Determination of Temporal Stability and Instrument Performance of an airborne QCLAS via Allan-Werle-plots

Lenard L. Röder (✉ lenard.roeder@mpic.de)

   Max-Planck-Institute for Chemistry

Linda M. Ort

   Max-Planck-Institute for Chemistry

Jos Lelieveld

   Max-Planck-Institute for Chemistry

Horst Fischer

   Max-Planck-Institute for Chemistry

Research Article

Additional Declarations: No competing interests reported.

Determination of Temporal Stability and Instrument Performance of an airborne QCLAS via Allan-Werle-plots

Lenard L. Röder[1*], Linda M. Ort[1], Jos Lelieveld[1,2], Horst Fischer[1]

[1*]Atmospheric Chemistry Department, Max-Planck-Institute for Chemistry, Hahn-Meitner-Weg 1, Mainz, 55128, RLP, Germany.
[2]Climate and Atmosphere Research Center, Cyprus Institute, 1645, Nicosia, Cyprus.

*Corresponding author(s). E-mail(s): lenard.roeder@mpic.de;
Contributing authors: linda.ort@mpic.de; jos.lelieveld@mpic.de;
horst.fischer@mpic.de;

Abstract

Allan-Werle-plots are an established tool in infrared absorption spectroscopy to quantify temporal stability, maximum integration time and best achievable precision of a measurement instrument. In field measurements aboard a moving platform, however, long integration times reduce time resolution and smooth atmospheric variability. A high accuracy and time resolution are necessary as well as an appropriate estimate of the measurement uncertainty. In this study, Allan-Werle-plots of calibration gas measurements are studied to analyze the temporal characteristics of a Quantum Cascade Laser Absorption Spectrometer (QCLAS) instrument for airborne operation. Via least-squares fitting the individual noise contributions can be quantified and different dominant regimes can be identified. Through simulation of data according to the characteristics from the Allan-Werle-plot, the effects of selected intervals between in-flight calibrations can be analyzed. An interval of **30** min is found sufficient for successful drift correction during ground operation. The linear interpolation of the sensitivity increases the accuracy and lowers the measurement uncertainty from **1.1 %** to **0.2 %**. Airborne operation yields similar results during segments of stable flight but additional flicker and sinusoidal contributions. Simulations verify an appropriate interval of **30** min in airborne operation. The expected airborne measurement uncertainty is in the range **0.8 %** to **2.4 %**.

1 Introduction

Allan-Werle-plots have become a major tool in infrared spectroscopy applications for the characterization of instrumental stability and performance for field experiments [1–7]. Werle et al. [8] originally applied the two-sample-variance, or Allan Variance, to tunable diode laser absorption spectroscopy (TDLAS) to determine the maximum integration time and minimum achievable precision at this integration time.

However, a large maximum integration time is not necessarily needed in field experiments. For many species, the atmospheric variability is much higher than the instrument uncertainty [9]. In this case, averaging the data to a lower time resolution filters atmospheric features of the time series. In field experiments aboard a moving platform, e.g. aircraft-borne or drone-borne field experiments, a high temporal resolution is needed to enable localization of the measurements [10] and further data-driven analysis, e.g. calculation of fluxes via eddy covariance and wavelet transforms [11–13]. Furthermore, determination of the instrumental properties in the laboratory or in a stationary environment may not be transferable to in-flight performance [14].

According to Werle (2011) [15] many more features of the instrumental temporal dependencies can be extracted from an Allan-Werle-plot. Apart from uncorrelated, white noise, and random walk drifts, other noise contributions, e.g. flicker noise and sinusoidal interferences can be determined. This study will characterize instrumental temporal dependencies of the TRISTAR (tracer insitu TDLAS for atmospheric research) instrument described in [2, 16–18] and investigate differences in the performance during ground and airborne operation. Using the information provided via Allan-Werle-plots, suitable repetition times of in-flight calibrations and their implication to measurement uncertainty estimation will be discussed.

2 The TRISTAR instrument

The TRISTAR instrument has been originally designed by Wienhold et al. (1998) [16] as a three-channel TDLAS that accumulates three laser beam paths via beam splitters [16] and in a later design via pneumatic moving mirrors [17]. The combined beam path is guided though an astigmatic White cell [16, 19] with a base length of $0.5\,\mathrm{m}$ and 128 passes. The outgoing beam is split into a signal and reference beam and both are focused to individual detectors. A reference cell filled with a high concentration of the target gas is inserted into the reference beam path to enable line determination and line locking via laser current offset adaption. The instrument has been successfully deployed in various aircraft campaigns [2, 20–22] aboard a Learjet 35A and the High Altitude Long-range (HALO) Gulfstream G550 research aircraft. During these campaigns a range of trace gases has been measured including CO, CH_4, HCHO, NO_2 and N_2O. Measurement uncertainties of $0.5\,\%$, $0.5\,\%$ and $0.7\,\%$ have been achieved for CO, N_2O and CH_4, respectively [2].

2

During 25 years of operation, various modifications have been applied to the original setup, including a major revision since the last reported experimental setup by Tadic et al. (2017) [18]. The cryogenic cooling system has been abolished by replacing individual components that were cooled via liquid nitrogen, i.e. lasers and detectors, by thermoelectrically cooled components. The three-laser setup has been replaced by single-laser operation which increases data coverage of the target species. While these measures might decrease the optimal performance and limit the diversity of data output, a simpler instrument maintenance and more autonomous operation can be achieved. The latter is enabled by substituting the V25 electronics and DOS PC by a National Instruments real time controller paired with an ARK Windows PC running a LabView real time application and graphical user interface. Establishing a PC-based user interface enables remote operation via internet and reduces the required training time for airborne operators in aircraft measurement campaigns. Autonomous and remote operation has been demonstrated during the ground-based field experiment at the German Taunus Observatory in 2021 [23], requiring only maintenance and calibration gas source exchange as on-site activities.

2.1 Optics and gas flow

A room temperature quantum cascade laser (RT-QCLAS) inside a high heat load (HHL) housing (AlpesLasers, Switzerland) is driven in cw-mode and emits a divergent beam. The beam is captured and focused onto a first focal point via a three-mirror telescopic optics (IPM Freiburg). Via a pair of an off-axis parabolic (OAP) and elliptic (OAE)

mirror the beam is converted into a parallel beam with a width of approximately 14 mm. The large beam waist makes the setup less sensitive to aerosol particles and imperfections on the mirrors and turbulence in the optical path.

Several planar mirrors guide the path to the entrance of the White cell. There the parallel beam is focused to a second focal point using another pair of an OAE and OAP mirror. Then the divergent beam is injected through wedged CaF_2 windows into the double corner cube White cell, where it gets reflected 128 times.

The White cell has a length of 0.5 m and holds a volume of 3 l. The cell is drained by a scroll vacuum pump set to a constant gas flow of 3 slm (standard liters per minute). On the top of the cell a movable valve operated by a step motor regulates the cell pressure to 50 hPa. The inlet is connected to a T-junction, where calibration gas can be injected by a mass flow controller (Bronkhorst). Excess calibration gas exits through the ambient air inlet. A detailed schematic of the tubing is given in Tadic et al. (2017) [18], Figure 2.

The outgoing beam is focused on a third focal point and converted back into a parallel beam using another pair of OAE and OAP mirrors. The parallel beam is split via a CaF_2 beam splitter (50:50) into signal and reference beam. The reference beam is further passed through a reference cell filled with high concentration of the target gas. Both parallel beams are focused via coated CaF_2 lenses with focal length of 40 mm onto room-temperature photovoltaic infrared detectors (VIGO systems, Poland). All optical components are mounted on a rigid aluminum plate that is electrically heated to ensure thermal stability. This

makes the optical setup less sensitive to distance variations caused by thermal fluctuations.

2.2 Electronics and data processing

The instrument runs autonomously inside a NI LabView real time application on the real time controller (NI cRIO 9038). The software controls pump, inlet and MFC and writes data to disk. Through the FPGA module the scanning ramp and modulation is set at a clock frequency of 250 kHz with a modulation frequency of 31.25 kHz and a scanning frequency of 60 ramps per second. The ramp and modulation are input to the laser controller that operates the QCL at a set temperature and laser base current to obtain a laser frequency near the absorption central frequency.

The detector output is digitally demodulated at twice the modulation frequency for $2f$ wavelength modulation spectroscopy (WMS). The raw signal is down-sampled by a CIC-stage to reduce the amount of data. Individual spectra of the increasing (up) and the decreasing (down) part of the triangular scanning frequency are averaged separately to obtain a time resolution of 1 Hz. The averaged up and down spectra are saved to disk and further processed for online operation.

The described configuration of modulation frequency, scanning frequency and time resolution can be adjusted to the needs of the experimental requirements.

Throughout the experiments carried out in this study, the QCL is operated near $2190 \, \text{cm}^{-1}$, where a strong line pair of CO and N_2O is located [24]. The scanning range is optimized to enable simultaneous measurements of both species.

The saved data is post-processed using custom code written in IGOR Pro (Wavemetrics). The raw signal is filtered using a IIR low-pass filter, the line centers and widths are determined from the reference signal and the spectra are linearly fit against the most stable absorption spectrum obtained during a calibration measurement.

2.3 Instrumental performance measures

The total measurement uncertainty is composed of several metrics that each contribute differently to the limitation of the measurement device. The systematic error is defined by the accuracy and describes the overall deviation of all data points from the true value. Since the TRISTAR device uses the relative measurement mode of WMS, the uncertainty of the conversion factor from arbitrary units to mixing ratio contributes to the accuracy. The accuracy depends on the calibration gas mixing ratio uncertainty as determined in the laboratory or provided by the gas standard supplier.

The systematic deviations caused from drifts can be mitigated by regular calibrations. Considering changes in relative sensitivity, the drifts can be corrected via linear interpolation of the reciprocal mixing ratios. This interpolation scheme requires that the drift speed is not faster than the interval between calibrations. The reproducibility R refers to the relative spread of the calibrations over a complete measurement. This uncertainty cannot be lowered by further averaging.

The precision P refers to the relative statistical error at a given integration time after drift correction. It is obtained from the relative standard deviation of the calibration measurements. The total relative error at native time resolution is

a combination of P and R:

$$\frac{\sigma_\chi}{\chi} = \sqrt{P^2 + R^2} \qquad (1)$$

The detection limit refers to the lowest mixing ratio distinguishable from noise. It is estimated from the standard deviation of absorption-free gas injected into the absorption cell. By interpreting the precision as pure absolute error, the precision at the mixing ratio of the calibration gas gives an upper boundary and a worst case estimate to the detection limit.

Werle et al. [8] define the Allan variance as the mean squared difference between adjacent data batches for different integration times τ:

$$\sigma_\chi(\tau)^2 = \frac{1}{2} \sum_{k=0}^{M-2} (\chi_{k+1}(\tau) - \chi_k(\tau))^2 \quad (2)$$

$$\chi_k(\tau) = \frac{1}{\tau} \int_{t=k\tau}^{(k+1)\tau} \chi(t)dt \qquad (3)$$

As reported in [15], plotting the Allan variance as a function of the integration time (Allan-Werle-plot) provides insight into the temporal characteristics of an absorption spectroscopy measurement. The maximum integration time $\tau_{\max}$ refers to the minimum in the Allan-Werle-plot where the trade-off between the decrease from the integration of white noise and the build-up from integration of drifts is optimal. Considering only uncorrelated white noise σ_W and random walk noise σ_B the Allan-variance σ^2 and $\tau_{\max}$ can be modeled as

$$\sigma^2(\tau) = \frac{\sigma_W^2}{\tau} + \frac{\sigma_B^2}{2}\tau \text{ and} \qquad (4)$$

$$\tau_{\max} = \sqrt{2}\frac{\sigma_W}{\sigma_B}, \qquad (5)$$

respectively.

3 Ground operation inside the HALO aircraft

The TRISTAR instrument is part of the TRIHOP rack that also carries the hydrogen peroxide and higher organic peroxide monitor HYPHOP [25]. The rack is mounted into the research aircraft HALO. The inlet is connected to a T-junction that may provide scrubbed cabin air during zero cycles. Scrubbing is shut off during CO measurements. The calibration gas tube is connected to a 6 l composite bottle (AUER). During ground operation the measurement devices run on external power while the aircraft auxiliary power unit and main engine remain shut off. As a consequence the integrated air conditioning of the aircraft also remains off and air conditioning has to be provided externally by ventilation. This lead to a slowly increasing temperature inside the aircraft over a whole day of ground operation up to 35 °C cabin temperature.

3.1 Performance of calibration gas measurement

Several functional tests were carried out during ground operation. Determining the precision and long-term stability of the instrument is important to estimate the expected performance during the research flights. In order to collect suitable data the cell was flushed with calibration gas regularly and also for longer time periods. Extended measurements of calibration gas provide a good estimate of the statistical measurement uncertainty directly from the standard deviation. A time series and corresponding Allan-Werle-plot of such an extended

period of calibration gas measurement is shown in Figure 1. Calibration gas was injected over a period of 2 hours. The absolute and relative standard deviation of the measurement are 1.7 ppbv and 1.1 %, respectively. The Allan-variance $\sigma^2(\tau)$ can successfully be modeled by pure noise σ_W and drift σ_B contributions. The optimal fit parameters are given in the figure caption of Figure 1 and the maximum integration time results to 20 s according to (5).
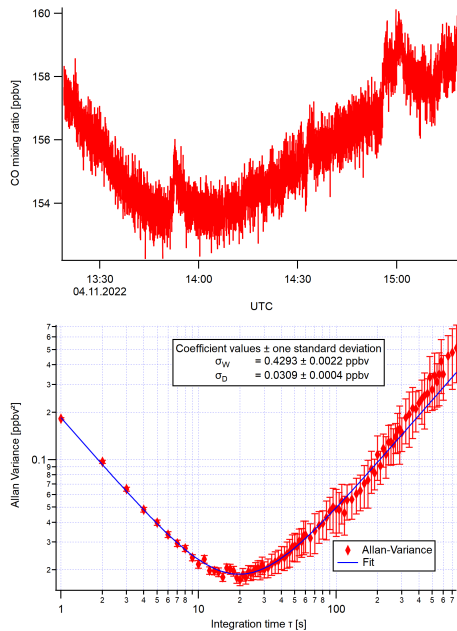

Fig. 1 Calibration gas measurement during ground operation inside the HALO aircraft. Top: time series of the 2 hour segment of calibration gas. Bottom: Corresponding Allan-Werle-plot with fit to (4). The optimal fit parameters are $\sigma_W = (0.4293 \pm 0.0022)$ ppbv and $\sigma_B = (0.0309 \pm 0.0004)$ ppbv.

3.2 Performance of ambient gas measurement

Determining the measurement uncertainty as indicated in the previous section is not possible during ambient measurements, however, since the standard deviation is a mixture of atmospheric variability and instrument uncertainty. Instrument drifts are also hard to distinguish from changes in the mixing ratios of the sampled air masses. Determining an appropriate interval of calibrations to minimize the effect of drifts and estimate the measurement uncertainty is a non-trivial task during field experiments and airborne field experiments, especially.

Using the Allan-Werle-plot in Figure 1, similar data can be simulated that shares the temporal characteristics obtained from the measurement. This can be achieved with the following simulation setup, where χ_0 refers to the mixing ratio of the calibration gas.

$$h_t = h_{t-1} + \epsilon_B \qquad (6)$$
$$\chi_t = \chi_0 + h_t + \epsilon_W \qquad (7)$$
$$\text{with } \epsilon_B \sim \mathcal{N}(0, \sigma_B^2),$$
$$\epsilon_W \sim \mathcal{N}(0, \sigma_W^2)$$

A resulting time series can be found in Figure 2. Starting from this simulated data, points can be artificially flagged as calibration and ambient to simulate the effect of different intervals between calibrations during a real field measurement scenario. The effects on the correction of the drifts and the estimation of the true measurement uncertainty can be studied. For each individual simulation, R is calculated from the spread of the means of the flagged calibrations. The reciprocal means are interpolated to estimate the sensitivity drift. The interpolated sensitivity is multiplied to the amplitudes to

obtain the drift-corrected amplitude. P is obtained from the relative standard deviation of the drift-corrected calibration amplitudes. The relative standard deviation of the corrected amplitudes which are flagged ambient, on the other hand, gives a good estimate of the residual measurement uncertainty (RMU) due to drifts and noise present in the data. An optimal interval between calibrations will decrease the RMU until P is reached. The estimated measurement uncertainty (EMU) as defined in (1) should optimally be close to the true relative standard deviation of the simulated dataset.

Figure 3 shows the described metrics for different intervals between calibrations. The observable features match the expectations: If the calibration frequency is too low, the drifts cannot be successfully corrected. Thus, the RMU is close to the original measurement uncertainty. Furthermore, the EMU is very unstable since there are too few calibration measurements in total. With more calibrations both of these effects are reduced. The EMU approximates the original measurement uncertainty better and the RMU decreases. In principle, more calibrations are always advantageous. However, more calibrations also result in less ambient measurements and a high calibration gas consumption. An empirical choice is made that subdivides the calibration frequencies into three regimes: Too many calibrations only contribute diminishing returns while decreasing the duty cycle, too few calibrations cause the effects described at the beginning of this paragraph, and the appropriate choice lies in the middle. The appropriate region is chosen within $10\tau_{max}$ and $100\tau_{max}$, where τ_{max} refers to the maximum integration time derived from the Allan-Werle-plot.

Accordingly a calibration interval of 30 min lies within the desired regime. On another day of ground operation, regular calibrations were carried out at that interval. The resulting values for R and P are 1.0% and 0.2%, respectively, and accurately estimate the total measurement uncertainty. Furthermore, a successful correction for the random walk drifts can be assumed considering the results of the simulation shown in Figure 3. The residual measurement uncertainty of the ambient data can therefore be estimated with the precision P, reducing the uncertainty from a value of 1.1% to 0.2% as a consequence of regular calibrations.

4 Airborne operation

In the previous section, the instrument has been characterized during ground operation aboard the HALO aircraft. An appropriate interval for calibrations has been chosen that effectively corrects for random walk drifts and allows for a reduction of the measurement uncertainty. Now the question arises: Is the ground-based performance transferable to airborne operation?

In order to study the instrument response to disturbances caused by airborne operation, e.g., vibration, acceleration, orientation and pressure, a test flight was performed where TRISTAR was constantly measuring calibration gas. The data acquired can be used to qualitatively and quantitatively analyze the effects of this new environment to the instrumental performance.

Figure 4 shows the time series and the corresponding Allan-Werle-plot of this test flight. The time series indicates different environmental conditions throughout the flight. During takeoff and landing
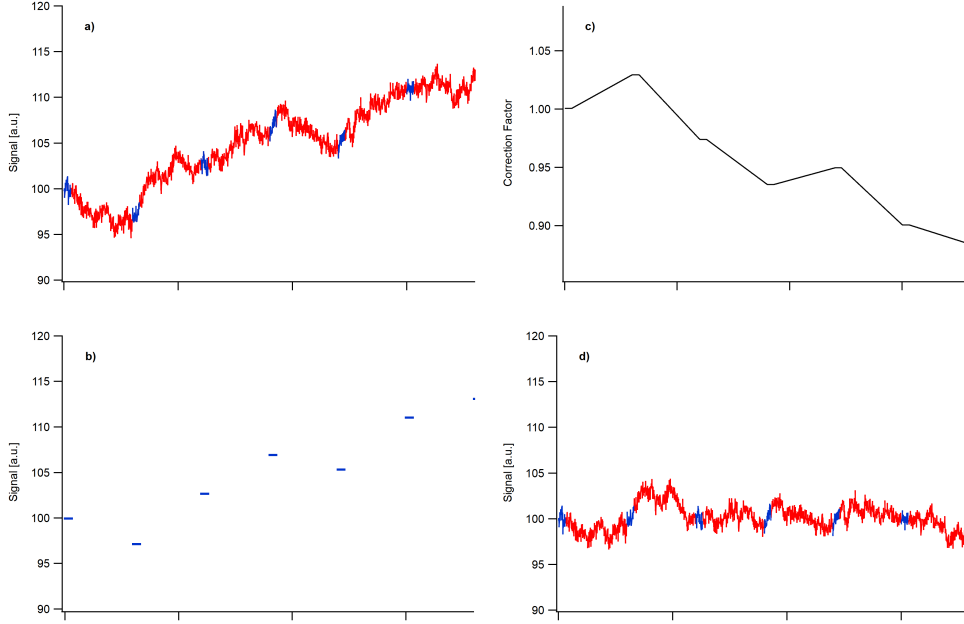
Fig. 2 Example simulation process carried out to retrieve the metrics P, R, RMU and EMU. Top left: Simulated time series with flagged ambient (red) and calibration (blue) segments. Bottom left: individual means of flagged calibrations. The standard deviation of these values is used to calculate R. Top right: Interpolated sensitivity that is connected to the reciprocal of the calibration means. Bottom right: Drift-corrected time series. The standard deviation of the data flagged calibration yields P, the standard deviation of the data flagged ambient yields RMU.

periods, rapid accelerations impact the optical setup and lead to distortions in the signal. At high altitudes, the pressure regulation of the inlet starts to oscillate and results in oscillations of the gas concentration. In post processing these oscillations can be corrected by assuming a linear dependence on the cell pressure. On intermediate flight levels the system is more stable. Therefore, calculating the Allan variance for the complete data set is no longer meaningful, as the experimental conditions and the underlying distribution is changing. Thus, the Allan variance is calculated individually for the different conditions derived from the time series and metadata.

The Allan variance of the total data set is a combination of the individual variances. This obstructs the information extractable in the individual Allan-Werle-plots. The takeoff and landing phase is dominated by a strong sine wave and drift that are caused by the mechanical system reacting and recovering from the strong changes in acceleration. The high altitude section is dominated by a strong sine wave that matches the oscillations caused by the pressure regulation. The flight segments on intermediate flight levels are quite stable. During these segments a series of left and right turns and various altitude changes were carried out to study the instrument performance.
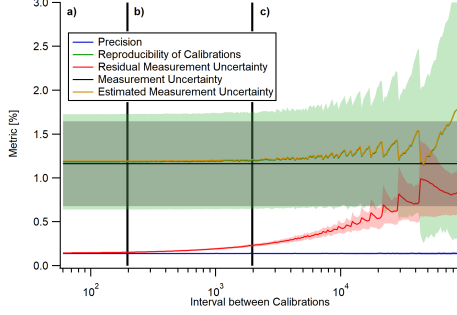
8

Fig. 3 Performance metrics for different intervals between calibrations. The simulation was repeated 100 times. The solid lines indicate the mean values and the shaded areas indicate the standard deviations of each metric for the 100 repetitions. The metrics shown are R (green), P (blue), RMU (red) and EMU (yellow). The black horizontal line shows the total measurement uncertainty of the simulated time series. The black vertical lines separate the three regimes a): too many calibrations, b): appropriate interval between calibrations and c): too few calibrations. The spikes in the results originate from discrete changes in the number of calibrations that fall inside the total number of data points in the time series for each given calibration interval.

Fig. 4 Calibration gas measurement during test flight #01 on the 22nd of November, 2022. The data is segmented according to changes in the environmental conditions into the takeoff and landing phase (yellow), first plateau segment (green), high altitude segment (blue), second plateau segment (grey), complete data set (red). Top: time series of the 5 hour flight, separated into these segments, plotted against the left axis and GPS altitude (black) plotted against the right axis. Bottom: Corresponding Allan-Werle-plots with fits to (8). The optimal fit parameters are summarized in Table 1.

Again, the Allan variances are fit against theoretical models to quantify the temporal features. This time the composition is more complex, so pure white noise and random walk is not sufficient to reproduce the measured features. Flicker noise σ_F is added that acts as a noise floor according to [15]. Sinusoidal variations contribute a term proportional to $A_{\sin}$ and dependent on the period $T_{\sin}$. Additionally, constrained random walk noise that behaves like white noise for longer time spans is introduced with amplitude A_{MH} and transition time T_{MH}. The theoretical model results in (8), and the most appropriate fit is chosen with the least number of parameters. The resulting fit parameters are summarized in Table 1.

During the stable segments the values extracted from the Allan-Werle-plots
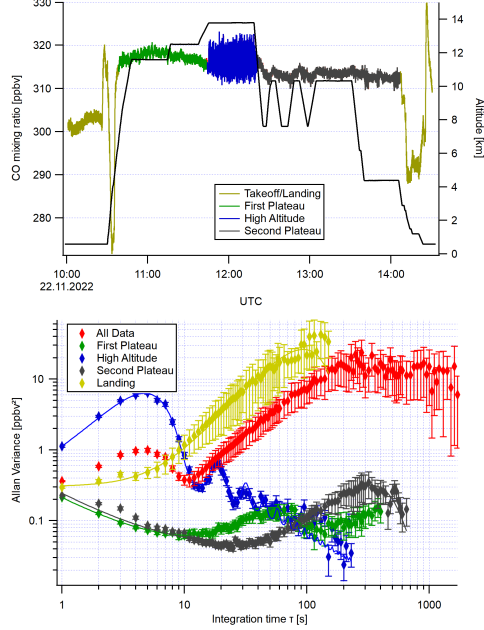
for the white noise and random walk contributions are similar to the values obtained from the Allan-Werle-plot of the ground operation. This indicates similar effects influencing the instrumental performance on both modes of operation. Features exclusive to airborne operation are manifested through additional flicker and sinusoidal contributions during the stable segments. More noticeable features are visible in the time series and enable segmentation of the time series.

9

$$\sigma^2(\tau) = \frac{\sigma_W^2}{\tau} + \frac{\sigma_B^2}{2}\tau + \sigma_F^2 + A_{\mathrm{sin}}^2 \frac{\sin\left(\pi\frac{\tau}{T_{\mathrm{sin}}}\right)^4}{\left(\pi\frac{\tau}{T_{\mathrm{sin}}}\right)^2} + \frac{A_{\mathrm{MH}}^2}{\sqrt{2}}\left(\sqrt{\frac{T_{\mathrm{MH}}-1}{\tau}} + \sqrt{\frac{\tau}{T_{\mathrm{MH}}}}\right)^{-2} \quad (8)$$

Similar to the analysis carried out in the previous section, data was simulated that follows the temporal characteristics of the two stable segments of the flight. The results of the simulation procedure is shown in Figure 5. Since the Allan Variance in these cases contains flicker and sinusoidal components, the results differ from the simulation results of the ground operation. The flicker and sinusoidal components contribute further short-term uncertainties that increase the RMU on shorter periods even with high frequency of calibrations. Thus, the measurement uncertainty on ambient data can no longer be estimated based on the precision only but needs to be approximated by the EMU (1). Accurate estimation of the EMU is again possible within the appropriate regime of calibration intervals. This regime can again be calculated from (5) and lies inbetween $10\tau_{\mathrm{max}}$ and $100\tau_{\mathrm{max}}$. For the first plateau segment, τ_{max} is equal to 26 s. For the second plateau segment, no value for σ_B could be extracted from the Allan-Werle plot. A time interval of 30 min again lies inside the desired range.

During a second test flight, the operation of the TRISTAR instrument was switched to an operation scheme used in real research flights. This way the instrumental behavior during gas switching and pressure regulation at different altitudes were tested. Ambient gas measurements were carried out and regular calibration measurements were performed for drift correction and measurement uncertainty determination.

The time series of Testflight #2 is shown in Figure 6. The precision P and reproducibility R as evaluated from the repeated calibrations are equal to 0.35 % and 1.50 %, respectively. The estimated measurement uncertainty lies between the relative standard deviation of the total Testflight #1 and the reduced data set, where takeoff and landing phase are discarded. These values are equal to 1.5 %, 2.4 % and 0.76 %, respectively.

5 Summary and Conclusion

Allan-Werle-plots have become a key tool in infrared absorption spectroscopy to investigate the temporal stability of a measurement instrument. However, these plots are commonly used to determine the maximum integration time and the best achievable precision. In this study, Allan-Werle-plots were used to analyze the temporal characteristics of the newest iteration of the TRISTAR instrument both for ground operation inside the research aircraft and for airborne operation. Determination of the instrumental stability through the maximum integration time is not beneficial due to the desired time resolution during research flights. Obviously, such unstable measurement environment like a moving aircraft can lead to strong drifts.

Instead, the Allan-Werle-plots were used to identify the dominant noise types present during the measurement. Via least-squares fitting the contributions of
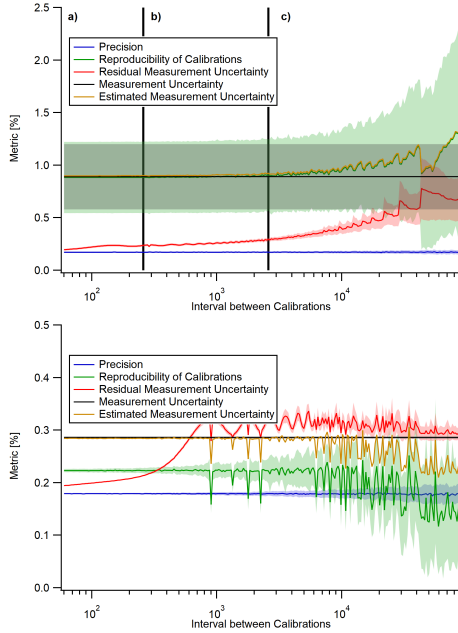
Fig. 5 Performance metrics for different intervals between calibrations. The simulation was repeated 100 times. Top: First plateau segment of Testflight #1. Bottom: Second plateau segment of Testflight #1. The solid lines indicate the mean values and the shaded areas indicate the standard deviations of each metric for the 100 repetitions. The metrics shown are R (green), P (blue), RMU (red) and EMU (yellow). The black horizontal line shows the total measurement uncertainty of the simulated time series. The black vertical lines (only applies to top plot) separate the three regimes a): too many calibrations, b): appropriate interval between calibrations and c): too few calibrations. The spikes in the results originate from discrete changes in the number of calibrations that fall inside the total number of data points in the time series for each given calibration interval.

Fig. 6 Ambient and calibration gas measurement during Testflight #02 on the 25th of November, 2022. The plot shows the ambient mixing ratio (red), calibration gas mixing ratio (blue) plotted against the left axis and GPS altitude (black) plotted against the right axis.

white noise, random walk noise, flicker noise, meta-stable random walks and sinusoidal contributions were obtained for the different measurement conditions.

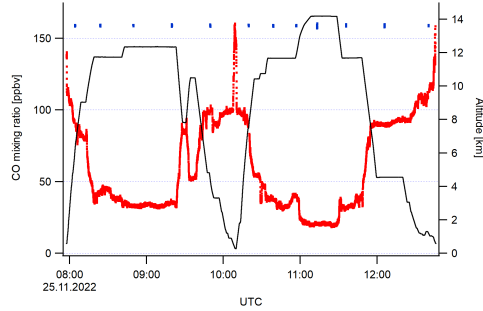Simulation of data that recreates the characteristics observed in the Allan-Werle-plots allowed for an extensive analysis of the choice of the interval between calibrations. Several metrics were defined to identify the capability of the procedure to correct for drifts and estimate the true measurement uncertainty. An interval of 30 min was found sufficient for complete drift correction during ground operation. Thus, estimation of the measurement uncertainty with only the precision was enabled. The expectations were verified using a second data set with the appropriate calibration frequency.

The same analysis was carried out using data from two test flights. Segmentation of the time series according to the sudden changes in the measurement conditions enabled resolution of the dominant noise types for each individual segment during the flight. Dominant sinusoidal contributions were observed during a high altitude segment due to oscillations in the absorption cell pressure regulation. Low-frequency contributions dominated the takeoff and landing phases. Investigation of stable segments during the flight showed similar characteristics to ground operation. This designated sufficient intervals between calibrations of 30 min during research flights

for measurement uncertainty estimation and drift correction. However, apparent flicker and sinusoidal contributions lead to a lowered accuracy for reasonable calibration frequencies. Therefore, the measurement uncertainty had to be estimated considering the reproducibility of calibrations. A second test flight verified the estimated measurement uncertainty for the chosen calibration frequency.

The results of this study show yet another great application of Allan-Werle-plots to increase the understanding and the accuracy of infrared absorption spectrometers. The implications for drift correction and measurement uncertainty are also transferable to other measurement instruments for trace gas measurements.

Declarations

References

[1] Weibring, P., Richter, D., Fried, A., Walega, J., Dyroff, C.: Ultra-high-precision mid-IR spectrometer II: system description and spectroscopic performance. Applied Physics B **85**, 207–218 (2006)

[2] Schiller, C., Bozem, H., Gurk, C., Parchatka, U., Königstedt, R., Harris, G., Lelieveld, J., Fischer, H.: Applications of quantum cascade lasers for sensitive trace gas measurements of CO, CH_4, N_2O and HCHO. Applied Physics B **92**(3), 419–430 (2008)

[3] Tuzson, B., Mangold, M., Looser, H., Manninen, A., Emmenegger, L.: Compact multipass optical cell for laser spectroscopy. Optics letters **38**(3), 257–259 (2013)

[4] Ehlers, P., Johansson, A.C., Silander, I., Foltynowicz, A., Axner, O.: Use of etalon-immune distances to reduce the influence of background signals in frequency-modulation spectroscopy and noise-immune cavity-enhanced optical heterodyne molecular spectroscopy. J. Opt. Soc. Am. B **31**(12), 2938–2945 (2014) https://doi.org/10.1364/JOSAB.31. 002938

[5] Giglio, M., Patimisco, P., Sampaolo, A., Scamarcio, G., Tittel, F.K., Spagnolo, V.: Allan Deviation Plot as a Tool for Quartz-Enhanced Photoacoustic Sensors Noise Analysis. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control **63**(4), 555–560 (2016) https://doi.org/10.1109/TUFFC. 2015.2495013

[6] Dong, L., Li, C., Sanchez, N.P., Gluszek, A.K., Griffin, R.J., Tittel, F.K.: Compact CH_4 sensor system based on a continuous-wave,

low power consumption, room temperature interband cascade laser. Applied Physics Letters **108**(1), 011106 (2016) https://doi.org/10.1063/1.4939452

[7] Wilson, D., Phair, J.W., Lengden, M.: Performance Analysis of a Novel Pyroelectric Device for Non-Dispersive Infra-Red CO_2 Detection. IEEE Sensors Journal **19**(15), 6006–6011 (2019) https://doi.org/10.1109/JSEN.2019.2911737

[8] Werle, P., Mücke, R., Slemr, F.: The limits of signal averaging in atmospheric trace-gas monitoring by tunable diode-laser absorption spectroscopy (TDLAS). Applied Physics B **57**(2), 131–139 (1993)

[9] Junge, C.E.: Residence time and variability of tropospheric trace gases. Tellus **26**(4), 477–488 (1974) https://doi.org/10.3402/tellusa.v26i4.9853

[10] Liao, J., Wolfe, G.M., Hannun, R.A., St. Clair, J.M., Hanisco, T.F., Gilman, J.B., Lamplugh, A., Selimovic, V., Diskin, G.S., Nowak, J.B., Halliday, H.S., DiGangi, J.P., Hall, S.R., Ullmann, K., Holmes, C.D., Fite, C.H., Agastra, A., Ryerson, T.B., Peischl, J., Bourgeois, I., Warneke, C., Coggon, M.M., Gkatzelis, G.I., Sekimoto, K., Fried, A., Richter, D., Weibring, P., Apel, E.C., Hornbrook, R.S., Brown, S.S., Womack, C.C., Robinson, M.A., Washenfelder, R.A., Veres, P.R., Neuman, J.A.: Formaldehyde evolution in US wildfire plumes during the Fire Influence on Regional to Global Environments and Air Quality experiment (FIREX-AQ).

Atmospheric Chemistry and Physics **21**(24), 18319–18331 (2021) https://doi.org/10.5194/acp-21-18319-2021

[11] Yuan, B., Kaser, L., Karl, T., Graus, M., Peischl, J., Campos, T.L., Shertz, S., Apel, E.C., Hornbrook, R.S., Hills, A., Gilman, J.B., Lerner, B.M., Warneke, C., Flocke, F.M., Ryerson, T.B., Guenther, A.B., Gouw, J.A.: Airborne flux measurements of methane and volatile organic compounds over the Haynesville and Marcellus shale gas production regions. Journal of Geophysical Research: Atmospheres **120**(12), 6271–6289 (2015) https://doi.org/10.1002/2015JD023242

[12] Wolfe, G.M., Hanisco, T.F., Arkinson, H.L., Bui, T.P., Crounse, J.D., Dean-Day, J., Goldstein, A., Guenther, A., Hall, S.R., Huey, G., Jacob, D.J., Karl, T., Kim, P.S., Liu, X., Marvin, M.R., Mikoviny, T., Misztal, P.K., Nguyen, T.B., Peischl, J., Pollack, I., Ryerson, T., St. Clair, J.M., Teng, A., Travis, K.R., Ullmann, K., Wennberg, P.O., Wisthaler, A.: Quantifying sources and sinks of reactive gases in the lower atmosphere using airborne flux observations. Geophysical Research Letters **42**(19), 8231–8240 (2015) https://doi.org/10.1002/2015GL065839

[13] Nussbaumer, C.M., Place, B.K., Zhu, Q., Pfannerstill, E.Y., Wooldridge, P., Schulze, B.C., Arata, C., Ward, R., Bucholtz, A., Seinfeld, J.H., Goldstein, A.H., Cohen, R.C.: Measurement report: Airborne measurements of NO_x fluxes over Los Angeles during the RECAP-CA 2021 campaign. EGUsphere **2023**, 1–20

(2023) https://doi.org/10.5194/egusphere-2023-601

[14] Richter, D., Weibring, P., Walega, J.G., Fried, A., Spuler, S.M., Taubman, M.S.: Compact highly sensitive multi-species airborne mid-IR spectrometer. Applied Physics B **119**(1), 119–131 (2015)

[15] Werle, P.: Accuracy and precision of laser spectrometers for trace gas sensing in the presence of optical fringes and atmospheric turbulence. Applied Physics B **102**, 313–329 (2011)

[16] Wienhold, F., Fischer, H., Hoor, P., Wagner, V., Königstedt, R., Harris, G., Anders, J., Grisar, R., Knothe, M., Riedel, W., *et al.*: TRISTAR–a tracer in situ TDLAS for atmospheric research. Applied Physics B **67**(4), 411–417 (1998)

[17] Kormann, R., Fischer, H., Gurk, C., Helleis, F., Klüpfel, T., Kowalski, K., Königstedt, R., Parchatka, U., Wagner, V.: Application of a multi-laser tunable diode laser absorption spectrometer for atmospheric trace gas measurements at sub-ppbv levels. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy **58**(11), 2489–2498 (2002) https://doi.org/10.1016/S1386-1425(02)00066-5

[18] Tadic, I., Parchatka, U., Königstedt, R., Fischer, H.: In-flight stability of quantum cascade laser-based infrared absorption spectroscopy measurements of atmospheric carbon monoxide. Applied Physics B **123**, 1–9 (2017)

[19] White, J.U.: Optical System Providing a Long Optical Path. 2779230, January 1957

[20] Hoor, P., Fischer, H., Lange, L., Lelieveld, J., Brunner, D.: Seasonal variations of a mixing layer in the lowermost stratosphere as identified by the CO-O_3 correlation from in situ measurements. Journal of Geophysical Research: Atmospheres **107**(D5), 1–1111 (2002) https://doi.org/10.1029/2000JD000289

[21] Engel, A., Bönisch, H., Brunner, D., Fischer, H., Franke, H., Günther, G., Gurk, C., Hegglin, M., Hoor, P., Königstedt, R., Krebsbach, M., Maser, R., Parchatka, U., Peter, T., Schell, D., Schiller, C., Schmidt, U., Spelten, N., Szabo, T., Weers, U., Wernli, H., Wetter, T., Wirth, V.: Highly resolved observations of trace gases in the lowermost stratosphere and upper troposphere from the Spurt project: an overview. Atmospheric Chemistry and Physics **6**(2), 283–301 (2006) https://doi.org/10.5194/acp-6-283-2006

[22] Tomsche, L., Pozzer, A., Ojha, N., Parchatka, U., Lelieveld, J., Fischer, H.: Upper tropospheric CH_4 and CO affected by the South Asian summer monsoon during the Oxidation Mechanism Observations mission. Atmospheric Chemistry and Physics **19**(3), 1915–1939 (2019) https://doi.org/10.5194/acp-19-1915-2019

[23] Dewald, P., Nussbaumer, C.M., Schuladen, J., Ringsdorf, A., Edtbauer, A., Fischer, H., Williams, J., Lelieveld, J., Crowley, J.N.: Fate of the nitrate radical at the summit of

a semi-rural mountain site in Germany assessed with direct reactivity measurements. Atmospheric Chemistry and Physics **22**(10), 7051–7069 (2022) https://doi.org/10.5194/acp-22-7051-2022

[24] Li, G., Gordon, I.E., Rothman, L.S., Tan, Y., Hu, S.-M., Kassi, S., Campargue, A., Medvedev, E.S.: Rovibrational line lists for nine isotopologues of the CO molecule in the $X1\sigma+$ ground electronic state. The Astrophysical Journal Supplement Series **216**(1), 15 (2015)

[25] Hamryszczak, Z., Hartmann, A., Dienhart, D., Hafermann, S., Brendel, B., Königstedt, R., Parchatka, U., Lelieveld, J., Fischer, H.: HYPHOP: a tool for high-altitude, long-range monitoring of hydrogen peroxide and higher organic peroxides in the atmosphere. Atmospheric Measurement Techniques Discussions **2023**, 1–25 (2023)

Parameter	Plateau Left	Plateau Right	Pressure Variation	Landing	Ground
σ_W [ppbv]	0.421 ± 0.007	0.456 ± 0.005	–	–	0.4293 ± 0.0022
σ_B [ppbv]	0.0230 ± 0.0013	–	–	–	0.0309 ± 0.0004
σ_F [ppbv]	0.189 ± 0.006	0.1918 ± 0.0025	–	0.552 ± 0.008	–
$A_{\sin}$ [ppbv]	0.385 ± 0.014	0.730 ± 0.022	3.31 ± 0.04	7.03 ± 0.37	–
$T_{\sin}$ [s]	144 ± 5	897 ± 26	12.82 ± 0.06	251 ± 17	–
A_{MH} [ppbv]	–	–	1.68 ± 0.04	–	–
T_{MH} [s]	–	–	3.07 ± 0.17	–	–

Table 1 Optimal fit parameters of each Allan-Werle-plot shown in Figure 4 when fitting against (8). Parameters marked – are neglected since they do not contribute major features of the Allan variance. For comparison, the fit results from ground operation is also appended to the table.

15