Deliverable D3.16: Implementation report on increasing number of NA3 stations and instruments providing data in near-real-time

Markus Fiebig[1], Paul Eckhardt[1], Nikolaos Kalivitis[2], Nikolaos Michalopoulos[2], Cathrine Lund Myhre[1], Richard Olav Rud[1], Thomas Tuch[3], Alfred Wiedensohler[3]

[1]NILU – Norsk Institutt for Luftforskning, Kjeller, Norway
[2]University of Crete, Crete, Greece
[3]TROPOS – Institute for Tropospheric Research, Leipzig, Germany

Work package no	**WP3**
Deliverable no.	**D3.16: Implementation report on increasing number of NA3 stations and instruments providing data in near-real-time**
Lead beneficiary	**NILU**
Deliverable type	☑ R (Document, report) ☐ DEC (Websites, patent fillings, videos, etc.) ☐ OTHER: please specify ....
Dissemination level	☑ PU (public) ☐ CO (confidential, only for members of the Consortium, incl Commission)
Estimated delivery date	**M44**
Actual delivery date	***31/01/2019***
Version	**1**
Comments	

ACTRIS (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRAIA-2014-2015, Grant Agreement number: 654109

Table of Contents

ACTRIS (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRAIA-2014-2015, Grant Agreement number: 654109

Implementation Report on Increasing Number of NA3 Stations and Instruments Providing Data in Near-Real-Time

Abstract

Through its predecessor projects, ACTRIS has a 10 year history of providing observations of advanced atmospheric parameters from in situ surface stations in real-real-time (RRT, delay < 3 h, previously called near-real-time), focussing on observations of atmospheric aerosol particles. ACTRIS in situ RRT data provision is achieved through a centralised workflow, ensuring a homogeneous data product with efficient and cost effective use of resources. With a central infrastructure for provision of atmospheric aerosol RRT data already in place, the critical task in increasing the number of RRT stations and instruments for these parameters is to make participation as easy as possible for stations and to offer services connected to RRT data serving as incentive for participation. Providing RRT data includes provision of annotated raw (level 0) data from an instrument with hourly frequency to the data centre. To facilitate this data provision, ACTRIS WP3 provides or supports software packages for shared use at stations completing this task. Due to these measures, the fraction of potential RRT aerosol instruments being operated at sites contributing to ACTRIS in situ data that provide RRT data has increased from 14% to 55% during ACTRIS-2. ACTRIS services connected to RRT data include a station/model data comparison and validation tool, a RRT data showcase, and a closure tool for advanced quality control. In the future, RRT data provision will be the default for online ACTRIS surface in situ instruments.

ACTRIS (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRAIA-2014-2015, Grant Agreement number: 654109

1. Motivation and background

As one of several responses to the 2008 economic crisis, the European Commission published in2010 its “Digital Agenda for Europe”, intending to create economic growth by establishing a digital single marked across Europe. Building on this strategy, and also to connect space-borne with surface-borne Earth observation, the Copernicus programme, and here specifically the Copernicus Atmospheric Services (CAMS), implement user-oriented services based on data from satellite and in situ observations of the Earth, for the atmosphere and relating to ACTRIS most notably services on air quality and climate forcing. Operational services normally require their input and validation data to be available in real-real time (RRT, delay after observation < 3 h).

ACTRIS has a long history of providing data from its stations in RRT. This history dates back to its predecessor projects, the EU-FP5 project CLOUDnet and the EU-FP6 project *European Supersites for Atmospheric Aerosol Research* (EUSAAR). CLOUDnet set up a RRT data production chain for cloud properties derived from cloud radar observations, whereas EUSAAR implemented a demonstrator for RRT provision of atmospheric aerosol properties (particle scattering and absorption coefficients, particle number size distribution) measured at a number of surface in situ sites. This ACTRIS-2 WP3 deliverable reports on the continuation and extension of the efforts on atmospheric aerosol RRT data provision measured at ACTRIS surface in situ sites as started in EUSAAR.

2. Architecture of the ACTRIS in situ RRT data production workflow

The ACTRIS surface in situ RRT data production workflow follows a centralised architecture. The online instruments operating at the sites contributing to ACTRIS surface in situ data provide a raw data stream. This raw data stream is annotated with discovery and use metadata by the data acquisition system maintained by the station operator, generating level 0 data submissions at the turn of each hour covering the past hour. Immediately after generation, these level 0 data submissions are transferred to the in situ data centre (DC) unit by FTP. From the incoming level 0 RRT data submissions, the in situ DC unit produces data levels 1b (including calibrations and basic, automatic quality control) and 2b (final RRT data product, for surface in situ data usually hourly averaged, including measure of atmospheric variability). The “b” in the data level designation indicates data originating from the RRT data production workflow having seen automatic quality control. Level 2b has formerly been called level 1.5.

The centralised RRT data production workflow has been chosen against two alternatives:

ACTRIS (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRAIA-2014-2015, Grant Agreement number: 654109

- **De-centralised data production**: Here, data production is located not centrally in the DC, but de-centrally at the stations. RRT data products can be disseminated directly from the station, or centrally from the DC. To ensure a homogeneous, comparable data product in this architecture, it needs to be ensured that all de-central data production instances use identical software at any given time, and that all processing is done is the same comparable way.
- **Data production in the cloud**: Level 0 data are inserted into a cloud network immediately after production. De-central data production nodes collect them from the cloud, produce levels 1b and 2b, and re-insert level 2b into the cloud. In this architecture, data often have a limited lifetime in the cloud. For a homogeneous data product, all data production nodes need to run identical software at any given time. Avoiding a single point of failure by having several data production nodes is often mentioned as key advantage of data production in the cloud. It needs to be assured that all metadata parts, discovery and use metadata, survive the passage through the cloud.

These aspects were essential for weighing the different options for the RRT architecture:

- **Homogeneous data production**: For de-central data production, it needs to be assured that all data production nodes, either at stations or in the cloud, run identical production software versions at any given time. This requires an advanced software management system that is resource demanding in both implementation and maintenance. In a centralised data production architecture, implementing a homogeneous data production, even if several instances are deployed, is straight forward.
- **Resource efficiency**: Data production in the cloud excels for use cases with huge data amounts, distributed data access, and huge access volumes. However, none of these three aspects apply to ACTRIS. In turn, cloud data production architectures come at the added resource expense of ensuring correct routing, as well as ensuring data and metadata integrity while passing through the cloud. RRT data production in the cloud requires cloud computing resources with very high uptime requirements, usually by private providers, which are costly to procure. In practice, the cloud architecture selling point of avoiding single points of failure has little relevance since uptime can be assured by other means.
- **Documenting provenance**: In data management, provenance comprises the production history of data, aspects such as “which source data have been used?”, “which software in which version has been used to generate the data?”, “which hardware in which location was used?”, etc. Provenance information makes data production transparent and traceable, and facilitates tracking of errors. ACTRIS data production will document provenance in the future,

ACTRIS (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRAIA-2014-2015, Grant Agreement number: 654109

as required by the ENVRI Reference Model for European environmental research infrastructures. While it is possible to implement provenance documentation in a de-centralised data production workflow, the resource requirements for implementing and maintaining this service are considerably lower in a centralised architecture.

3. Strategy towards increasing number of ACTRIS aerosol in situ RRT instruments

Once the central infrastructure for RRT data production is set up, which for ACTRIS in situ data was already accomplished during the EUSAAR project, the essential task remaining is to facilitate surface in situ stations to generate and deliver hourly submissions of annotated raw data, i.e. level 0 data, to the DC from the participating instruments. Normally, it is the data acquisition system at the station responsible for collecting data from the instrument that conducts this task. The station needs to commit to maintaining RRT submissions.

In order to minimise the effort involved for stations to participate in RRT data production, ACTRIS WP3 decided to provide or support software packages for data acquisition at the stations that not only store the data locally, but also perform the tasks of annotating the data with discovery and use metadata, putting them into the correct format, and uploading them to the DC with hourly frequency. Table 1 lists these software packages, the author or provider, and their status of development.

Table 1: List of shared instrument data acquisition software packages facilitating RRT data reporting in ACTRIS.

Instrument	Solutions by:	Status
TSI 3563 nephelometer	NILU NOAA[1]	Ready Ready
ECOTECH nephelometer	U. Crete NOAA[1]	Testing Ready
Thermo Mult-Angle Absorption Photometer (MAAP)	TROPOS NOAA[1]	Ready Ready
Magee Aethalometer AE31	NILU NOAA[1]	Ready Ready
Magee Aethalometer AE33	NILU	Testing
Radiance Ressearch PSAP	NILU NOAA[1]	Testing Ready
TROPOS Mobility Particle Size Spectrometer (MPSS)	TROPOS	Ready

[1] The NOAA software package is provided by the Aerosol Group of the U.S. NOAA Earth System Research Lab. Global Monitoring Division.

ACTRIS (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRAIA-2014-2015, Grant Agreement number: 654109

All software packages with “ready” status are available from the ACTRIS in situ DC web pages for download or by reference.[2]

4. Increase of ACTRIS aerosol in situ RRT instruments in numbers

For assessing the increase in the number of stations and instrument participating in the ACTRIS in situ RRT data production, the numbers at the start of ACTRIS-2 are taken as reference. From the predecessor projects EUSAAR and ACTRIS-FP7, there were 7 stations reporting RRT data to the ACTRIS DC in situ node from in total 10 instruments. Considering the numbers of potentially contributing stations (29) and instruments (71), these numbers amounted to a relative contribution of 24% in the number of stations and 14% in the number of instruments (see Figure 1).

ACTRIS-2 start (summer 2015):
7 stations, 10 instruments

Autum 2018:
21 stations, 39 instruments

Potential NRT system:
29 stations, 71 instruments

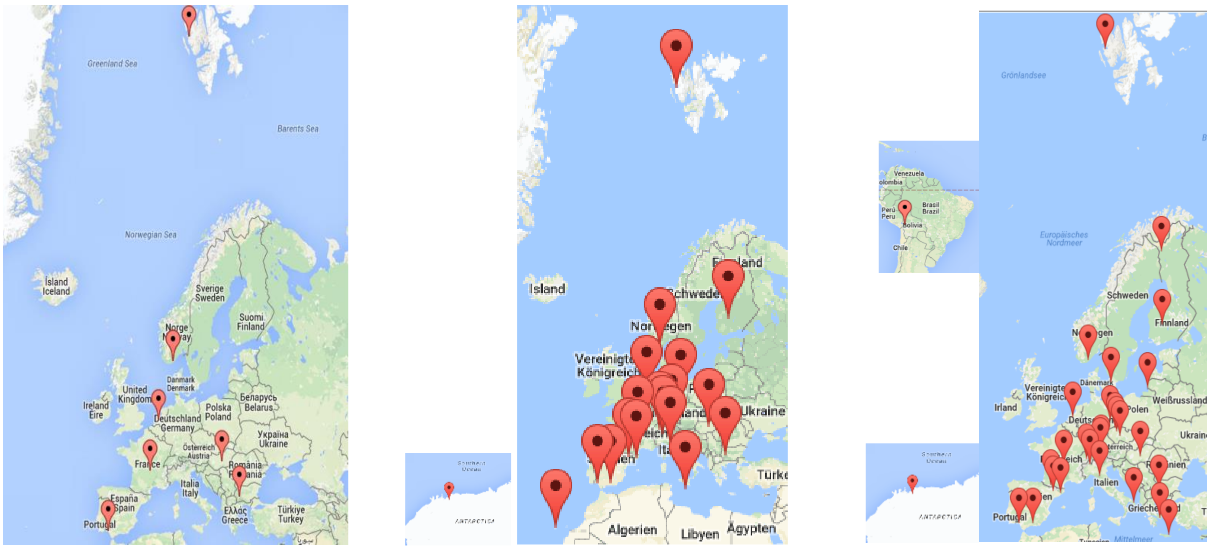

Figure 1: Maps of sites contributing to ACTRIS aerosol in situ data reporting RRT data, at the start of the ACTRIS-2 project (left), towards the end of the project (middle), and the potential extension as counted by the total number of instruments of types currently supported by ACTRIS in situ RRT.

Towards the end of ACTRIS-2 in autumn 2018, these numbers had increased to 21 participating stations contributing with 39 instruments reporting RRT data, corresponding to 72% in the number of stations and 55% in the number of instruments. This significant increase in the number of ACTRIS in situ RRT stations and instruments is also illustrated by Figure 2, which displays the numbers of participating

[2] https://ebas-submit.nilu.no/Near-real-time-data-submissions
https://ebas-submit.nilu.no/software-tools.

ACTRIS (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRAIA-2014-2015, Grant Agreement number: 654109

stations and instruments over time. The number of RRT stations has reached a preliminary saturation level, whereas the number of RRT instruments continues to increase.

The increase in sites reporting aerosol in situ RRT data delivery by ACTRIS WP3 can be considered a success story. The number of RRT ACTRIS in situ instruments is expected to increase further as RRT data reporting will be compulsory in ACTRIS in the future.

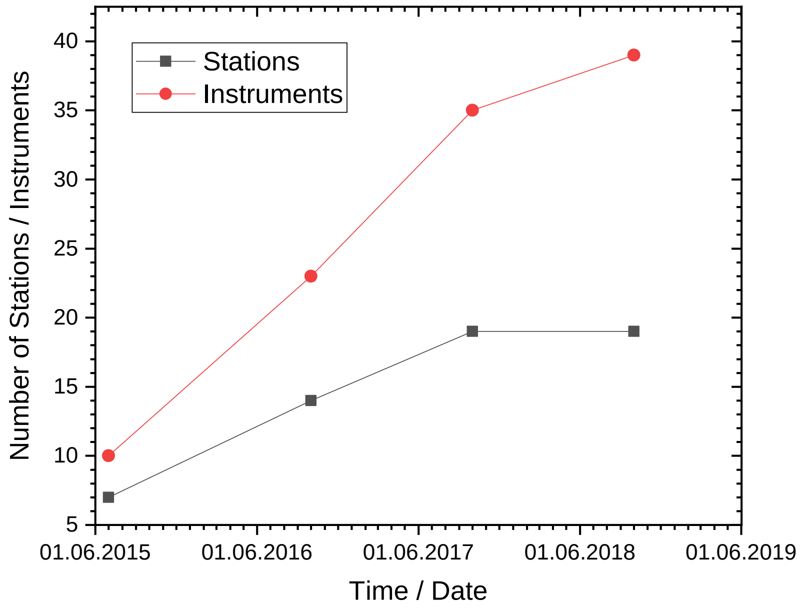

Figure 2: Number of sites contributing to ACTRIS in situ data and instruments delivering RRT data as a function of time and parameter.

Another likely reason for the increase in ACTRIS WP3 station and instrument numbers delivering data in RRT are the services connected to RRT data that benefit data providers.

5. Services based on ACTRIS aerosol in situ RRT data

According to the Digital Agenda of the European Commission, openly and immediately available data will contribute to stimulating economic growth and to providing new and previously impossible services to the public. This applies also at a smaller scale to scientific users of research data. ACTRIS-2, also beyond WP3, has provided several examples of such research services involving RRT data, which are summarised here.

1. **Service for comparing RRT data to corresponding output of a forecast model.**
   In ACTRIS JRA3, ECMWF, together with the ACTRIS DC in situ node, has worked on a demonstrator for a website comparing RRT observations of atmospheric aerosol optical properties with the same parameters calculated by the development version of the ECMWF

ACTRIS (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRAIA-2014-2015, Grant Agreement number: 654109

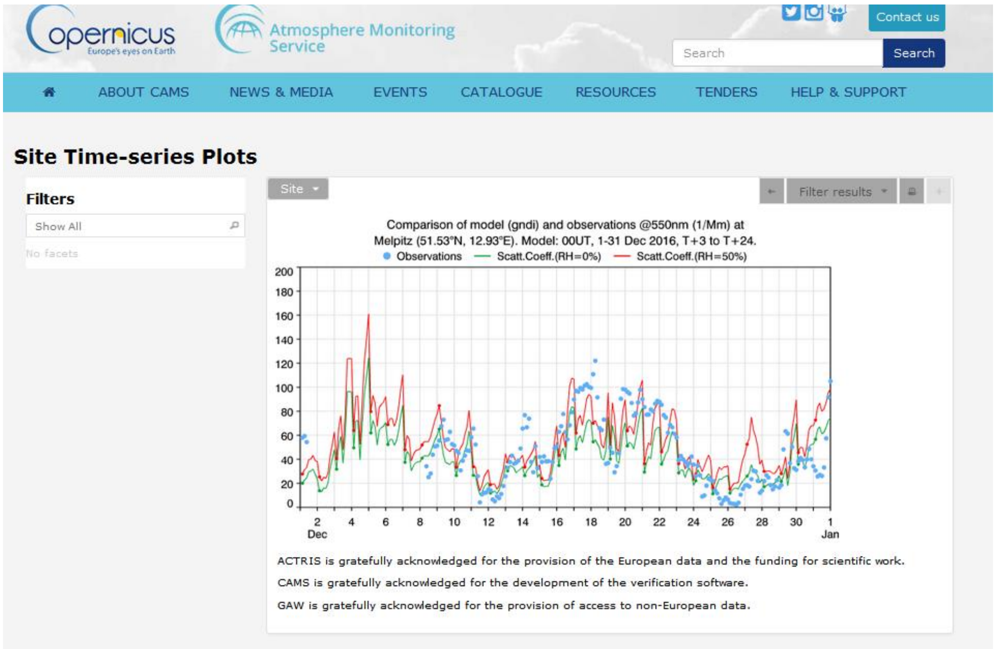

Figure 3: Demonstrator of service comparing optical aerosol parameters calculated by the development version of the ECMWF forecast model with observations at ACTRIS stations reported in RRT.

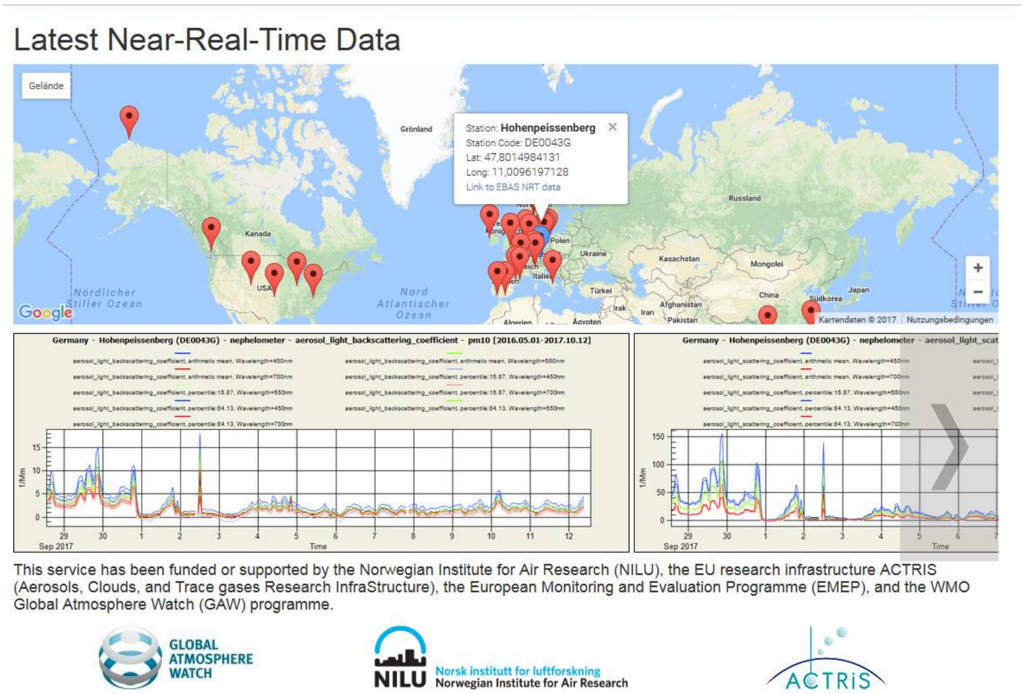

Figure 4: Showcase displaying data collected by the ACTRIS Data Centre in RRT from ACTRIS stations and stations part of partner networks.

ACTRIS (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRAIA-2014-2015, Grant Agreement number: 654109

forecast model (Figure 3, Letertre-Danczak et al., 2017). Such a service can be used both for model validation and for interpreting episodes occurring in observations, also for assessing instrument performance.

2. **Showcase webpage displaying ACTRIS RRT data.**

   The ACTRIS DC has set up a webpage giving a simple display of RRT data available from ACTRIS and collaborating frameworks such as WMO Global Atmosphere Watch (GAW) (see Figure 4), also beyond aerosol properties. The page contains a map of participating RRT stations, where each station marker links to a series of time-series charts visualising the latest available data. The page is designed to be included a frame into other websites. The links to the continuously updated time-series charts are permanent to give stations the opportunity to include them in their own displays of their RRT data.

3. **Closure service for advanced quality control.**

   In this context, closure means the comparison of one parameter obtained by 2 independent means, here by 2 independent observations. In the implemented example, which is also reported in ACTRIS-2 deliverable D3.13 (Rud et al., 2017), the aerosol particle scattering coefficient from a direct measurement is compared with the corresponding number calculated from the observed particle number size distribution, where the calculation makes an assumption on the particle refractive index. The 2 results, both obtained from RRT data, are compared as time series strip charts and scatter plots. While factors like aerosol type can influence the closure result, it is still a sensitive measure of instrument performance, and can facilitate early detection of instrument performance issues (example see Figure 5).

6. Outlook: Further extension of ACTRIS RRT data production

The increase of participation in ACTRIS in situ RRT data delivery from 14% to 55% measured as fraction of potentially available online aerosol instrumentation at ACTRIS surface in situ stations can be considered a success. This could be achieved even though participation was not compulsory in ACTRIS-2, but by providing comparatively easy means of participation and incentives in form of additional services connected to RRT data delivery. The network of ACTRIS RRT stations and instruments will be extended further in the upcoming unfunded period of ACTRIS to demonstrate the infrastructure's viability.

In the future when ACTRIS will be an operational research infrastructure, RRT data delivery will the compulsory default for all online surface in situ instruments connected to ACTRIS.

ACTRIS (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRAIA-2014-2015, Grant Agreement number: 654109

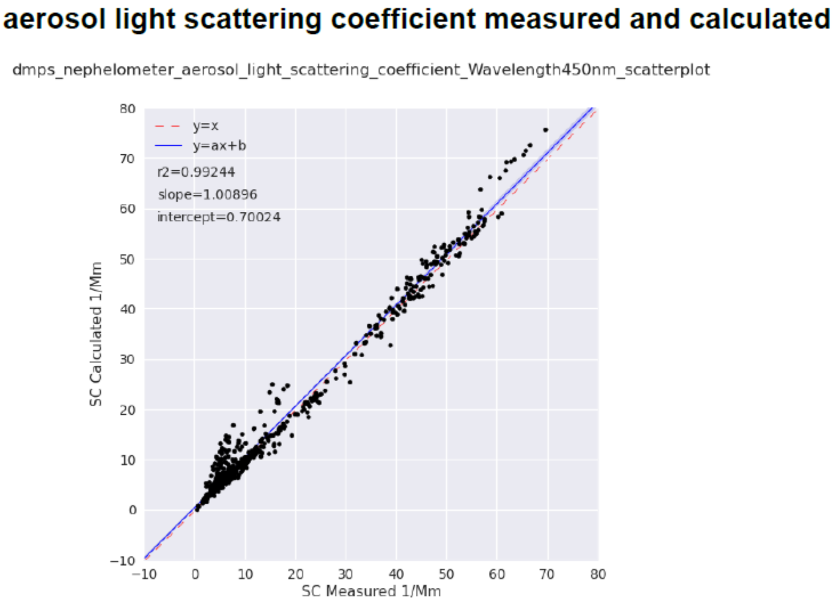

Figure 5: Example of QC product based on data collected from sites contributing to ACTRIS aerosol in situ RRT. The product performs a closure, i.e. a comparison of the same parameter obtained by independent means, here two independent observations. This serves as a sensitive tool to assess the performance of the instruments involved, and facilitates timely corrective measures if needed.

ACTRIS (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRAIA-2014-2015, Grant Agreement number: 654109

References

Julie Letertre-Danczak, Luke Jones, Cihan Sahin, Angela Benedetti, Enric Terradellas, Sara Basart, Enza Di Tomaso, Oriol Jorba, Augustin Mortier, Jan Griesfeller, Michael Schulz (2017): ACTRIS-2 Deliverable D13.2: Demonstration of online evaluation capabilities, https://www.actris.eu/Portals/46/Documentation/actris2/Deliverables/public/WP13_D13.2_M21_revised-M31.pdf?ver=2017-11-16-155729-977

Richard Olav Rud, Paul Eckhardt, Cathrine Lund Myhre, and Markus Fiebig (2017): ACTRIS-2 Deliverable D3.13: Implementation of closure studies for particle light scattering coefficient, https://www.actris.eu/Portals/46/Documentation/actris2/Deliverables/public/WP3_D3%2013_M28.pdf?ver=2017-10-10-155340-470

ACTRIS (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRAIA-2014-2015, Grant Agreement number: 654109