

Milestone M29 (M5.1)

Overview of measurement
data and pilots, recommendations for data curation



RI-URBANS

**Research Infrastructures Services Reinforcing Air
Quality Monitoring Capacities in European Urban &
Industrial Areas (GA n. 101036245)**

By

NILU, TROPOS, INERIS & IAGOS



24th November 2022

Milestone M29 (M5.1): Overview of measurement data and pilots, recommendations for data curation

Authors: Markus Fiebig, Yong Lin, Lise Eder Murberg, Richard Olav Rud, Wenche Aas, Cathrine Lund Myhre (all NILU)

| | |
|------------------------------------|--|
| Work package (WP) | WP5: Strategic guidance for upscaling RI-URBANS STs |
| Milestone | M29 (M5.1) |
| Lead beneficiary | NILU |
| Means of verification | Workshop held |
| Estimated delivery deadline | M14 (30/11/2022) |
| Actual delivery deadline | 24/11/2022 |
| Version | Final |
| Reviewed by | WP5 leaders |
| Accepted by | RI-URBANS Project Coordination Team |
| Comments | Report describing the content and major outputs of the workshop to implement the RI-URBANS' Data Management Plan (DMP) by building capacity on the data provider end. A data manager knows which and how information is to be archived alongside the data as metadata, but the data provider holds the information content to be archived. Thus, capacity building for data providers is an essential part of any data management concept. |

Table of Contents

| | |
|--|----------|
| 1. ABOUT THIS DOCUMENT | 1 |
| 2. WORKSHOP AGENDA..... | 1 |
| 3. WORKSHOP MINUTES | 2 |
| 4. ACTION ITEMS | 5 |
| 5. APPENDIX I: DATASETS EXPECTED IN RI-URBANS | 6 |
| 6. APPENDIX II: LINKS TO PRESENTATIONS | 6 |

1. About this document

This RI-URBANS project milestone documents a virtual workshop dedicated to implementing the project's data management plan (DMP). Apart from other purposes, the DMP lists the datasets expected to be collected or generated in the project, and describes how and by whom these will be handled to achieve [FAIR data management](#) principles. The workshop described here implements the DMP by building capacity on the data provider end. A data manager knows which and how information is to be archived alongside the data as metadata, but the data provider holds the information content to be archived. Thus, capacity building for data providers is an essential part of any data management concept and to make data FAIR with the required traceability and documentation.

This is a public document, available in the RI-URBANS website (<https://riurbans.eu/work-package-5/#milestones-wp5>). The document will be distributed to all RI-URBANS partners for their use and submitted to European Commission as an RI-URBANS milestone M29 (M5.1).

2. Workshop agenda

9:30 – 9:40 Welcome

9:40 – 9:55 Data management in RI-URBANS (Richard Rud)

- The RI-URBANS data landscape and the DMP
- How we collect and archive data in RI-URBANS – Collaboration with ATMO-ACCESS

09:55 – 13:00 Curation of surface in situ data:

09:55 – 10:15 General principles (Lise Eder Murberg)

10:15 – 11:00 Aerosol particle phase:

- Data reporting for aerosol particle light absorption coefficient / equivalent black carbon, aerosol particle number concentration, aerosol particle number size distribution (Lise Eder Murberg)
- Data treatment of eBC data (Thomas Müller)
- Data treatment of aerosol particle number concentration / aerosol particle number size distribution data (Thomas Müller)
- Data reporting for online particle chemical composition (non-refractory) (Jean-Eudes Petit)
- Data reporting for offline particle chemical composition, elemental /organic carbon (Yong Lin)
- Discussion

11:00 – 11:30 Break

11:30 – 12:00 Gas phase: (Yong Lin)

- volatile organic compounds (VOCs)
- NH₃, O₃, CO₂, NO_x, SO₂
- Discussion

12:00 – 12:45 Break-out sessions: 1) particle phase; 2) gas phase, 3) working with data reporting in Jupyter notebook

12:45 – 13:00 Summary of discussions

13:00 – 13:30 Break

13:30 – 14:00 Data curation of ground-based aerosol remote sensing data (Lucia Mona)

14:00 – 14:30 Data curation of data from airborne platforms (Damien Boulanger)

3. Workshop Minutes

Number of registered participants: 30

Number of attended participants: 35

9:30 – 9:40 Welcome

- See presentation, no questions from audience.

9:40 – 9:55 Data management in RI-URBANS (Richard Rud)

- The RI-URBANS data landscape and the DMP
- How we collect and archive data in RI-URBANS – Collaboration with ATMO-ACCESS
 - Tuukka emphasizes the importance of WP5 collaboration between WP1 and WP4 and the importance of the task 5.1 and the data compilation, archiving and access. The RI-URBANS legacy is the data that will be provided by the collaborators after the end of the project. We need to be active to initiate the data flow and help the collaborators to develop their workflows to submit data also after RI-URBANS has ended.
 - Markus: Focus today is on data reporting and data management. Probably need more capacity building concerning data QC control and operating procedures.
 - Jelle Hofman: How do we handle data from mobile surface platforms such as bicycles? Example of eBC / aerosol particle absorption coefficient data collected with bicycles as platform. Do we store individual bike rides? Or aggregated maps of the data?
 - Markus: Data seems to have more demands than previously reported in the DMP, but absolutely possible to solve. Needs discussions inside data centre nodes. Send requests to the ATMO-ACCESS portal, will be handled from there.

09:55 – 13:00 Curation of surface in situ data:

09:55 – 10:15 General principles (Lise Eder Murberg)

- Markus: this covers data included in ATMO-ACCESS homeless data portal, being surface in situ data, and being routed to the ACTRIS In Situ DC node / EBAS database.
- Markus: EBAS offers traceable data reporting using data levels 0, 1, and 2, where 2 is the final data product. The levels document the production history of the data product. For established ACTRIS stations all levels must be submitted, but for new RI-URBANS stations we are focusing on level 2 data, also called quality assured data. RI-URBANS data will be associated to a RI-URBANS label in EBAS.
- Are there software tools to convert measurements to EBAS formats?
 - <https://git.nilu.no/ebas/ebas-io>

10:15 – 11:00 Aerosol particle phase:

- **Data reporting for aerosol particle light absorption coefficient / equivalent black carbon, aerosol particle number concentration, aerosol particle number size distribution (Lise Eder Murberg)**
- **Data treatment of eBC data (Thomas Müller)**
 - *Jelle Hofman: Will provide mobile aethalometers: AE51 (possibly MA200).*
 - *Markus: there is a need for separate WS on data QC, which will also cover the task of converting eBC to aerosol particle absorption coefficient for “non-standard” filter absorption photometers.*
 - *Angela: can we submit data with other AE models?*
 - *T60 Teflon-coated borosilicate glass fiber*
 - *mobile aethalometers: AE51 (possibly MA200)*
 - *Which correction factors to use, which type of filter medium is used, same as AE31*
 - *Need a follow up session with TROPOS on how to convert this from eBC to Aerosol Absorption Coefficient*
 - *Which task in WP4 to submit source apportionment data, Task 4.1 is mandatory to use AE33.*
 - *Marco Pandolfi: there is info for the MA200 Filter Material / Capacity MA200 Filter Tape Cartridge with Polytetrafluoroethylene (PTFE) material*
- **Data treatment of aerosol particle number concentration / aerosol particle number size distribution data (Alfred Wiedensohler)**
 - *Are the flags used for the full dataset? and if so, are datasets with different flags need to be submitted separately? All flags are submitted as a part of the NasaAmes1001 file in the variables (numflag), so you just submit one file. Furthermore, if a flag is part of only some of the measured times in the hourly averaged data, the flag is given to the whole hourly averaged data value for level 2 data.*
 - *Question about flags: Is it ok to use this flag: ‘256 – I – Invalidated by database coordinator’?*
Answer: This flag is taken out of the (short) flag list for the shown templates and should therefore not be used when reporting data.
- **Data reporting for online particle chemical composition (non-refractory) (Jean-Eudes Petit, Olivier Favez)**
 - *Question if Q-ACSM Export Tool is working on Linux platform?*
 - *Olivier does not think so.*
 - *If NRT data should be provided to EBAS, if yes, then how? If EBAS, likely should be done at the end of the pilot phase.*
 - *What is the R package for importing AE33 data?*
 - *Part of the real-time data stream, a part of the data centre and the topic centre.*
 - *Markus mentions the need a separate workshop on QA and submission of these ACSM data.*
- **Data reporting for offline particle chemical composition, elemental /organic carbon. (Yong Lin)**
- **Discussion.**

11:00 – 11:30 Break

11:30 – 12:00 Gas phase: (Yong Lin)

- **volatile organic compounds**
- **NH₃, O₃, CO₂, NO_x, SO₂**
- **Discussion**
 - *Work on defining PTR-MS template, including vocabulary, is ongoing, work currently in process*
 - *Will be sites with ammonia*
 - *Anne/Yong responds with a suggestion to look into templates for passive samplers.*
 - *IMT will submit on behalf of the sites that provided the data – data is already collected*

- *Look into filter-pack templates (there are already some NH3 data in EBAS using the filter pack template)*
- *Submit initial request in ATMO-ACCESS portal.*

12:00 – 12:45 Break-out sessions:

1) particle phase

- *data of sub-locations in an urban campaign. One instrument is moved between various hotspots in dedicated periods. Here, we should define separate location codes for each location to underline that these are in fact different locations.*
- *Question of national data reporting routines. Some urban sites / networks need to report their data to national authorities, and now in addition to RI-URBANS. Can this be streamlined, so data provider doesn't have to report data multiple times? Issue relates also to storing copies of data in different repositories, which are never in sync. Ideally, data should be stored in one primary archive, and streamed from there in its latest version. Topic needs to be taken up at higher level in WP5 since it is policy related.*

2) gas phase

- *Question of protection / access restriction to data: Is it possible to report data to RI-URBANS while still have it non-public? Data collected in RI-URBANS are publicly funded, and thus should be publicly available. Data don't belong to the scientist, but to the organisation (RPOs) where the scientist is working/ affiliated. RPOs are usually in line with the EU policy of making data public as much as possible. Data have attribution license and unique DOI, these take care of the attribution interests of the scientist. If absolutely needed, password can be used and a temporary embargo period can be setup.*
- *Defining reporting for PTR-MS instruments in progress. Work on defining variable names ongoing.*

3) working with data reporting in Jupyter notebooks

- *The repository was shown to the participants with information on RI-URBANS recommended workflow (recommended to use ordinary templates as a reference point), how to add metadata elements in the scripts and how to create, check and submit the files.*
- <https://git.nilu.no/ebas/ebas-io/-/blob/master/Examples/Doc/Notebooks/WriteNasaAmes.ipynb>
- *Question about the version available now, as the wiki contradicts the version available. This will hopefully be fixed asap. Update after the meeting: it is fixed.*

12:45 – 13:00 Summary of discussions

- *See above under the reports of each break-out group.*

13:00 – 13:30 Break

13:30 – 14:00 Data curation of ground-based aerosol remote sensing data (Lucia Mona)

- *Markus asks if ARES expect any lidar through RI-URBANS that do not follow Single Calculus Chain (SCC) Not now, but we can get through campaigns etc.*
- *Lucia responds that these instruments should also be reported through the ATMO-ACCESS data portal.*

14:00 – 14:30 Data curation of data from airborne platforms (Damien Boulanger, Valerie Thouret)

- *No specific comments in this session*

4. Action items

- Separate WS is needed on instrument operation and data QC, which will also cover the task of converting eBC to aerosol particle absorption coefficient for “non-standard” filter absorption photometers.
- Need for a separate workshop on instrument operation, as well as data QC and data reporting of ACSM data for RI-URBANS community.
- Take up issue of double data reporting to national networks / authorities on one side, and European networks on the other side, at RI-URBANS WP5 level. Are there policy options for co-ordinating data management between national and European level?

5. Appendix I: Datasets expected in RI-URBANS

https://nilu365-my.sharepoint.com/:x/g/personal/ror_nilu_no/Ef13wtJ8Zm1LmR0JVME9Ne4BTaOzk2pecCT_Bq4wssAsUQ

6. Appendix II: Links to presentations

Presentations can be accessed at this link.

<https://riurbans.eu/results/data-submission-workshop-presentations/>

1. Data management in RI-URBANS (Richard Rud)
 2. General principles for in situ data submission procedures (Lise Eder Murberg)
 3. Data reporting for aerosol particle light absorption coefficient (Lise Eder Murberg)
 4. Data treatment of eBC data and aerosol particle number concentration/ aerosol particle number size distribution data (Thomas Müller)
 5. Data reporting for online particle chemical composition (non-refractory) (Jean-Eudes Petit)
 6. Data reporting for offline particle chemical composition (elemental/organic carbon) and gas phase (Yong Lin)
 7. Data curation of ground-based aerosol remote sensing data (Lucia Mona)
 8. Data curation of data from airborne platforms (Damien Boulanger)
-
1. **An introduction to Data Management in RI-URBANS** (Richard Olav Rud)
 2. **General In-situ Data Submission Procedures** (Lise Eder Murberg)
 3. **Data treatment of eBC data** (Thomas Müller)
 4. **Data treatment of aerosol particle number concentration / aerosol particle number size distribution data** (Alfred Wiedensohler)
 5. **Data reporting for online particle chemical composition (non-refractory)** (Jean-Eudes Petit, Olivier Favez)
 6. **In-situ Data Submission Procedures – EC/OC & Trace gases** (Yong Lin)
 7. **Data curation of ground-based aerosol remote sensing data** (Lucia Mona)
 8. **Data curation of data from airborne platforms** (Damien Boulanger, Valerie Thouret)