when they were still in a test phase.
- CAPTOR and RAPTOR nodes were deployed systematically for the first time in Italy and Austria.
- The CAPTOR and RAPTOR nodes were able to work mostly autonomously during the 2017 summer campaign, without as intensive technical support as was required in 2016. This is considered a major improvement.

Calibration of nodes:

- **A 2-step calibration procedure (at a central reference station, followed by a local reference station) was successfully implemented**, with positive results in terms of increased data availability and calibration of the nodes at concentrations closer to their target.
- A specific methodology was devised to calibrate the nodes, based on the RMSE of the individual sensors and on the correlation (R^2) between the calculated and reference ozone concentration time series.
- CAPTOR nodes: different behaviours were observed resulting from the intrinsic variability of the sensing devices. While in some nodes the best performing sensor was stable over time, for the majority of nodes this sensor was different in the Calibration 1 and Calibration 2 periods. This increased the complexity of selecting the most adequate beta coefficient for the calculation of the final ozone time series for the campaign period.

Performance of the nodes:

- **CAPTOR nodes: R^2 coefficients were relatively high, ranging $R^2=0.80-0.95$ for all of the 35 CAPTOR nodes** except for one ($R^2=0.78$). Correlation coefficients reached >0.90 for 14 nodes. For the 9 CAPTOR nodes deployed at reference stations, where the datasets were longer than 2 months (and this more prone to technical failures and higher data dispersion), R^2 values ranged between 0.82-0.91.
- **Certain CAPTOR nodes showed an upper detection limit at approximately $150-170 \mu\text{g}/\text{m}^3$.** This is a limitation of the sensing devices (hardware), which cannot be overcome with the calibration algorithm. It suggests that a non-linear regression approach may be more appropriate for this kind of sensor, as opposed to the linear approach used so far. Further research is necessary to address this issue. This upper limit was not detected for RAPTOR nodes, for which the linear regression seems adequate.
- **RAPTOR nodes: the comparability with reference data was also good, with $R^2=0.78-0.95$ for the 9 nodes at reference stations** (one exception with $R^2=0.70$).

Citizen science approach:

- The CAPTOR and RAPTOR nodes in 2017 were much more adequate for their deployment by Project funded by the European Union's Horizon 2020 research and innovation programme under Grant Agreement N° 688110

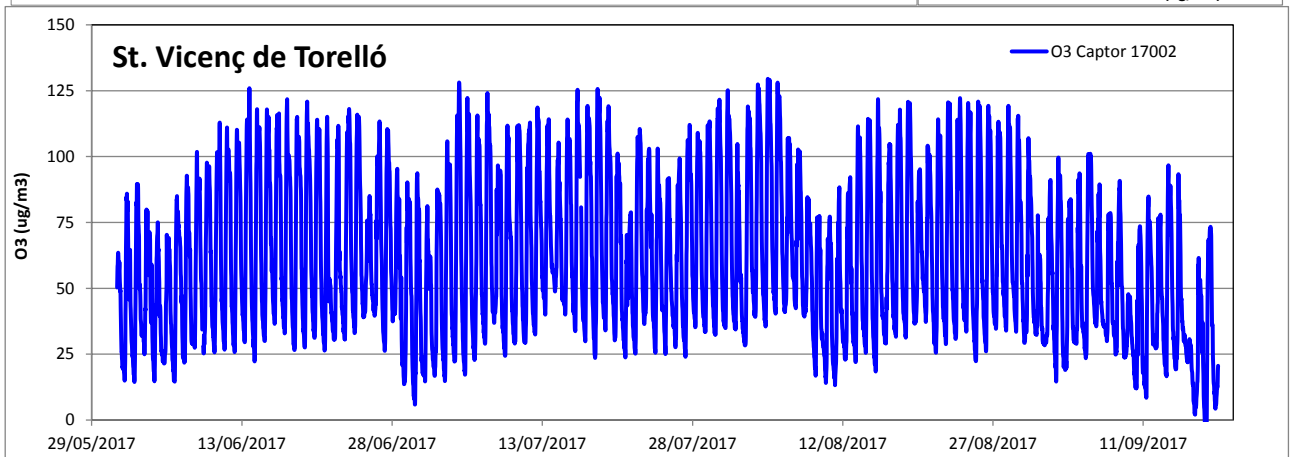
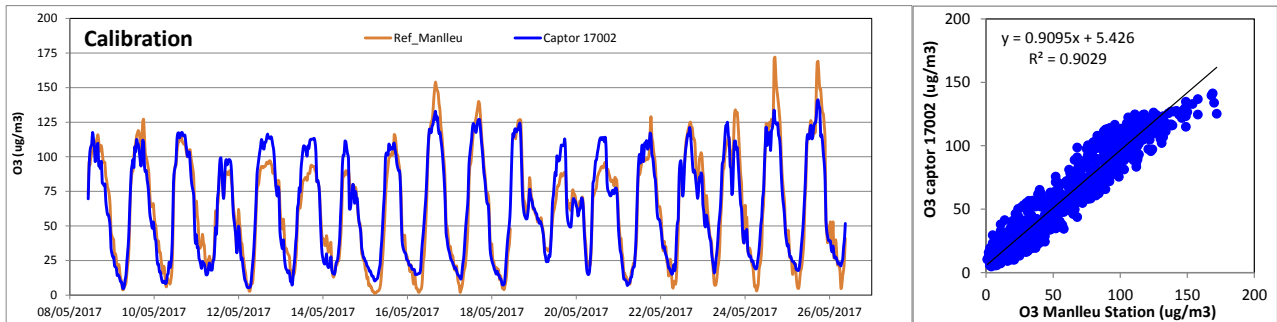
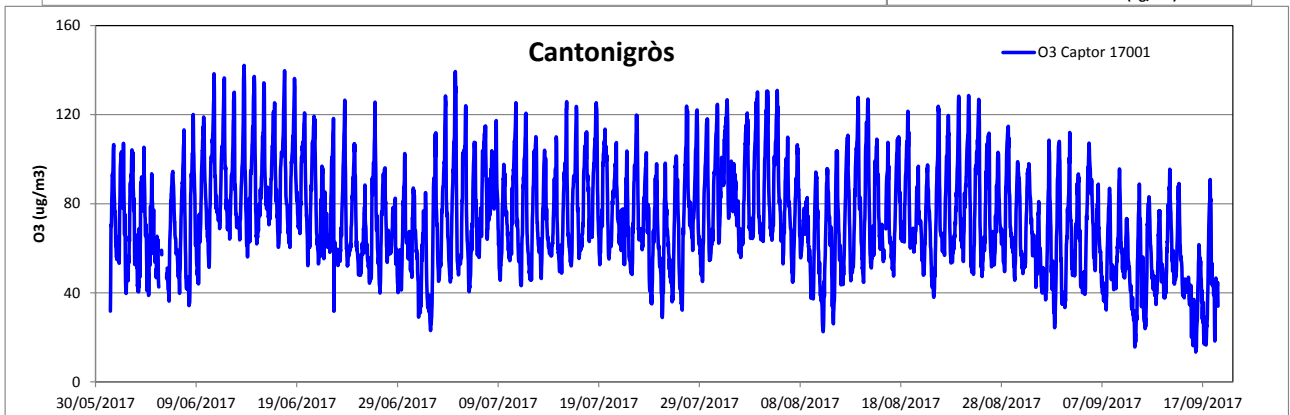
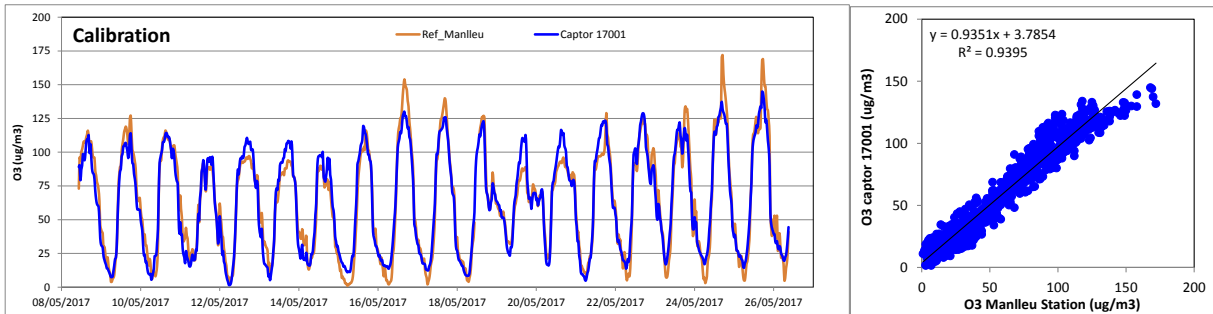
CAPTOR

citizens than in 2016. They used an internal 3G connection with which they were no longer reliant on the host's WiFi connection, and they required much less interaction with the user.

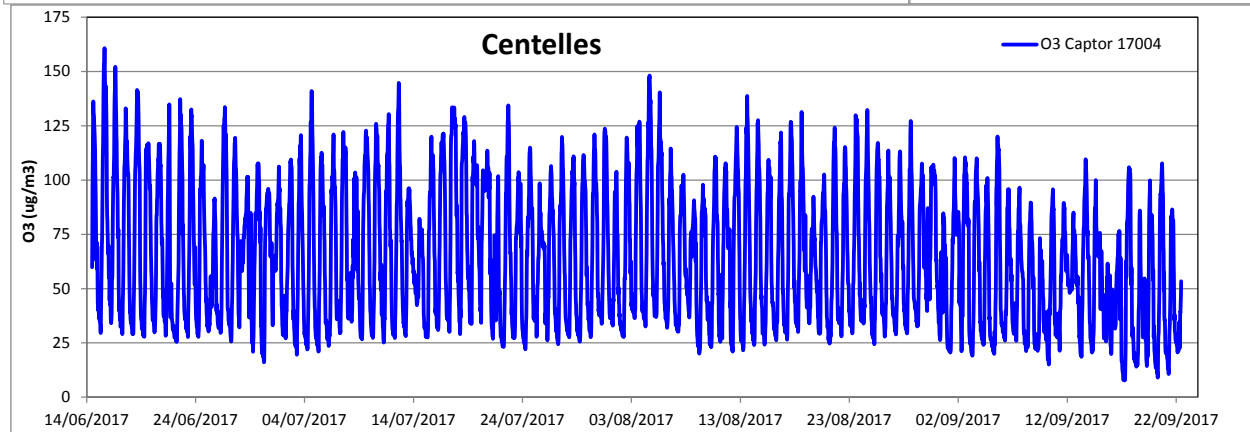
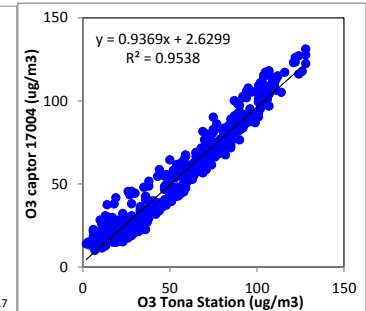
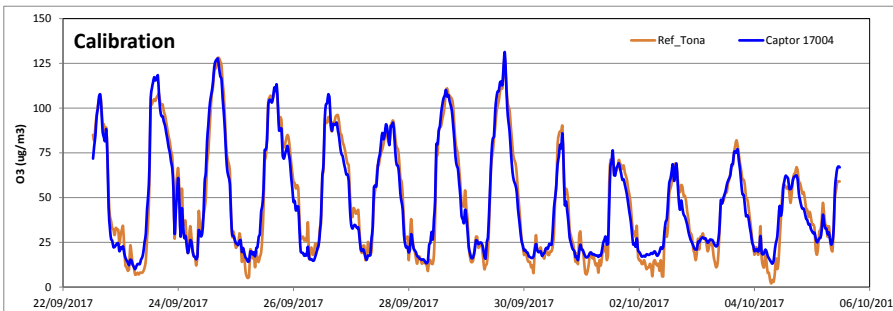
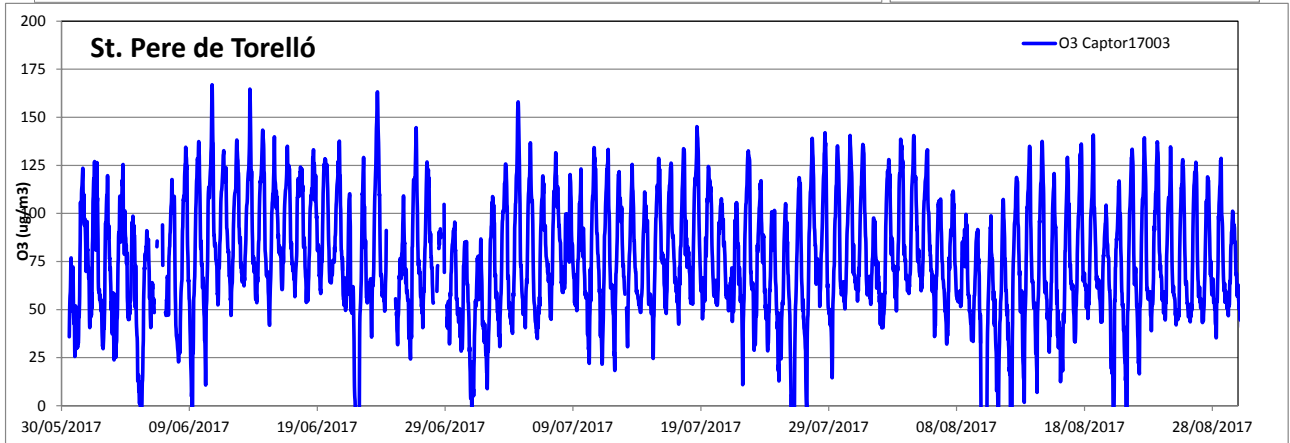
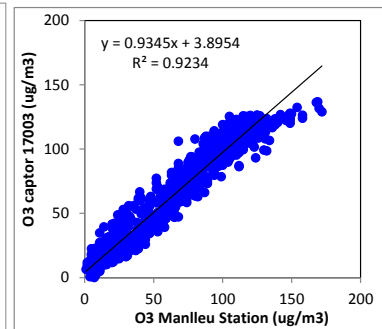
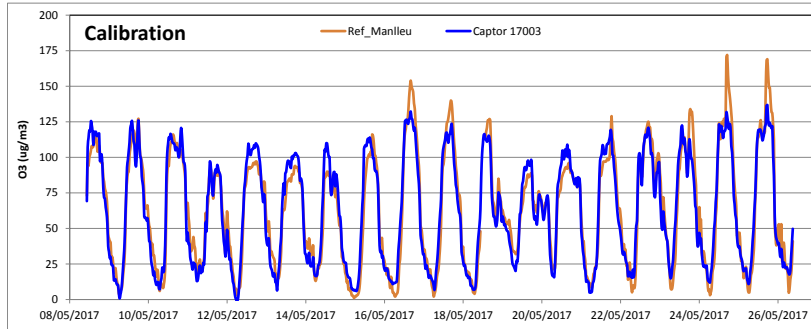
- Despite this, **neither the CAPTOR or the RAPTOR nodes are at present at a technology readiness level where they may be used solely by volunteers: the data processing work behind the final time series is still highly complex and requires high-level scientific and technical support by experts in the fields of air quality monitoring and data processing.**

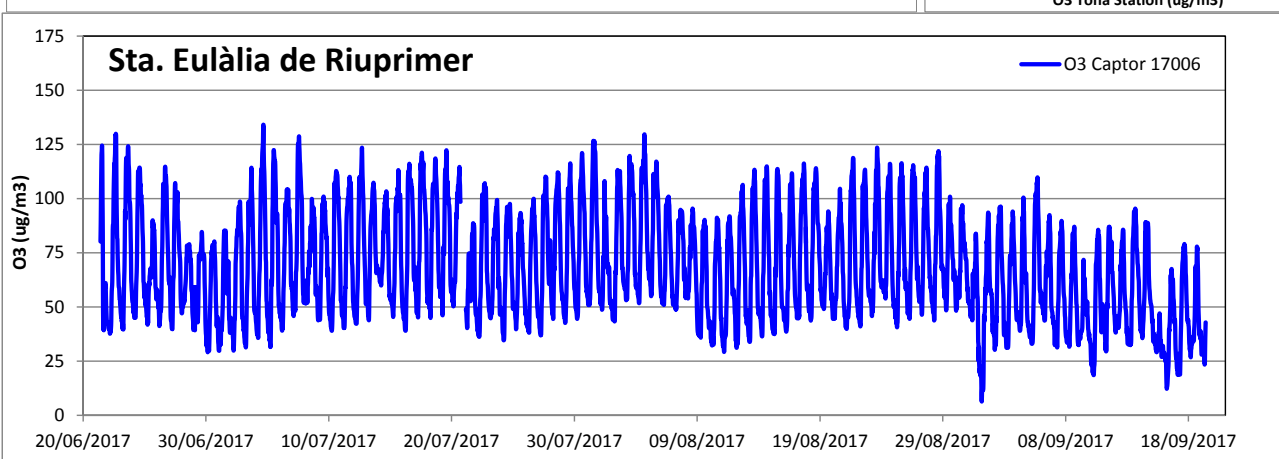
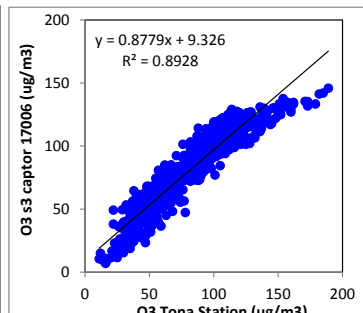
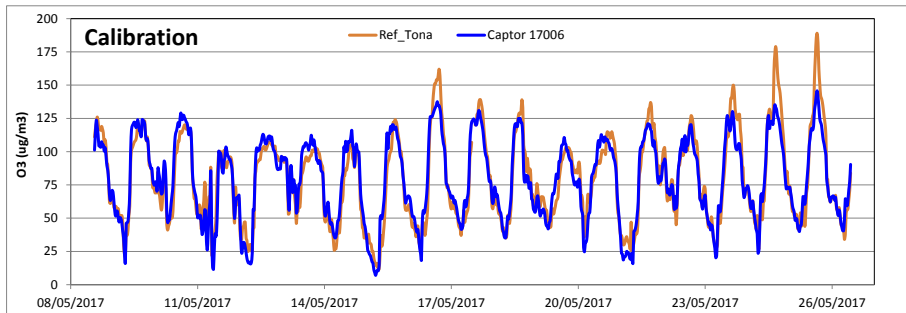
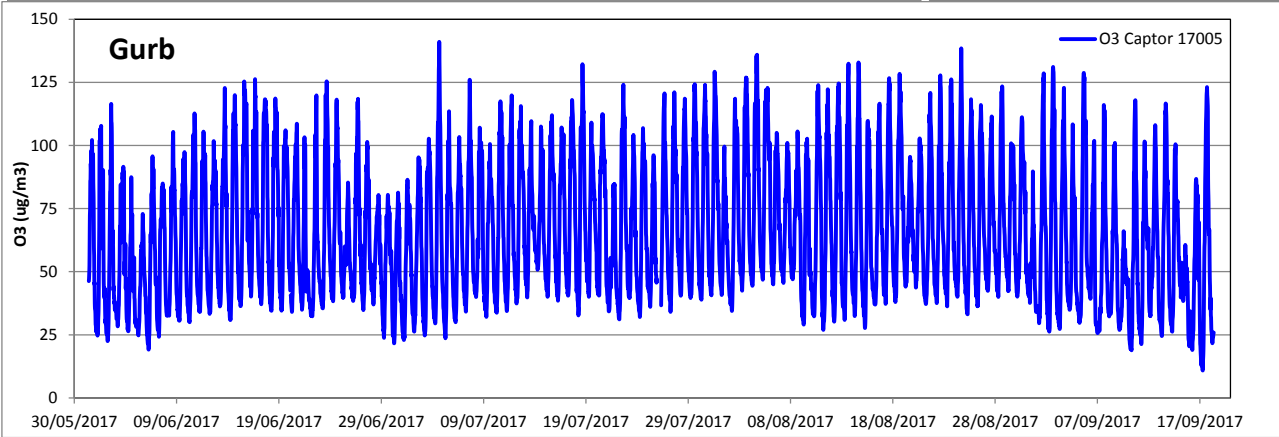
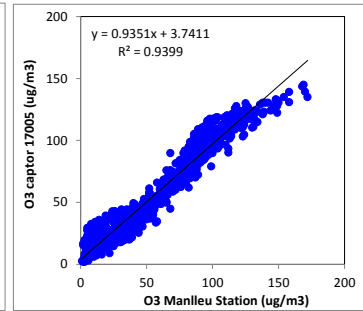
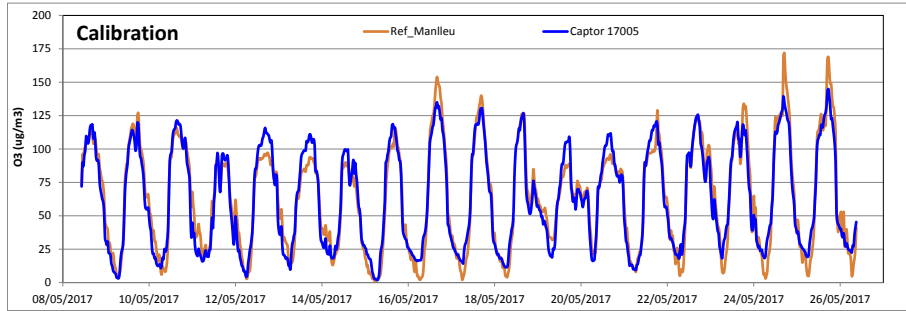
4. Annex

CAPTORS:

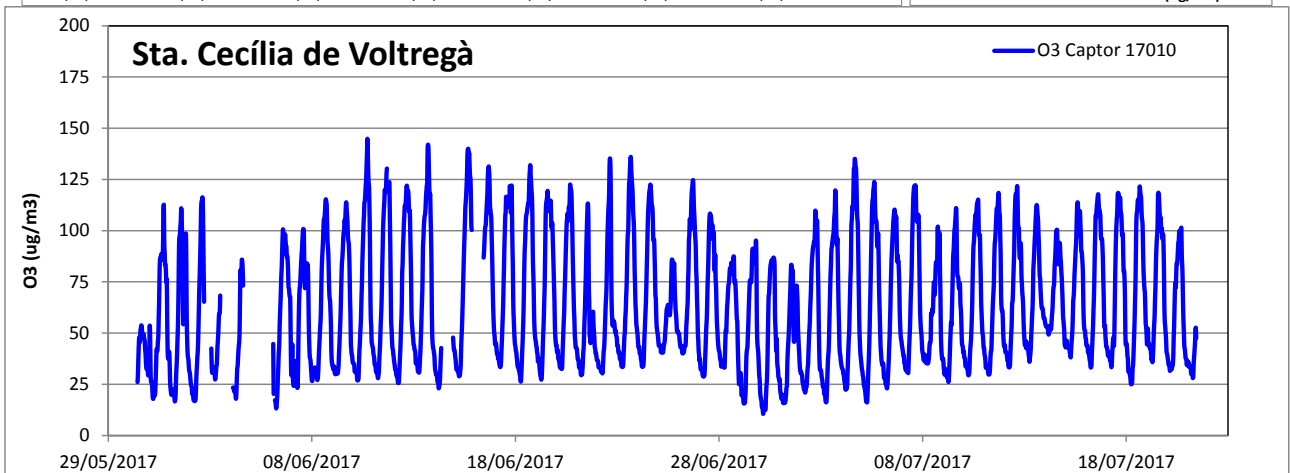
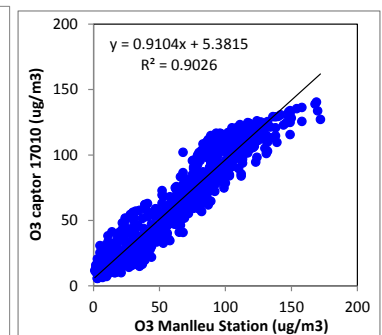
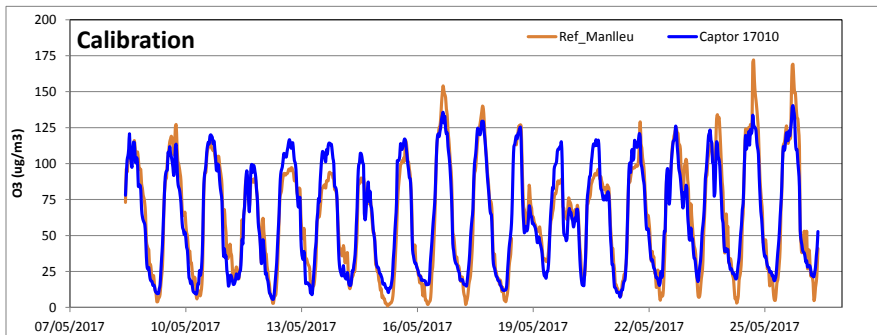
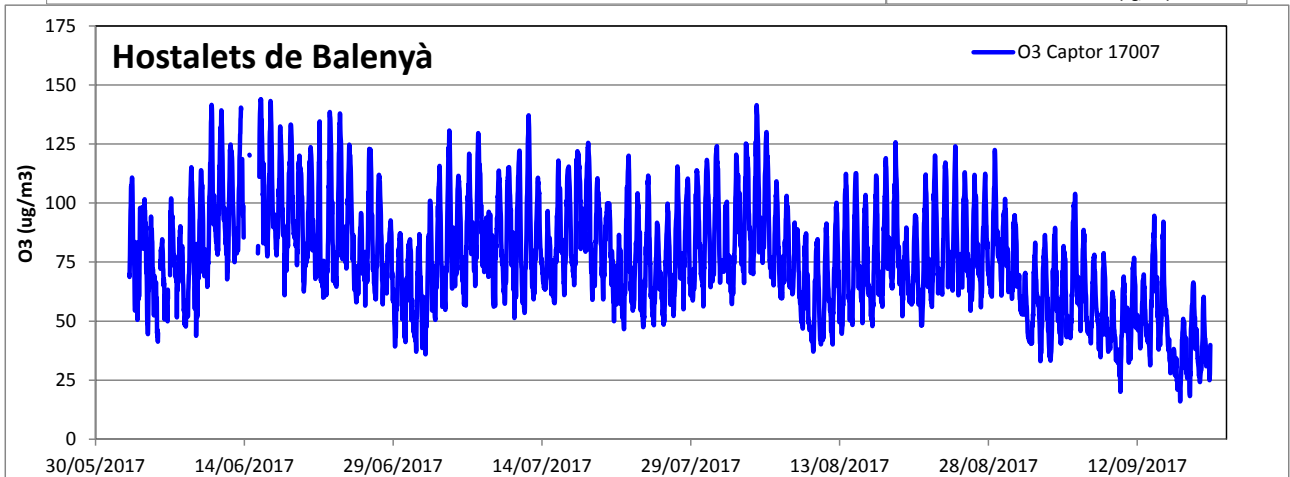
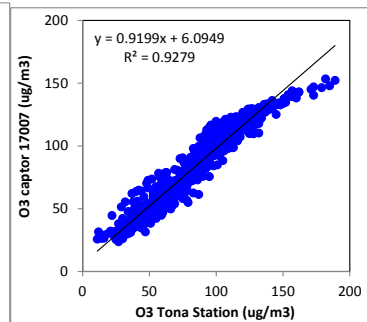
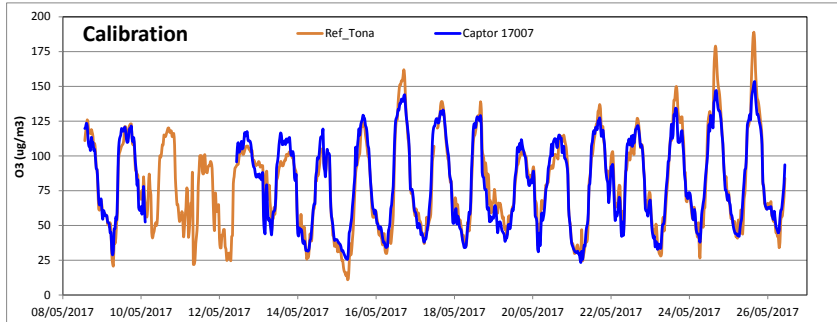


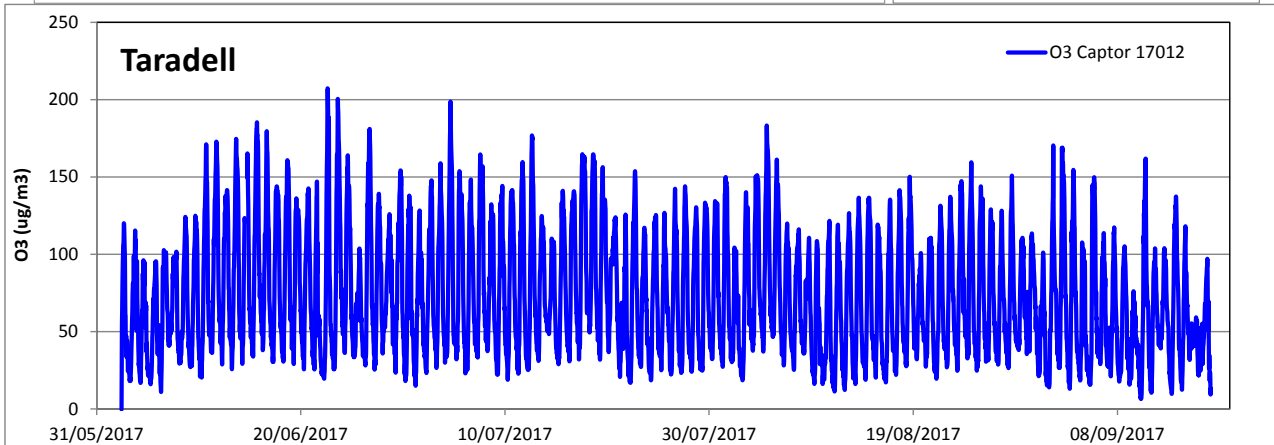
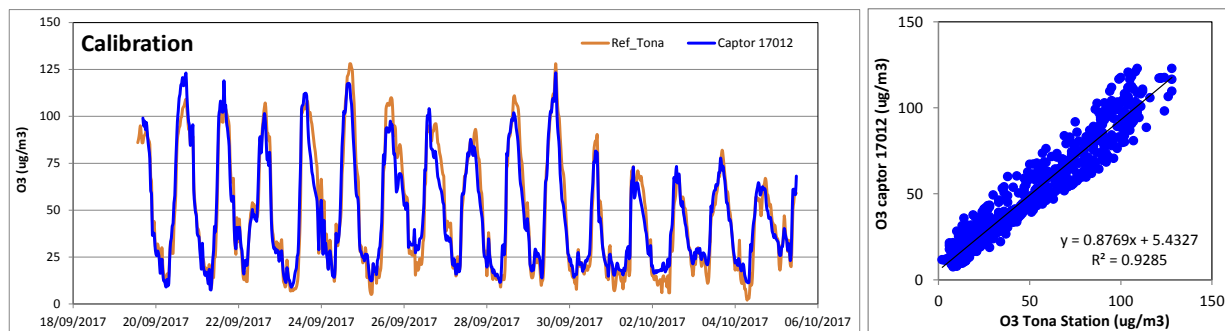
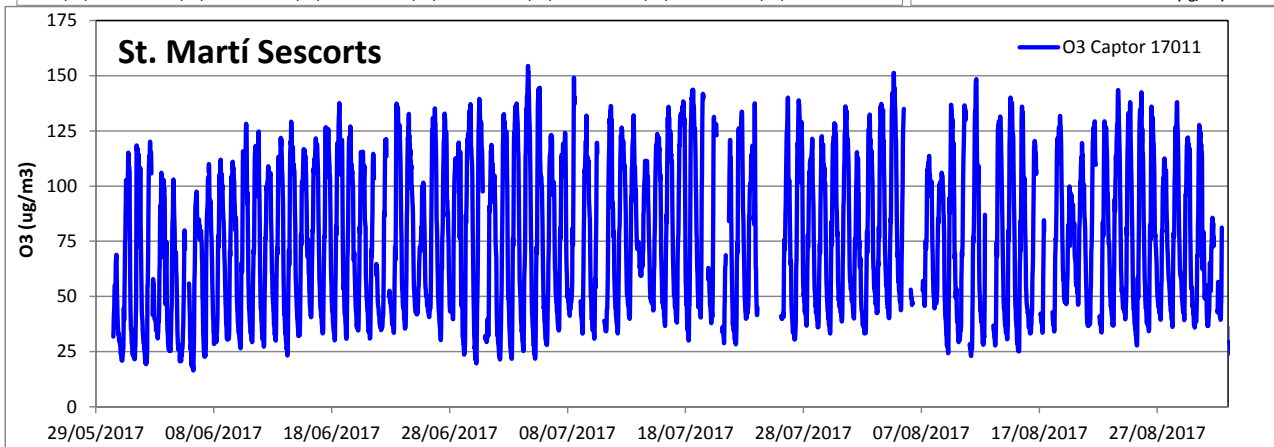
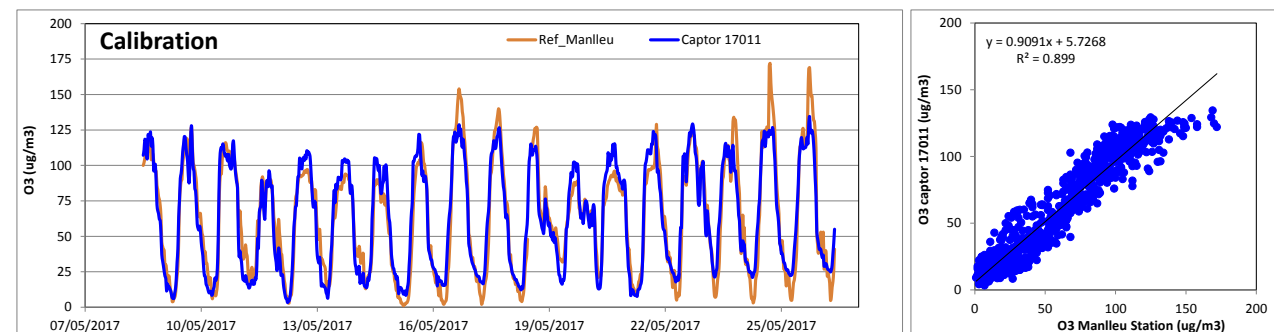
CAPTOR



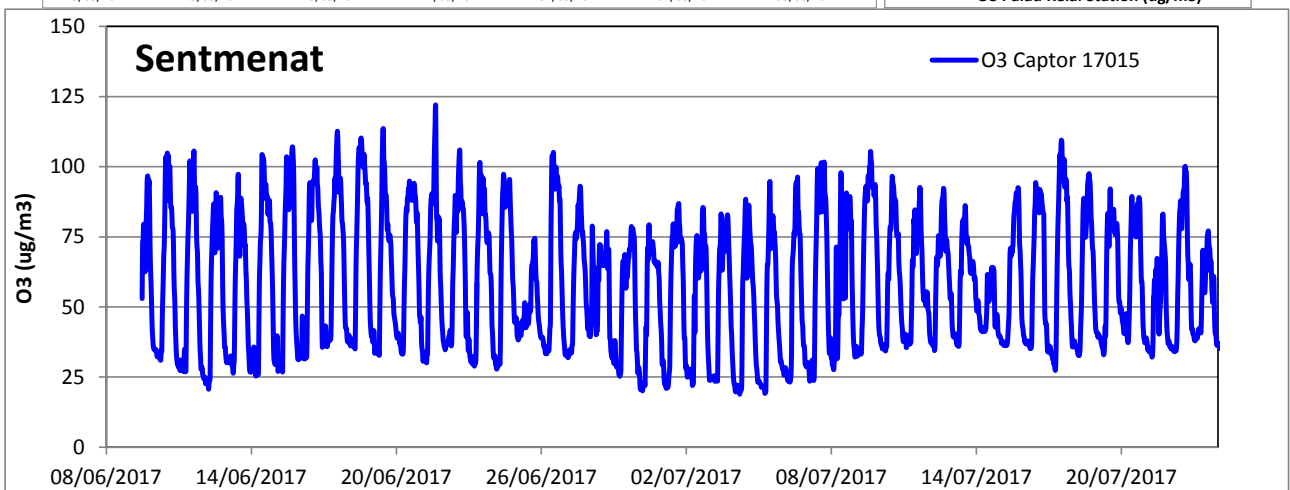
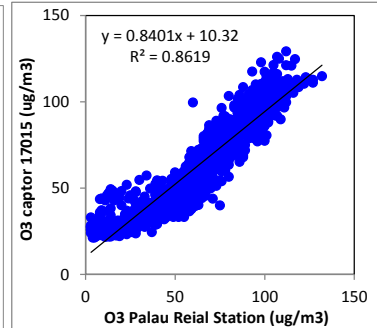
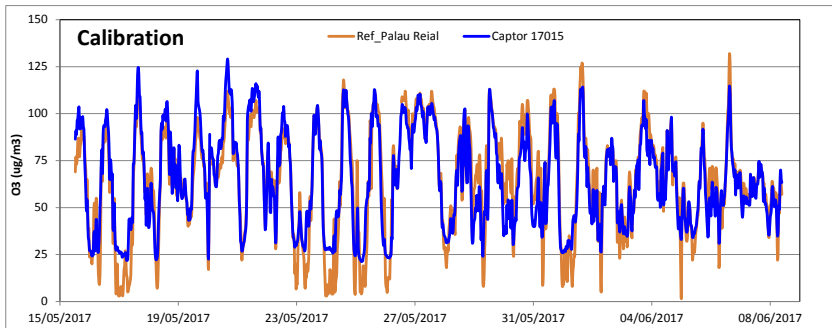
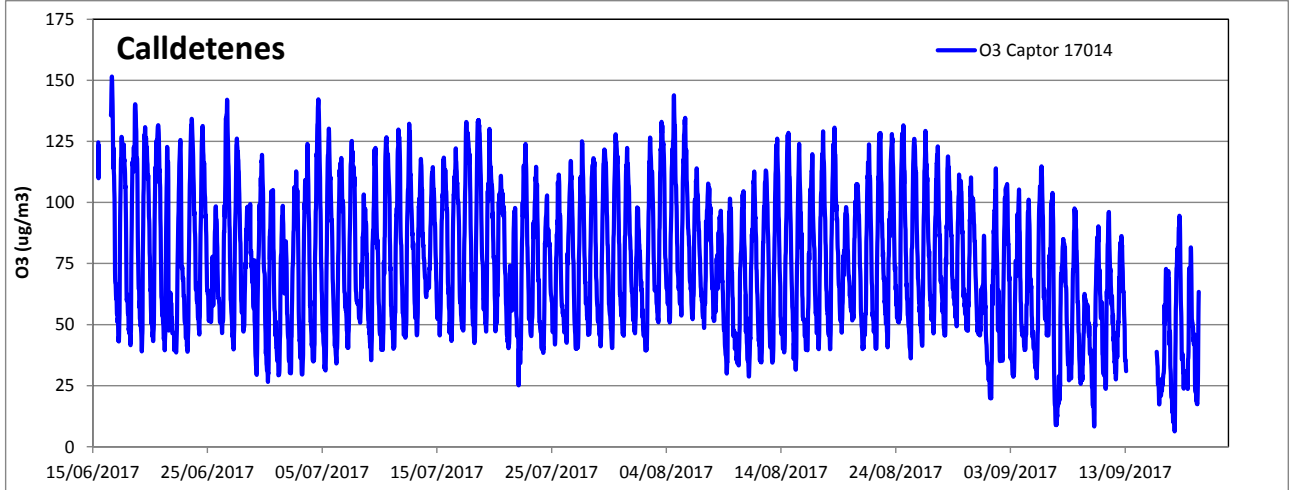
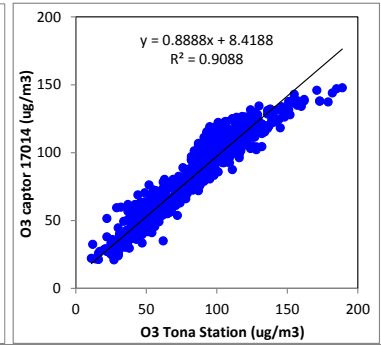
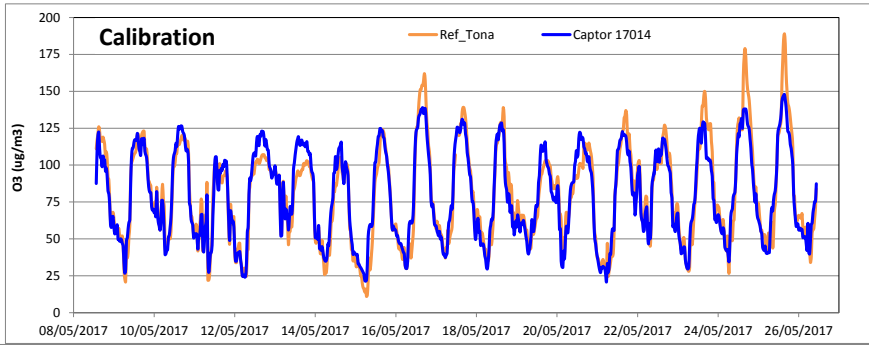


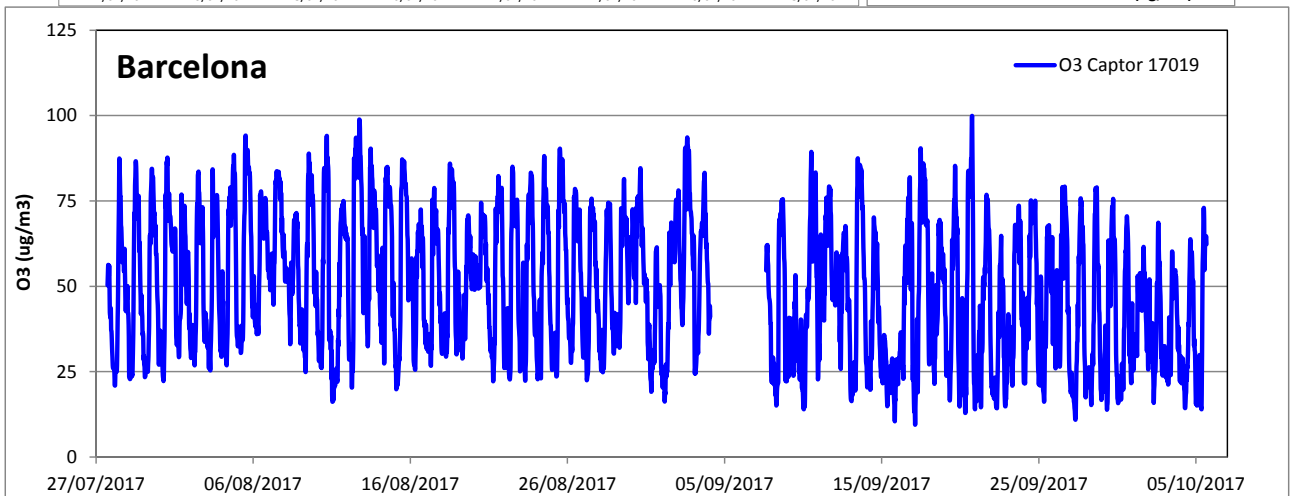
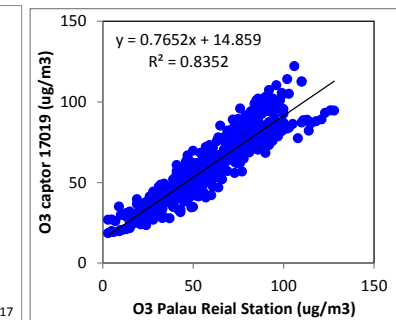
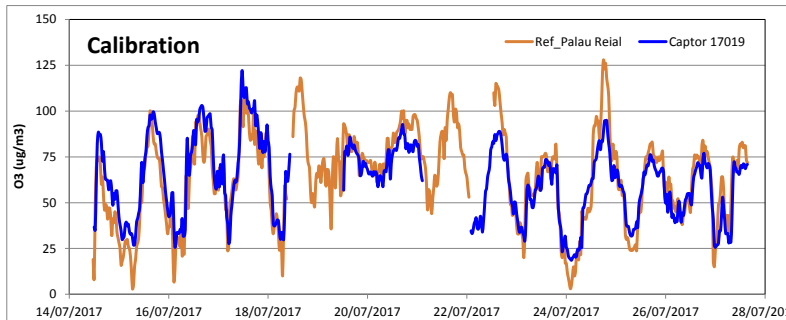
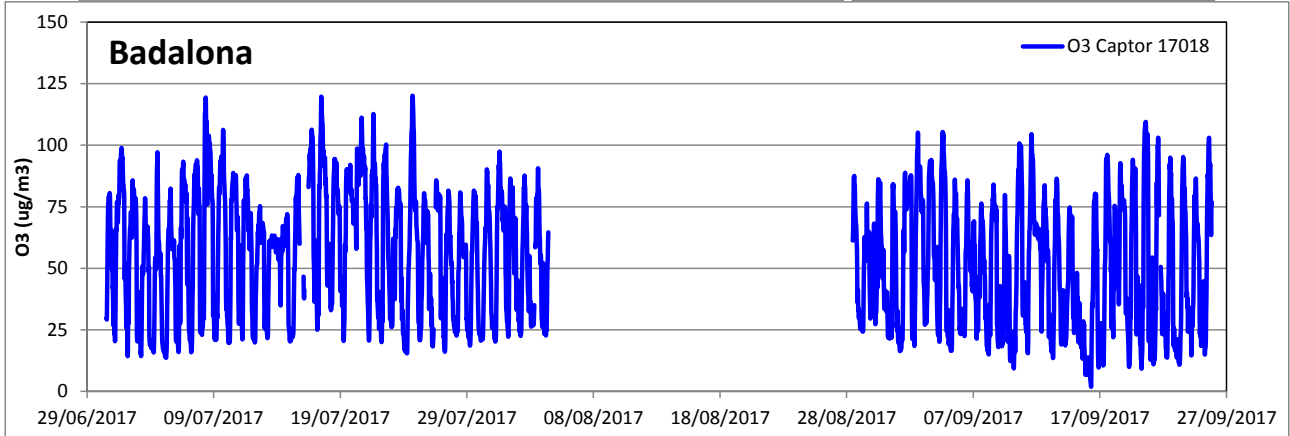
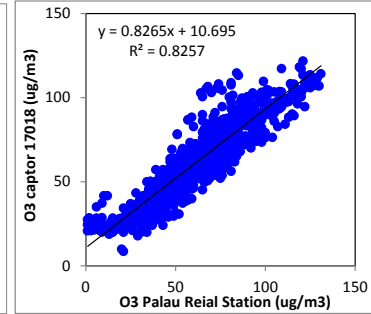
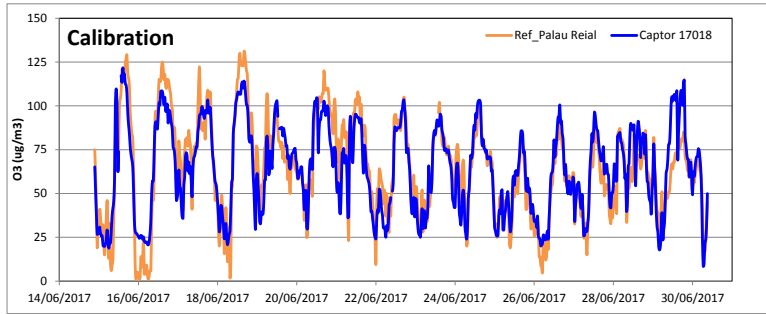
CAPTOR

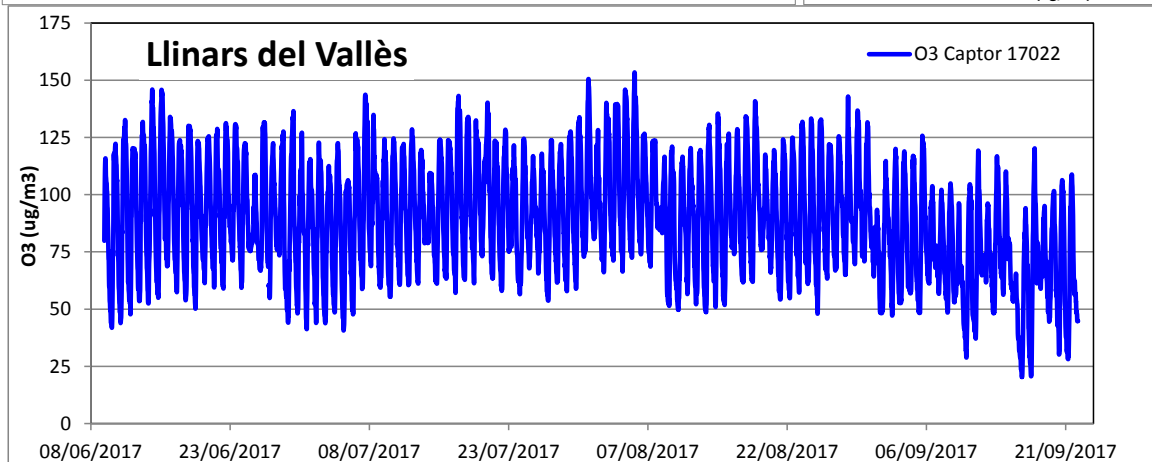
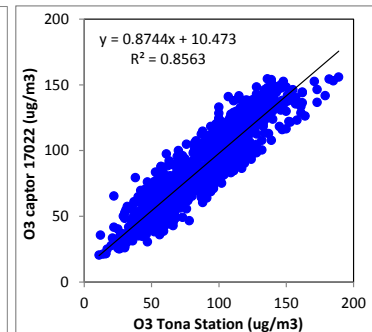
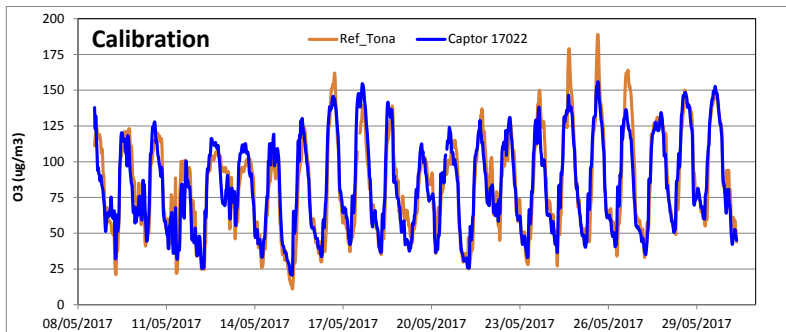
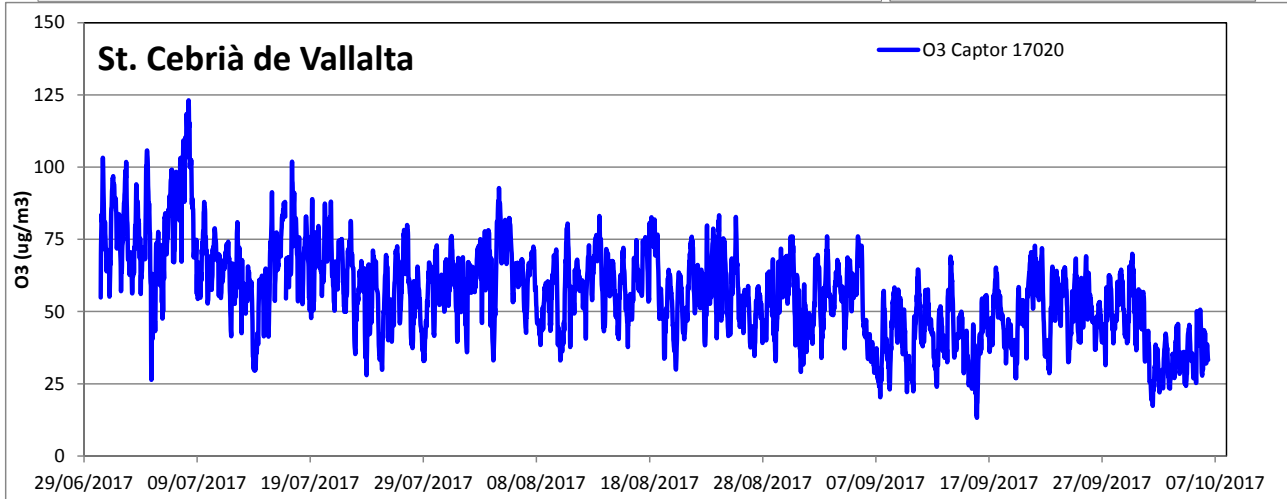
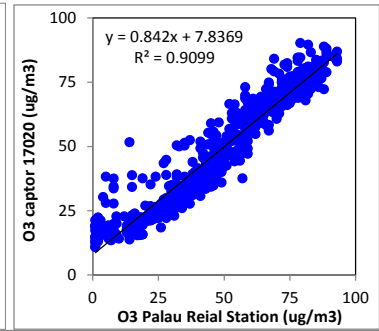
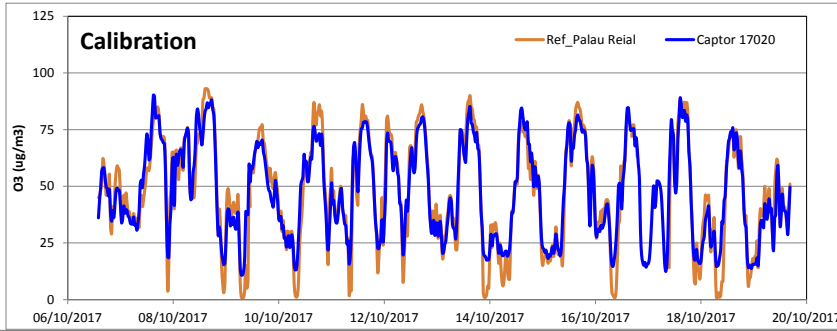


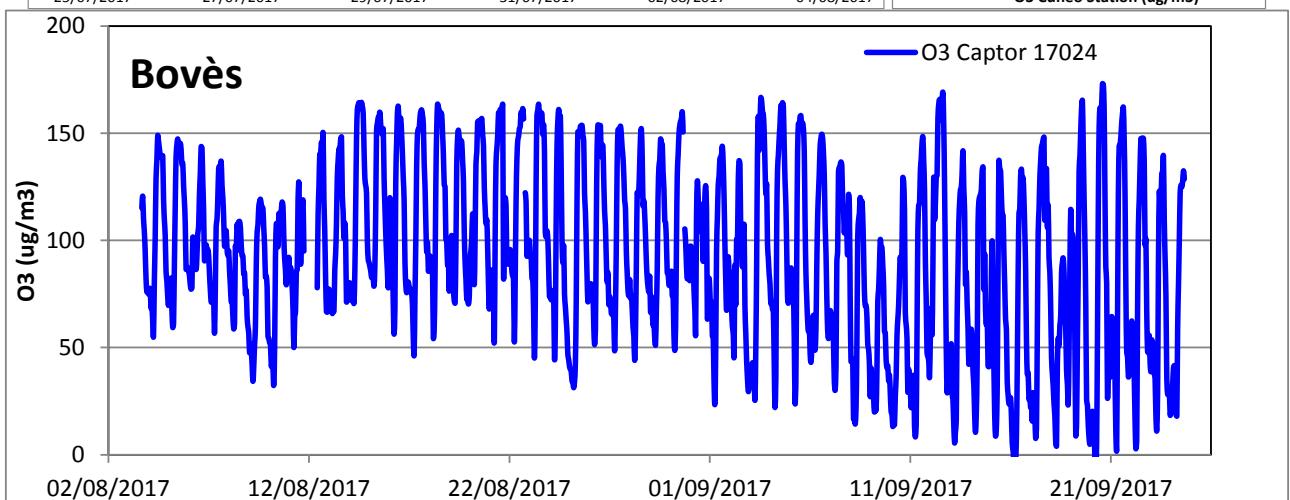
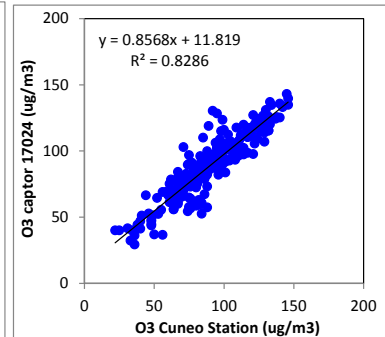
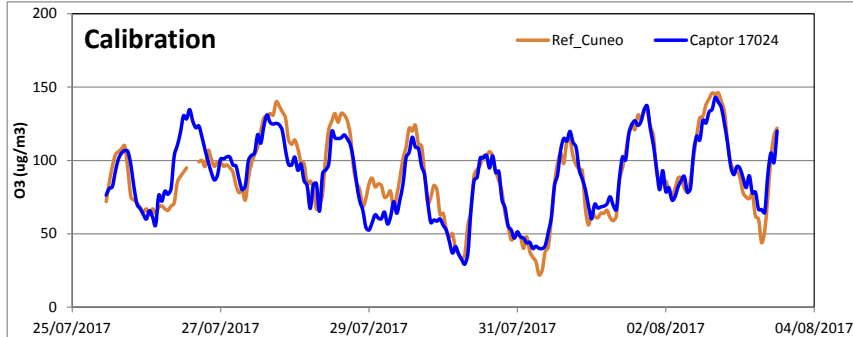
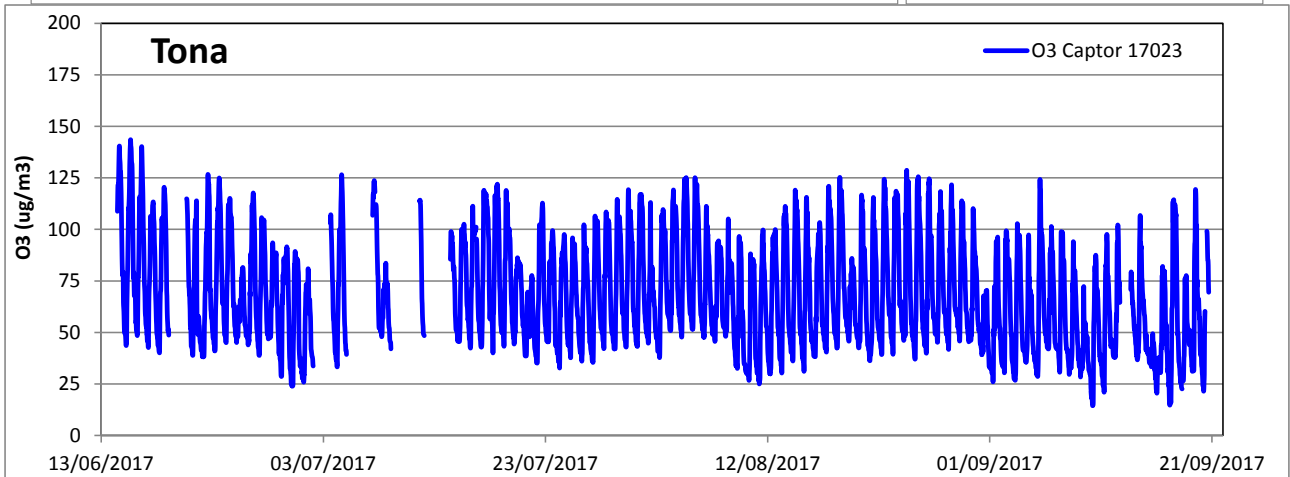
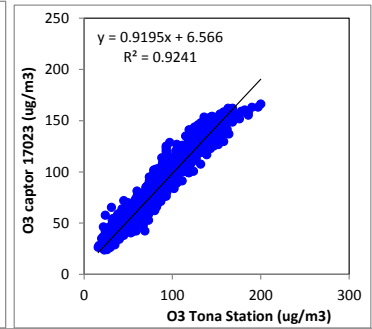
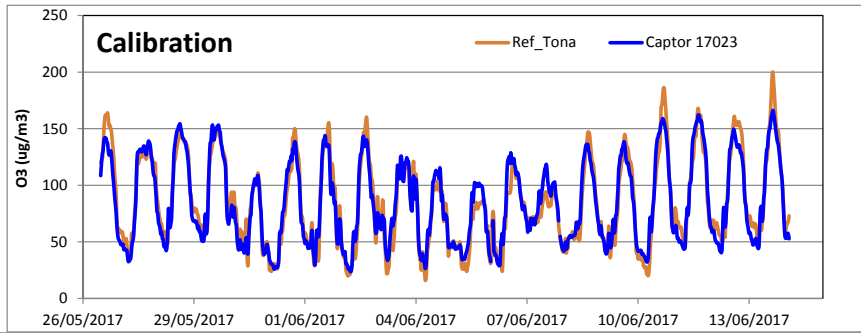


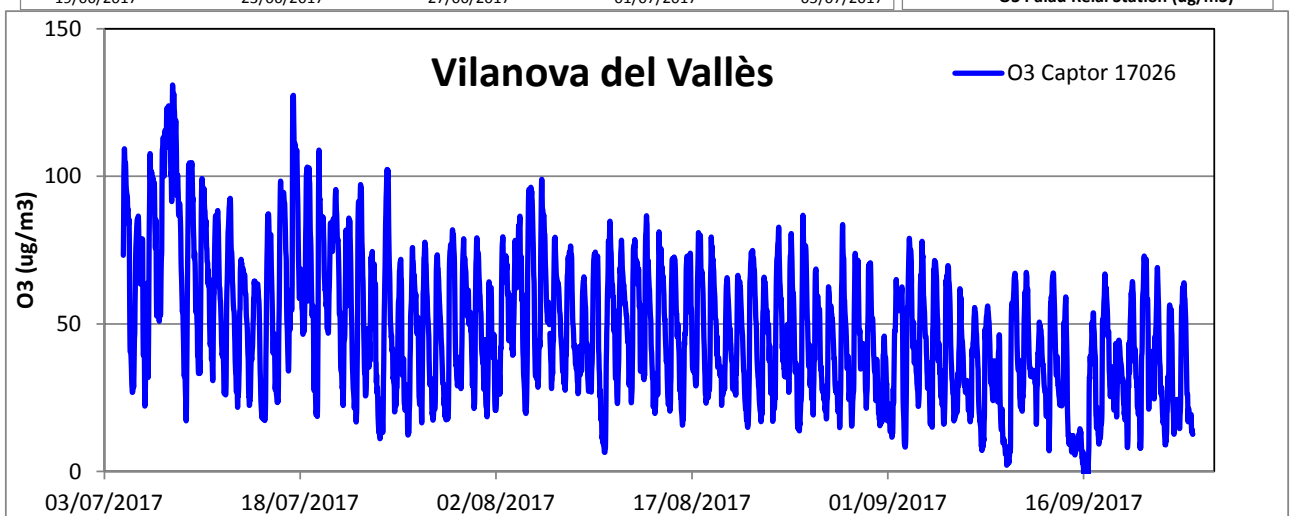
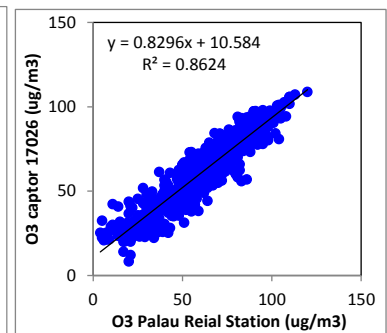
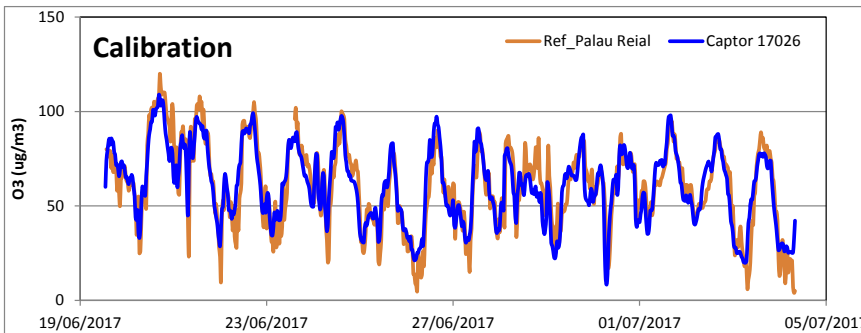
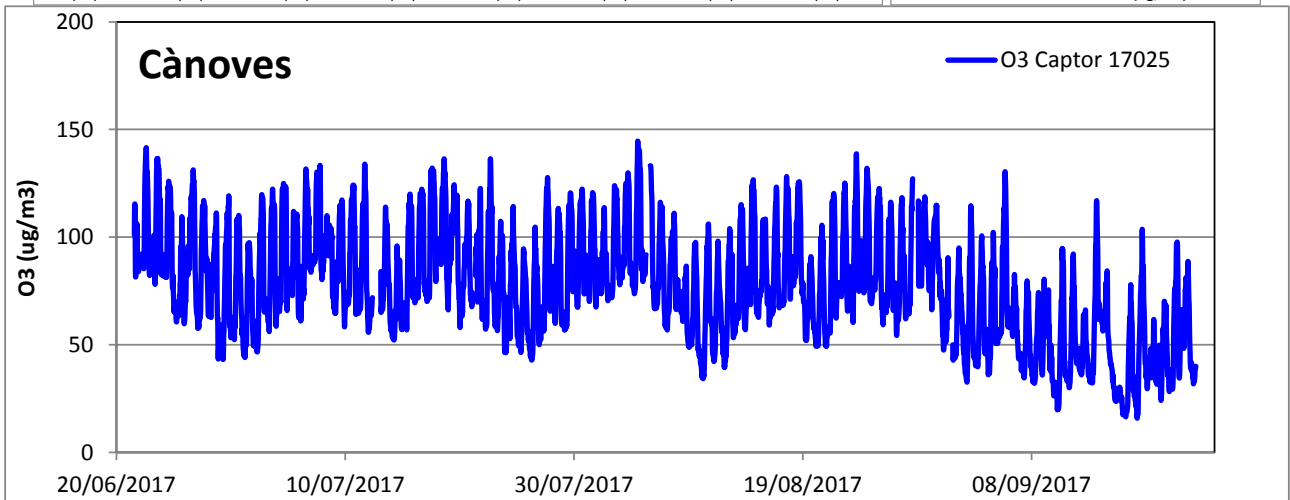
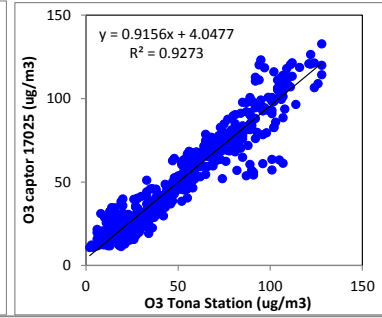
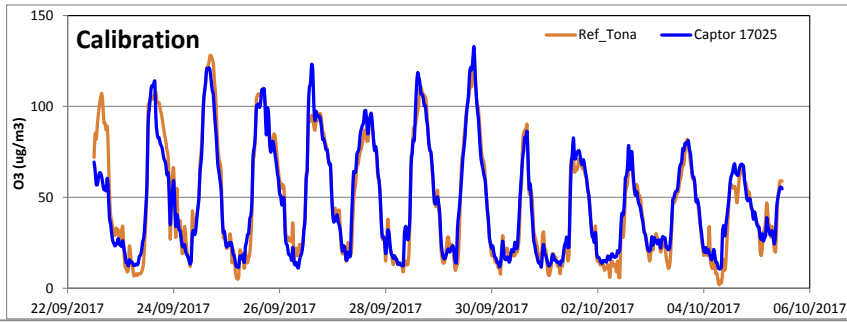
CAPTOR

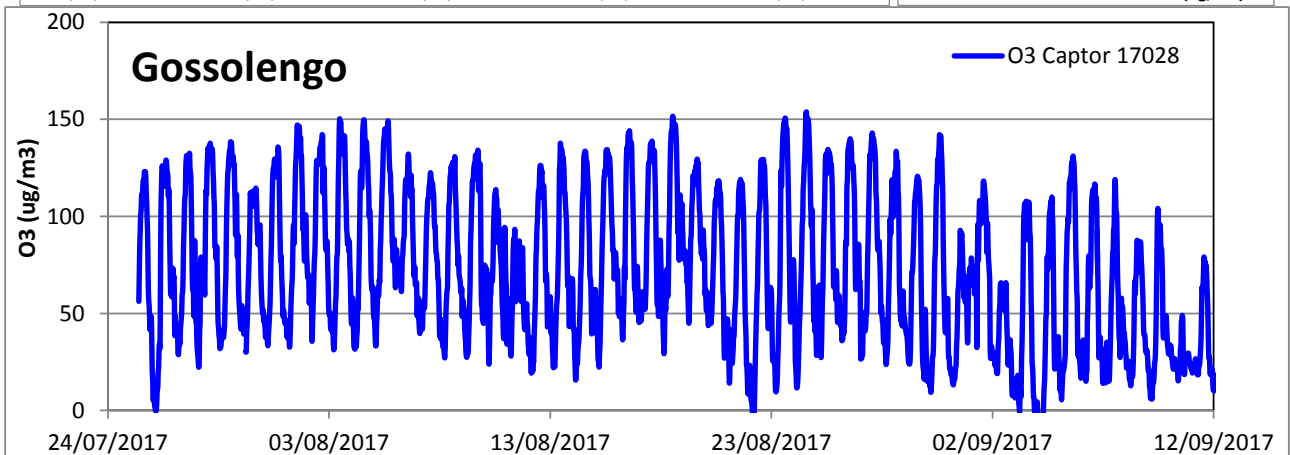
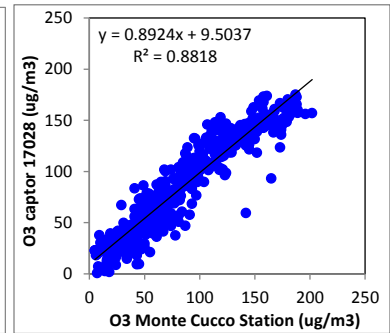
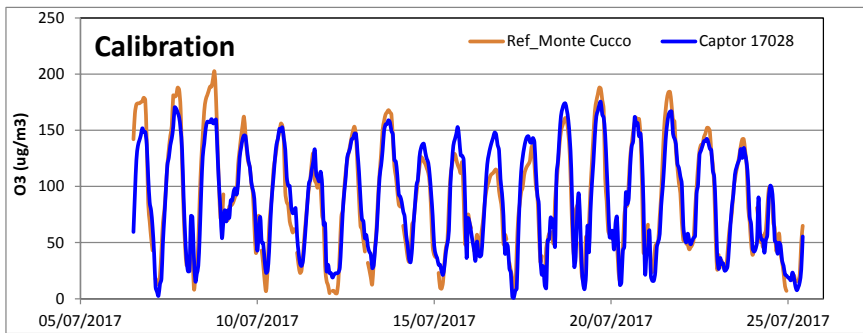
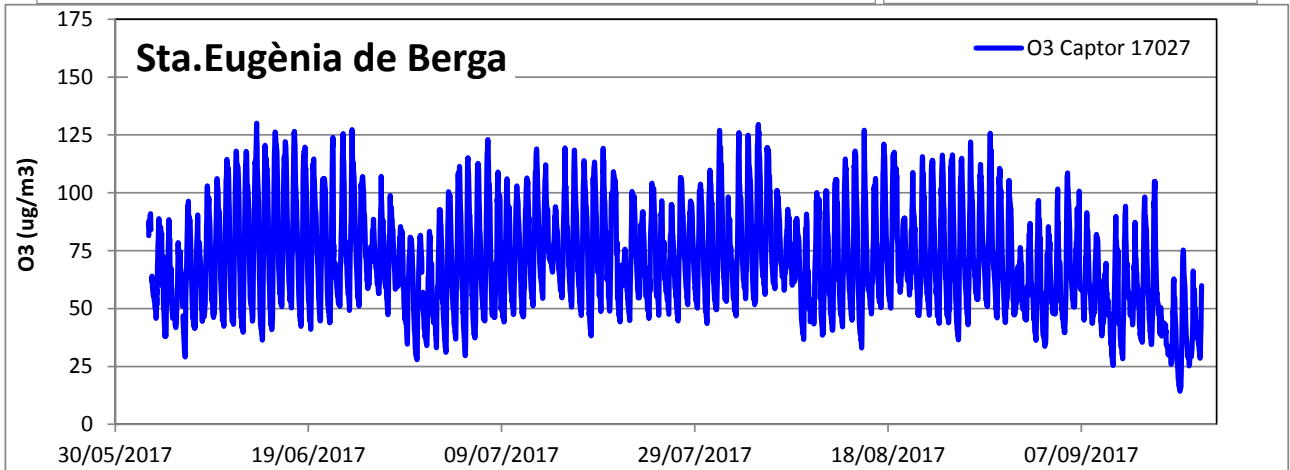
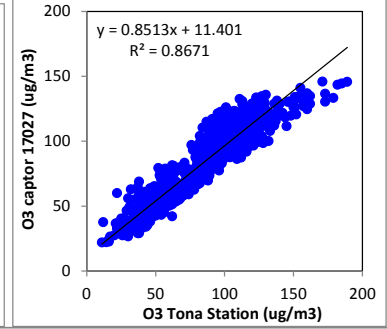
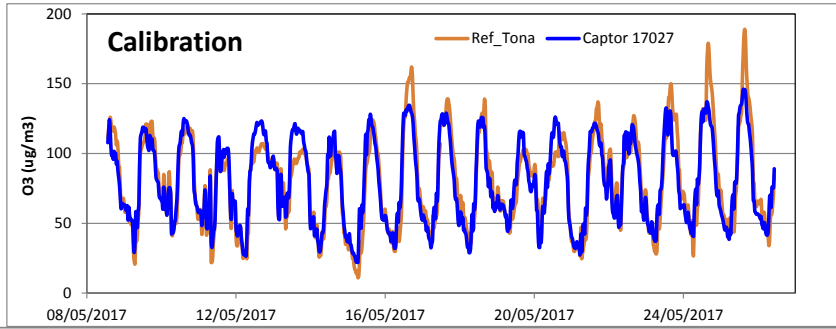


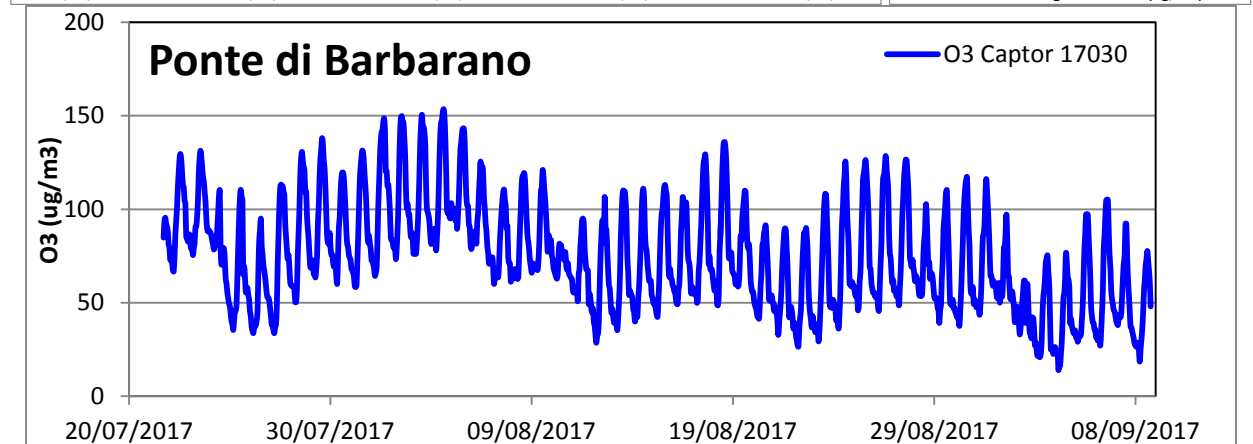
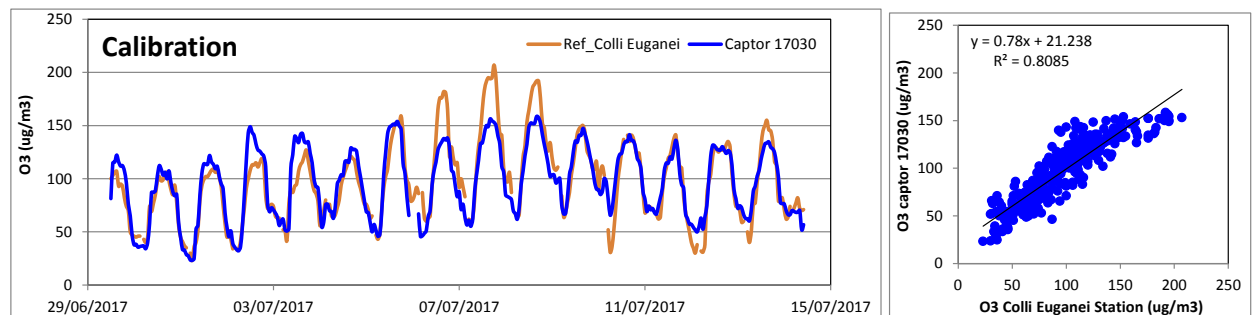
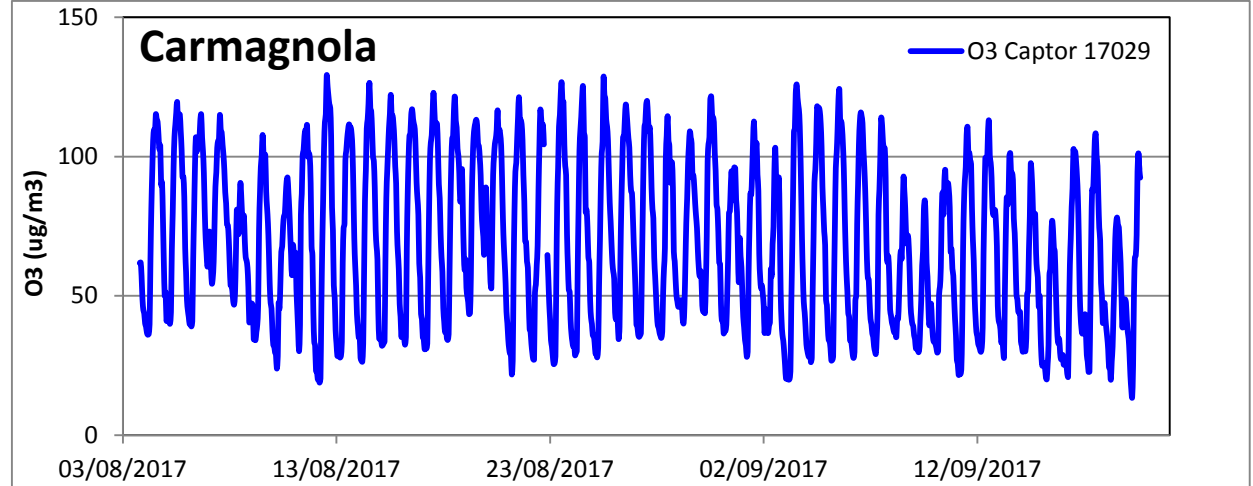
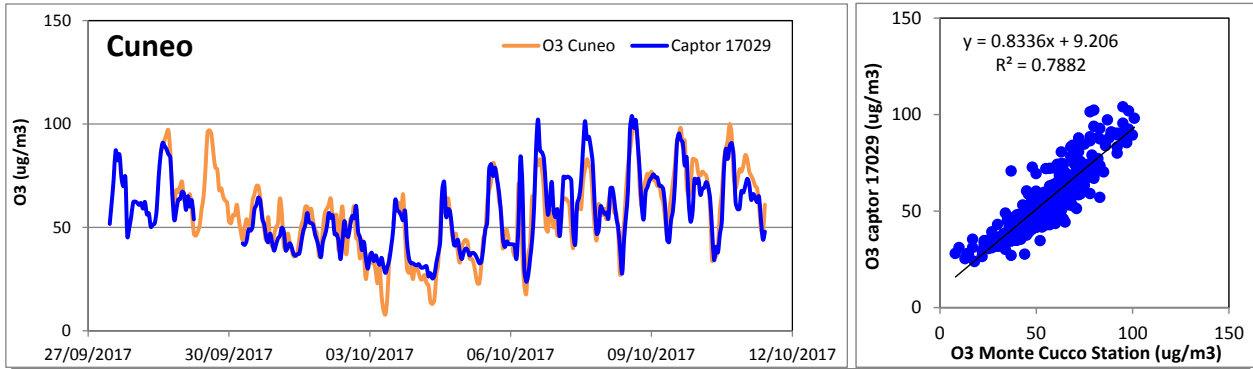


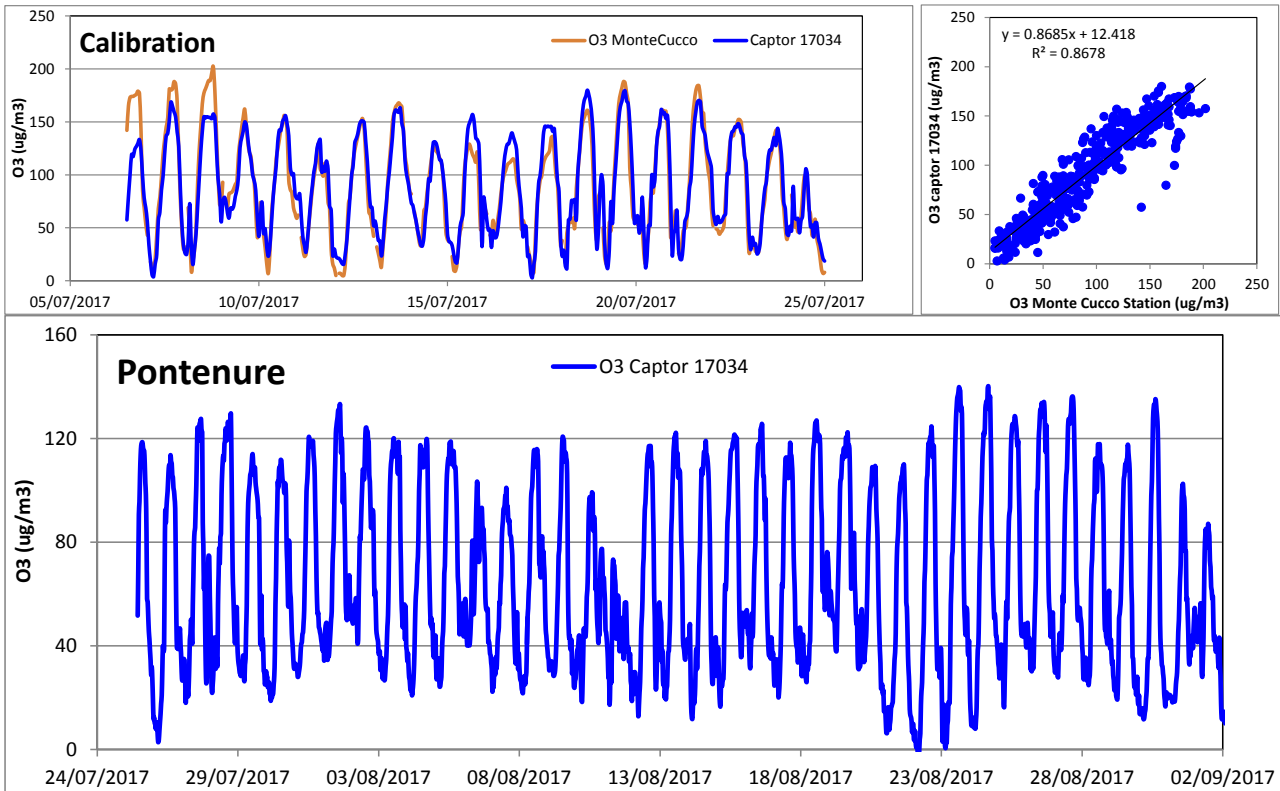












RAPTORS:

