ACTRIS In Situ Aerosol: Guidelines for Manual QC of Ecotech Aurora 4000/3000 Integrating Nephelometer Data

Th. Müller, TROPOS

M. Fiebig, NILU - Norsk Institutt for Luftforskning

The manual data QC is performed on the level 0 data, i.e. the raw data output, augmented with essential discovery and use metadata, and brought to a standardised data format. For the purpose of manual QC, temporary level 1 and 2 data versions are produced from the initial level 0 version. When performing manual QC for a given time period of data, e.g. a year for an annual submission of data, the parameters contained in the level 0, 1, or 2 data listed below are to be plotted as time series, and visually inspected in intervals of maximum 2 weeks at a time. Data sequences exhibiting issues are to be flagged with an appropriate flag contained in this list:

Group 0: Valid data

Flag	Validity	Description
000	V	Valid measurement

Group 1: Exception flags for accepted, irregular data

Flag	Validity	Description
110	V	Episode data checked and accepted by data originator. Valid measurement

Group 3: Flags for aggregated datasets (used for level 1.5 & 2 only)

Flag	Validity	Description
390	V	Data completeness less than 50%
392	V	Data completeness less than 75%
394	V	Data completeness less than 90%

Group 5: Chemical problem

Flag	Validity	Description
559	V	Unspecified contamination or local influence, but considered valid

Group 6: Mechanical or instrumental problem

Flag	Validity	Description

640	V	Instrument internal relative humidity above 40%
686	I	Invalid due to zero check. Used for Level 0
687	I	Invalid due to span check. Used for Level 0

Group 9: Missing flags[MF1][MF2]

Flag	Validity	Description
999	M	Missing measurement, unspecified reason

Regardless in which data level the issue is found, the flags are added to the initial level 0 data version, thereby producing level 0a (manually QCed level 0) as output of the QC process.

The flags for aggregated datasets in group 3 apply only to levels 1.5 and 2. They indicate which fraction of the averaging period is covered by active sample time of the instrument.

The following parameters are to be inspected for the following issues:

Level 0:

1. **Periods of zero and span checks**
   If not done automatically by the data acquisition software, periods of zero and span checks are to be flagged with separate flags (flags 686 and 687, respectively).
2. **Sample pressure, sample inlet temperature, sample outlet temperature**
   Sample pressure varies with ambient pressure. Other types of variations should not occur, e.g. variations with fluctuating sample flow, apart from periods with zero and span checks. Sample temperature at inlet and outlet normally varies only with lab temperature, and during zero and span checks. Other variations and spikes need to be inspected, the reason determined, and flagged according to issue if needed.
3. **Sample relative humidity at inlet and outlet**
   Sample relative humidity varies with ambient relative humidity and the temperature difference between ambient and lab. The sample should be dried so that the sample has RH < 40% already at the instrument inlet. If RH is higher, apply flag 640. Spikes can occur during zero and span checks. Other variations and spikes need to be inspected, the reason determined, and flagged according to issue if needed. The Aurora has an inlet heater upstream of T and RH sensor. The heater must not be used for drying the aerosol because of the loss of volatile material.
4. **Sample flow**
   Sample flow, both through the main inlet and the instrument, should be constant, with small variations caused by wind gusts. Sample flow under normal operation should typically be at least 3 l/min, with drops by 20% max. Spikes can occur during zero and span checks. Other variations and spikes need to be inspected, the reason determined, and flagged according to issue if needed.
   It is recommended to use an external pump instead of the internal blower, because the blower may be too weak.
5. **Zero scattering calibrations**
   These depend only on wavelength, and very slightly on sample pressure. Thus, they should essentially be constant. Spikes or shifts need to be examined carefully. Spikes can occur due to erroneous measurements caused by remaining CO_2 after a span check

(data period until next zero is invalid). Shifts can be caused by a leaky zero air filter, which can potentially be corrected if the leak rate is constant.
Warning! The internal pump for generating zero air fills the chamber with zero air and the pump/blower is switched off. It is possible that zero air flows back into the inlet and affects other instrument.

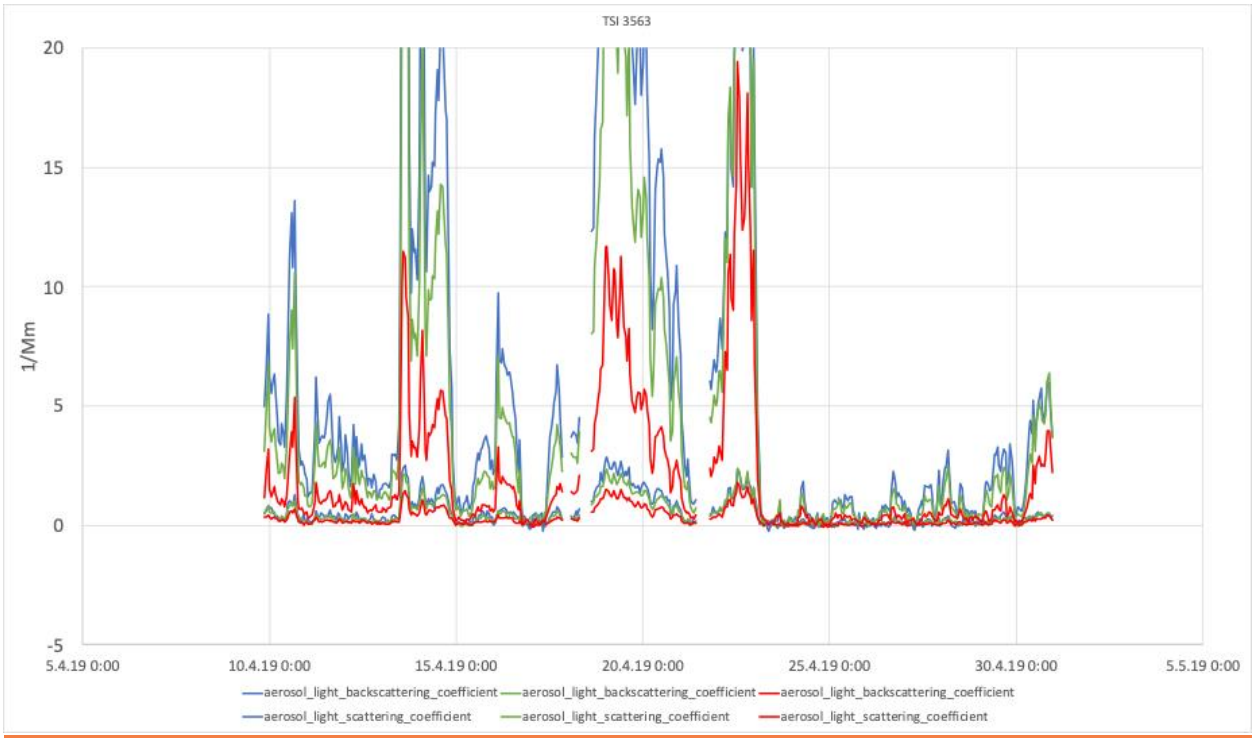
6. **Scattering and backscattering coefficients**
   These need to be inspected on 2 scales, logarithmic and linear. The (back)scattering coefficients change with air mass origin at the station. Thus, expected features in this time series are station dependent. Spikes are discovered on the logarithmic plot, and investigated using the station log and trajectory / backward plume analysis. Instrumental malfunctions are flagged with 999, local influence with 559, other episodes (dust, wildfires, long-range transport, ...) with 110. The linear plot displays data around or just below 0. At values near the detection limit, some noise around or below 0 is normal and expected. When averaging to hourly means, this noise cancels out to positive values around the detection limit. Negative spikes to values below the noise level are invalid and need to be removed (flag 999). These negative spikes are often caused by faulty zero measurements.
7. **Averaging filter**
   The Aurora can optionally smooth the time series of the scattering coefficients with a Kalman filter. This smoothing may lead to artificial jumps in the time series. These jumps occur independently for the three wavelengths and may be overlooked. For calibration (zero and span) the Kalman filter must be switched on, because otherwise the stability of the scattering coefficients is not sufficient and the calibration becomes faulty.
   The used averaging filter affects the effectiveness of the automatic zero measurement. The last value of the zero measurement is used as the current zero value for further measurements. Without Kalman filter this value is very uncertain due to the noise of the device. Solutions would be to perform the zero measurement without Kalman filter only as a zero test and the evaluation software has to consider this or the data logger switches on the Kalman filter for automatic zero measurements. For both variants there is currently no ready-to-use software solution.

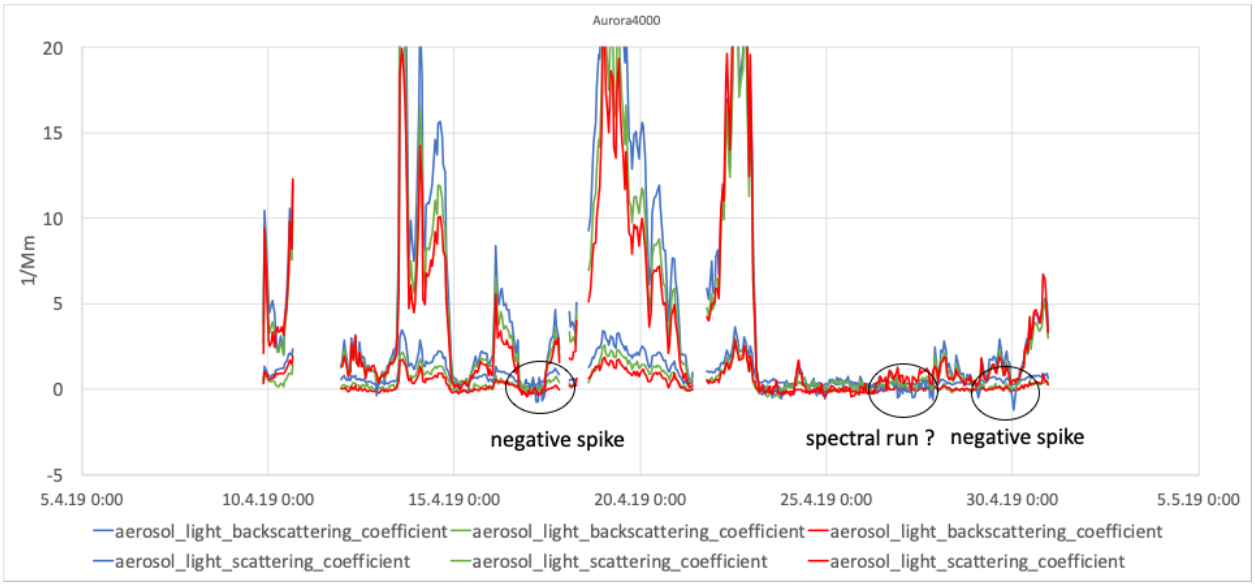
Level 2:

1. **Scattering and backscattering coefficients**
   These are plotted on a linear scale for the region around 0. If negative spikes have been flagged correctly in level 0, the level 2 (back)scattering data don’t exhibit negative values. If this is the case nevertheless, the flagging of the level 0 data needs to be revisited and corrected for the time sequence concerned. A new temporary level 2 data file is produced, and verified that the negative data have vanished.

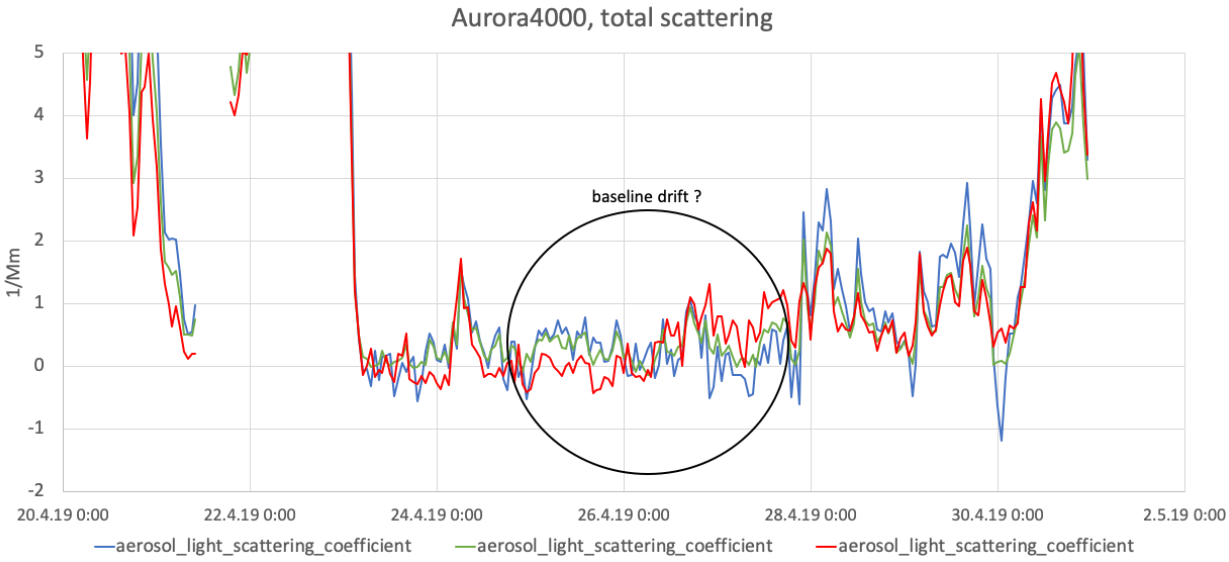
Appendix A: Examples of Conditions requiring flagging[MF3]

Example 1:





Comparison of scattering coefficients from TSI3563 and Aurora4000 nephelometers show good agreement. But in Aurora4000 data there are obviously negative spikes and one period with an unusual spectral run.

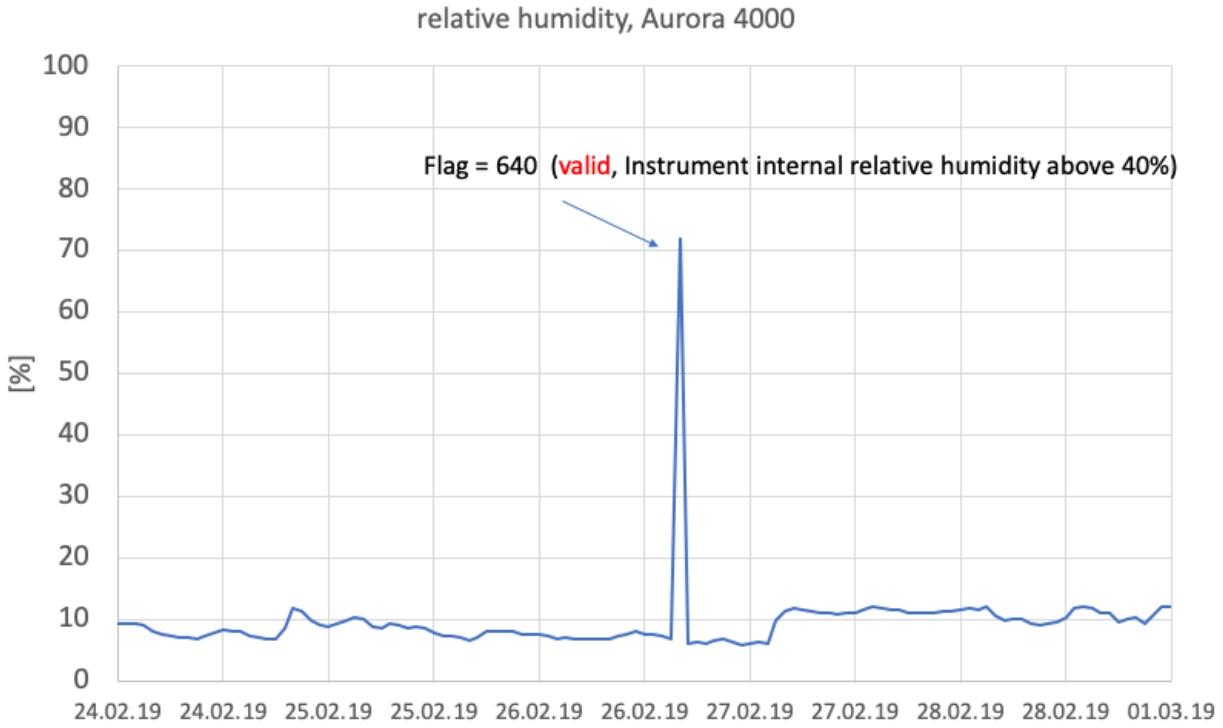


The negative spikes cannot be explained with the available data.

The unusual spectral run is probably because of a drift of the baseline of the red channel.

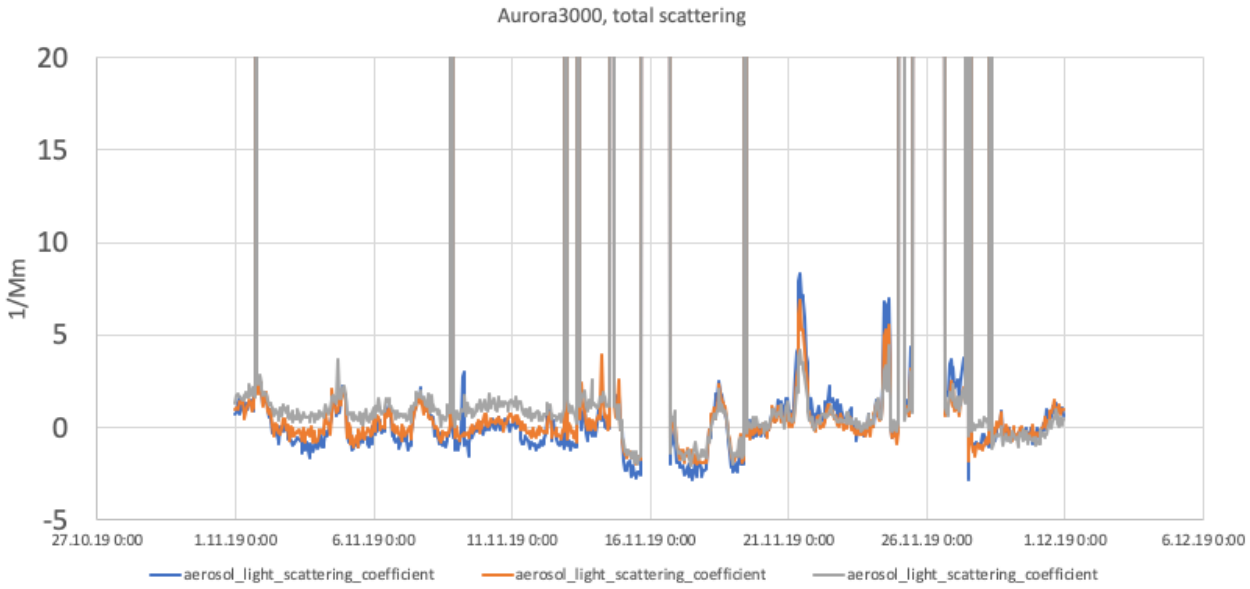
The detection limit (2 sigma) for 60 minute averages (without Kalman filter) for Aurora is about 0.04 1/Mm

Example 2:



Aurora internal humidity larger 70% for one hour average. Flag was correctly set to 640 (valid measurement). But, what can cause such a peak in the relative humidity? Is it a realistic aerosol measurement? Only the data provider can answer this question, and we must trust him.

Example3:



Obvious problems with baseline drift or baseline measurements. Without level 0 data with the baseline measurements it is not possible to determine what the problem is and which data needs to be flagged.