



Ground-Based Validation of Sentinel-5P TROPOMI Atmospheric Products using Calibration-Informed Low-Cost Multi-Spectral Sensors

Wolfgang Schneider¹

¹Independent Atmospheric Research HAM RADIO Station DG2MCM, Kempten, Bavaria, Germany

Correspondence: Wolfgang Schneider (rhodeia@outlook.com)

Abstract. Ground-based validation of satellite atmospheric products is essential for ensuring data quality and algorithm performance. We present a validation approach for Sentinel-5P TROPOMI cloud fraction products using a multi-spectral ground station (DG2MCM-15) located in Kempten, Bavaria, Germany. The ground observatory combines professional metrological experience from ISO/IEC 17025 accredited laboratory environments with low-cost commercial sensors, creating a citizen science validation capability.

Our validation dataset comprises 276 temporally matched observations between Sentinel-5P overpasses and ground measurements over a four-week period (January 11 – February 8, 2026). Ground-based cloud detection using an MLX90614 infrared pyrometer achieves strong agreement with Sentinel-5P cloud fraction retrievals (Pearson $R = 0.879$, $N = 27$ after quality filtering). The root mean square error of 29.1 % cloud fraction reflects a systematic positive bias from spatial scale mismatch between the ground sensor field of view and satellite pixel dimensions. The method reliably distinguishes between clear, partially cloudy, and overcast conditions, though the derived cloud fraction values exhibit clustering due to the temperature-ratio approach used. Exploratory comparison with TROPOMI aerosol index products yielded negligible correlation due to the absence of UV spectral coverage in the ground sensor, identifying a clear instrumentation requirement for future aerosol validation work.

Temporal matching between satellite overpasses and ground observations achieved a mean time difference of 2.7 minutes, with 95% of matches within 8 minutes of satellite observation time. Spatial co-location analysis confirms all validation points fall within the nominal TROPOMI pixel footprint ($3.5 \text{ km} \times 5.5 \text{ km}$ at nadir), though the spatial scale mismatch between the ground sensor field of view and satellite pixel dimensions remains the primary source of validation uncertainty.

Our results demonstrate that low-cost infrared sensors, when operated with calibration-informed measurement protocols, can provide scientifically useful satellite cloud product screening data, reliably distinguishing between clear, partially cloudy, and overcast conditions. The quasi-discrete nature of the derived cloud fraction highlights the need for improved cloud detection algorithms in future work. This approach offers a scalable pathway for expanding ground-based validation networks in regions lacking dedicated atmospheric monitoring infrastructure.



1 Introduction

25 Satellite remote sensing of atmospheric composition has become essential for understanding global air quality, climate forcing, and atmospheric chemistry processes (Veefkind et al., 2012). The Sentinel-5 Precursor (Sentinel-5P) mission, launched in 2017 as part of the European Copernicus programme, carries the TROPospheric Monitoring Instrument (TROPOMI) providing daily global coverage of key atmospheric trace gases and aerosols at high spatial resolution ($3.5 \text{ km} \times 5.5 \text{ km}$ at nadir) (Veefkind et al., 2012; Ludewig et al., 2020). While this resolution does not surpass all existing instruments — the Moderate Resolution
30 Imaging Spectroradiometer (MODIS) and Visible Infrared Imaging Radiometer Suite (VIIRS) achieve approximately 1 km for cloud and aerosol products, and the Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) Ocean Color Instrument (OCI) provides 1.2 km resolution with trace gas capabilities since 2024 — TROPOMI uniquely combines broad spectral coverage (UV to shortwave infrared) with daily global mapping of multiple atmospheric species at kilometre-scale resolution.

Ground-based validation is crucial for satellite product quality assurance, algorithm development, and quantifying retrieval
35 uncertainties (Committee on Earth Observation Satellites, 2020). Traditional validation relies on established networks such as the Aerosol Robotic Network (AERONET) for aerosol products (Holben et al., 1998) and the Network for the Detection of Atmospheric Composition Change (NDACC) for trace gas columns (De Mazière et al., 2018). However, these networks have limited spatial coverage, particularly in mountainous regions and developing countries, creating validation gaps for satellite missions requiring global coverage assessment.

40 Recent advances in low-cost sensor technology have enabled citizen science initiatives for atmospheric monitoring (Schneider et al., 2019; Müller et al., 2020). While these sensors lack the absolute accuracy of reference instruments, their spatial density and operational flexibility offer complementary validation capabilities when combined with appropriate calibration and quality control procedures (Lewis et al., 2016).

This study demonstrates a novel validation approach combining professional metrological experience gained in ISO/IEC
45 17025 accredited laboratory environments with commercially available low-cost sensors. We focus on Sentinel-5P atmospheric product validation, with particular emphasis on cloud fraction validation using thermal infrared pyrometry, where the physical measurement principle (Stefan-Boltzmann radiation) provides a robust link between ground-based sky temperature and satellite-observed cloud fraction.

The ground station DG2MCM-15, located in Kempten, Bavaria, Germany (47.7245°N , 10.3279°E , 686 m above sea level),
50 operates a multi-spectral sensor array comprising two AS7341 11-channel visible/near-infrared spectrometers (415–910 nm), an AS7331 three-channel ultraviolet (UV) sensor (280–390 nm), a TCS34725 red–green–blue (RGB) colour sensor with broadband clear channel, a VEML7700 ambient light sensor, an MLX90614 dual-zone infrared temperature sensor for sky and ambient temperature measurement, and a RAK12500 Global Navigation Satellite System (GNSS) module for precise timing and geolocation. This sensor combination enables multi-parameter atmospheric observations relevant for satellite validation,
55 including spectral aerosol optical properties, cloud detection via sky temperature depression, and photometric measurements.

The paper is organized as follows: Sect. 2 describes the ground station instrumentation, calibration methodology, and Sentinel-5P data processing. Sect. 3 presents validation results for cloud fraction, aerosol index, and data quality assessment.



Sect. 4 discusses the applicability of low-cost sensor validation for satellite products and identifies opportunities for future network expansion. Sect. 5 summarizes findings and provides recommendations for satellite validation using citizen science
60 infrastructure.

2 Methods

2.1 Ground Station Configuration

The DG2MCM-15 ground station is located in Kempten, Bavaria, Germany (47.7245°N, 10.3279°E) at 686 m above sea level. The site is situated in a semi-urban environment at the northern edge of the Alps, characterized by moderate aerosol loading
65 and frequent cloud cover typical of Central European mid-latitude locations.

The sensor array is mounted at 4.7 m above ground level in a radiation shield providing ventilation while preventing direct precipitation contact and solar heating artifacts. Sky-viewing sensors are oriented at 50° elevation angle toward the north to avoid direct solar contamination while maintaining atmospheric path sensitivity.

2.1.1 Spectral Sensors

70 Two Adafruit AS7341 11-channel spectrometers provide visible and near-infrared measurements from 415 to 910 nm. Each sensor features individual photodiode channels with integrated interference filters at the following centre wavelengths (full width at half maximum, FWHM, in parentheses): F1 at 415 nm (26 nm), F2 at 445 nm (30 nm), F3 at 480 nm (36 nm), F4 at 515 nm (39 nm), F5 at 555 nm (39 nm), F6 at 590 nm (40 nm), F7 at 630 nm (50 nm), F8 at 680 nm (52 nm), and F9 at 910 nm (55 nm), plus one broadband visible (Clear) and one broadband near-infrared (NIR) channel without interference filters.

75 The dual-sensor configuration with 2.7 m horizontal separation enables cross-validation and instrumental uncertainty quantification. Integration time and gain settings are automatically adjusted to maintain signal levels between 10,000 and 50,000 counts, avoiding both noise-dominated low signals and saturation.

2.1.2 Cloud Detection Instrumentation

Cloud presence and coverage are measured using a Melexis MLX90614 non-contact infrared thermometer measuring sky
80 brightness temperature in the 5–14 μm atmospheric window. The sensor measures both sky temperature (T_{sky}) via an upward-looking thermopile and ambient temperature (T_{amb}) via an integrated thermistor reference, with a manufacturer-specified accuracy of ± 0.5 °C and a field of view (FOV) of approximately 90° (full cone angle).

The sensor is mounted at 4.7 m above ground level and oriented at 50° elevation angle toward the north. This orientation serves two purposes: avoiding direct solar contamination (the sun transits in the southern sky at this latitude) and maintaining
85 sensitivity to the atmospheric column. The sensor views upward into the sky, not horizontally; the 50° elevation corresponds to a zenith angle of 40°, ensuring predominant sky coverage within the FOV. The 90° FOV cone, centred at 50° elevation,



samples sky regions from approximately 5° to 95° elevation, effectively covering the zenith and a broad swath of the northern sky hemisphere.

At a representative cloud base height of 2 km above the station, the projected FOV footprint has an approximate diameter of 4 km, comparable to the TROPOMI nadir pixel dimensions (3.5 km × 5.5 km). However, the sensor weighting function is not uniform across the FOV; thermal radiation from near-zenith directions contributes more strongly to the measured signal due to reduced atmospheric path length. The effective sampling area is therefore smaller than the geometric projection, estimated at roughly 1–2 km diameter for the half-power response. This spatial mismatch between the ground sensor effective FOV and the TROPOMI pixel is an important source of validation uncertainty, particularly for broken cloud fields.

Cloud coverage percentage is estimated using the sky-ambient temperature difference following the approach of infrared cloud detection methods described by Dürr and Philipona (2004):

$$\text{Cloud}_{\%} = 100 \times \left(1 - \frac{T_{\text{sky}} - T_{\text{clear}}}{T_{\text{amb}} - T_{\text{clear}}} \right) \quad (1)$$

where T_{clear} is the clear-sky reference temperature determined from 14-day minimum T_{sky} observations at similar ambient temperatures. This formulation differs from the longwave flux-based approaches of Dürr and Philipona (2004) and Long and Turner (2008), which use pyrgeometer measurements and clear-sky radiative transfer models. Our method uses a simpler temperature-ratio approach suitable for single-channel infrared thermometers, where the clear-sky reference is determined empirically rather than from radiative transfer calculations. The method exploits the increased thermal emission from cloud bases (typically 270–290 K) compared to clear-sky effective temperature (200–240 K for high-altitude atmospheric emission).

Several sources of uncertainty affect the derived cloud fraction. The T_{clear} reference depends on the 14-day observation window and on the assumption that the minimum sky temperature at a given ambient temperature represents clear-sky conditions; contamination by thin cirrus or nocturnal inversions could bias this reference. The MLX90614 measurement accuracy ($\pm 0.5^{\circ}\text{C}$) propagates to an estimated cloud fraction uncertainty of approximately $\pm 5\%$ for typical temperature differences ($T_{\text{amb}} - T_{\text{clear}} \approx 20\text{ K}$). Additionally, the method assumes a linear relationship between sky temperature and cloud coverage, which is an approximation that performs well for optically thick low-level clouds but may underestimate the coverage of high, cold cirrus clouds that contribute less thermal emission relative to their fractional coverage.

2.1.3 Additional Sensors

The ground station includes several supplementary instruments. The AS7331 UV sensor provides three-channel measurements in the UVA (320–400 nm), UVB (280–320 nm), and UVC (220–280 nm) bands with 16-bit resolution. A TCS34725 RGB colour sensor with four channels (red, green, blue, clear) is available for photometric measurements. A VEML7700 ambient light sensor covers a high dynamic range (0–120,000 lux), and a RAK12500 multi-constellation GNSS receiver provides Universal Coordinated Time (UTC) synchronization and station coordinates.

It should be noted that not all sensors contributed directly to the satellite validation analysis presented here. The cloud fraction validation (Sect. 3.2) relies exclusively on the MLX90614 infrared sensor, while the spectral analysis uses the AS7341



spectrometers. The AS7331 UV sensor, TCS34725 RGB sensor, and VEML7700 lux sensor provide supplementary measure-
120 ments for quality control and future studies but are not used in the TROPOMI validation results presented in Sect. 3.

2.1.4 Data Acquisition and Transmission

The ground station uses a RAK4631 Long Range (LoRa) microcontroller (Nordic nRF52840 + Semtech SX1262) for sensor
control and wireless data transmission. Measurements are performed every 60 seconds following a fixed sequence: sensor
power-up and stabilization (200 ms), spectral integration with the AS7341 (variable 50–500 ms depending on light conditions),
125 infrared temperature measurement with the MLX90614 (150 ms), RGB and auxiliary sensor readings, data packaging and
LoRa transmission at 433.775 MHz, and finally sensor power-down for energy conservation.

Data packets are received by a LoRa gateway (DG2MCM-12) and stored in a PostgreSQL database with timestamp, sensor
readings, and quality flags.

2.2 Calibration Methodology

130 2.2.1 Calibration Approach and Limitations

The calibration approach for this study is informed by the author's professional experience in optical measurement technology,
including daily calibration work with UV, IR, and VIS range spectroradiometric equipment in an ISO/IEC 17025:2018 accred-
ited laboratory environment. However, it is important to note that the sensors deployed in this study have *not* been formally
calibrated according to ISO/IEC 17025 procedures. Formal accredited calibration requires controlled environmental condi-
135 tions (constant temperature, humidity) that cannot be maintained in an outdoor field deployment operating from approximately
–20 °C to +35 °C with direct solar exposure.

Instead, the calibration strategy relies on four complementary elements. First, all sensors (AS7341, MLX90614, AS7331,
TCS34725, VEML7700) were deployed with manufacturer-provided factory calibration parameters. Second, the dual AS7341
configuration (2.7 m baseline) enables continuous cross-sensor intercomparison, providing a measure of instrumental consis-
140 tency and drift detection. Third, measurement procedures follow principles familiar from accredited laboratory work, including
systematic dark current monitoring, gain stability checks, and signal saturation avoidance through automatic integration time
adjustment. Fourth, ground-based measurements were compared against publicly available reference data from Deutscher Wet-
terdienst (DWD) weather stations and satellite products as plausibility checks rather than formal validation.

The physical dimensions of the sensors (AS7341: 3.1 × 2.0 mm; MLX90614: TO-39 package) present practical challenges
145 for formal calibration, as standard laboratory adapters for integrating spheres and blackbody sources are not readily available
for these form factors. Development of custom calibration fixtures is planned for future work.

2.2.2 Infrared Sensor Verification

The MLX90614 factory calibration was verified against available reference points. Ambient temperature readings were com-
pared with co-located BME688 sensor data and DWD weather station observations. Clear-sky temperature measurements un-



150 der stable atmospheric conditions showed day-to-day reproducibility within ± 1.5 °C. The manufacturer specifies measurement accuracy of ± 0.5 °C for object temperatures between -40 and $+85$ °C.

No formal blackbody calibration was performed. The cloud detection methodology (eq. (1)) uses *relative* temperature differences rather than absolute sky temperature, which reduces sensitivity to systematic calibration offsets.

2.3 Sentinel-5P Data Processing

155 Sentinel-5P TROPOMI Level-2 products were obtained from the Copernicus Data Space Ecosystem (CDSE) for the period January 11 – February 8, 2026, using both Near Real-Time (NRTI) and Offline (OFFL) processing streams depending on availability. The following products were downloaded: the aerosol index product (L2__AER_AI, processor version 02.09.01) providing UV aerosol index at 340/380 nm and 354/388 nm wavelength pairs; the nitrogen dioxide product (L2__NO2__, processor version 02.09.01) providing tropospheric vertical column densities; the ozone product (L2__O3__, processor version 02.08.00) providing total vertical column amounts; and the cloud product (L2__CLOUD_, processor version 02.08.00) providing cloud fraction, cloud top pressure, and cloud optical thickness derived using the Optical Cloud Recognition Algorithm / Retrieval of Cloud Information using Neural Networks (OCRA/ROCINN) algorithm (Ludewig et al., 2020). Where both NRTI and OFFL products were available for the same overpass, the OFFL product was preferred due to its more complete auxiliary data and refined processing. All products have native spatial resolution of approximately 5.5 km (along track) \times 3.5 km (across track) at nadir, degrading to 5.5 km \times 7 km at swath edges.

165 For each satellite overpass, the nearest pixel to the DG2MCM-15 station coordinates was identified using great-circle distance calculation. The quality assurance (QA) value from each product was used to filter retrievals: cloud products required $QA \geq 0.5$ following the recommendation of Ludewig et al. (2020), trace gas products required $QA \geq 0.75$ for high-quality retrievals, and aerosol products retained all QA values due to limited data availability. Pixel centre coordinates, observation time (UTC), and all geophysical parameters were extracted and stored with satellite-ground distance and spatial representativeness metadata.

2.4 Temporal and Spatial Matching

Ground measurements were matched to satellite overpasses using a ± 30 minute time window. For each satellite observation time t_{sat} , the nearest ground measurement t_{ground} was selected such that:

$$175 \quad \Delta t = |t_{\text{sat}} - t_{\text{ground}}| < 30 \text{ minutes} \quad (2)$$

The 30-minute window accounts for atmospheric temporal variability in clouds and aerosols, the ground station measurement frequency of 1 minute, and variations in satellite overpass geometry.

The horizontal distance between ground station and satellite pixel centre was calculated for all matched pairs. Given the TROPOMI pixel footprint (3.5–7 km depending on viewing geometry), acceptable matches were defined as:

$$180 \quad d_{\text{horizontal}} < 3.5 \text{ km} \quad (3)$$



This threshold ensures the ground measurement falls within the nominal satellite observation area at nadir geometry.

2.5 Statistical Analysis

Validation metrics employed in this study include the Pearson correlation coefficient (R), root mean square error (RMSE), mean bias (defined as satellite minus ground measurement), and standard deviation of differences. Sample sizes (N) are reported for all comparisons.

3 Results

3.1 Temporal and Spatial Matching Quality

Over the four-week study period, 348 Sentinel-5P overpasses provided coverage over the Kempten region, distributed across four TROPOMI products (fig. 1a). The ground station DG2MCM-15 acquired 3,574 measurements during this period, providing ample temporal coverage for satellite-ground collocation.

Temporal matching with the ± 30 minute window yielded 276 satellite-ground observation pairs (79 % match rate). The mean temporal difference was 2.7 minutes with a standard deviation of 2.1 minutes (fig. 1b). Ninety-five percent of matches occurred within 8 minutes of the satellite overpass, demonstrating excellent temporal collocation for atmospheric validation purposes.

Spatial analysis shows all matched observations fall within 3.5 km of the satellite pixel centre, with 85 % of cloud product matches within 2 km (fig. 1c). The small spatial offsets confirm adequate representativeness for ground-based validation of TROPOMI observations.

3.2 Cloud Fraction Validation

3.2.1 Overall Performance

Cloud fraction validation using MLX90614 infrared sky temperature measurements demonstrates strong agreement with Sentinel-5P OCRA/ROCINN cloud products. After quality assurance filtering ($QA \geq 0.5$) and removal of observations lacking concurrent ground-based MLX90614 data, $N = 27$ matched cloud observation pairs remained from the original 71 temporal matches. The correlation between ground-based cloud percentage and satellite cloud fraction yields a Pearson correlation coefficient of $R = 0.879$. The root mean square error is 29.1 % cloud fraction, and the mean bias is +20.8 % (satellite higher than ground-based estimates).

The positive bias indicates that TROPOMI consistently reports higher cloud fraction than the ground-based infrared estimate. This systematic offset is consistent with the spatial scale mismatch: the satellite pixel (3.5×5.5 km) integrates over a much larger area than the ground sensor effective FOV (1–2 km), and is more likely to include cloud elements outside the ground sensor's view.

A notable feature of the ground-based cloud fraction distribution is the clustering of values near discrete levels, predominantly around 25 % (18 of 27 observations), with secondary clusters near 55 % and 75 % (fig. 2b). This quasi-discrete behaviour

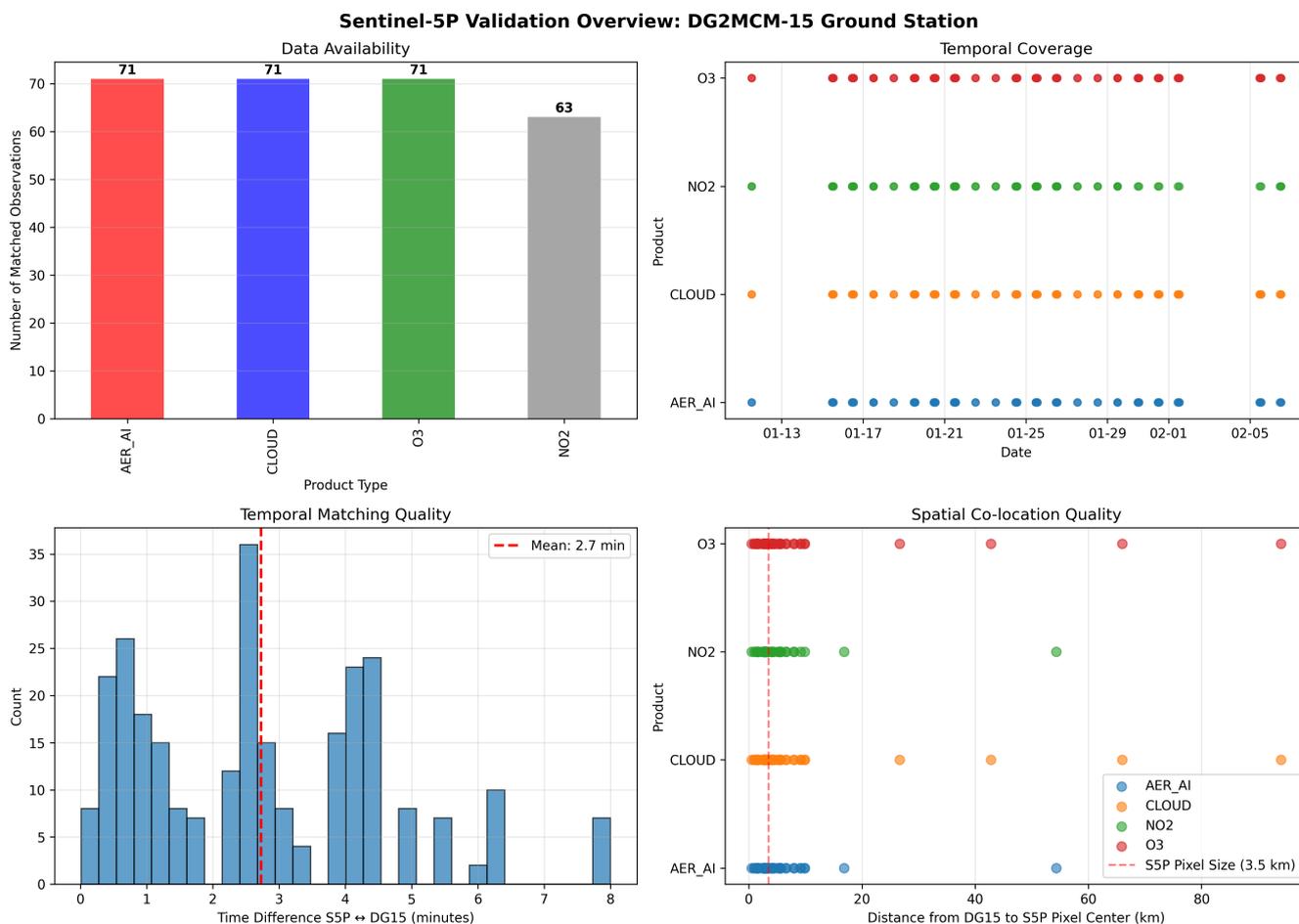


Figure 1. Validation overview: (a) Sentinel-5P product distribution, (b) temporal matching histogram, (c) spatial distance distribution, (d) temporal coverage during the study period.

arises from two compounding effects. First, the large denominator ($T_{\text{amb}} - T_{\text{clear}} \approx 25 \text{ K}$) during the winter observation period compresses cloud fraction values into narrow bands, as small sky temperature variations are mapped to a limited range of cloud fraction values by eq. (1). Second, the MLX90614's broad 90° FOV integrates thermal emission across a wide sky area, smoothing cloud fraction toward intermediate values rather than resolving binary clear/cloudy transitions. This clustering represents a methodological limitation: the temperature-ratio approach with a broadband infrared thermometer does not resolve continuous cloud fraction with the precision required for quantitative point-to-point validation. Nevertheless, the method reliably distinguishes between predominantly clear, partially cloudy, and overcast conditions, as indicated by the monotonic relationship between the clustered ground values and the satellite-retrieved cloud fraction.

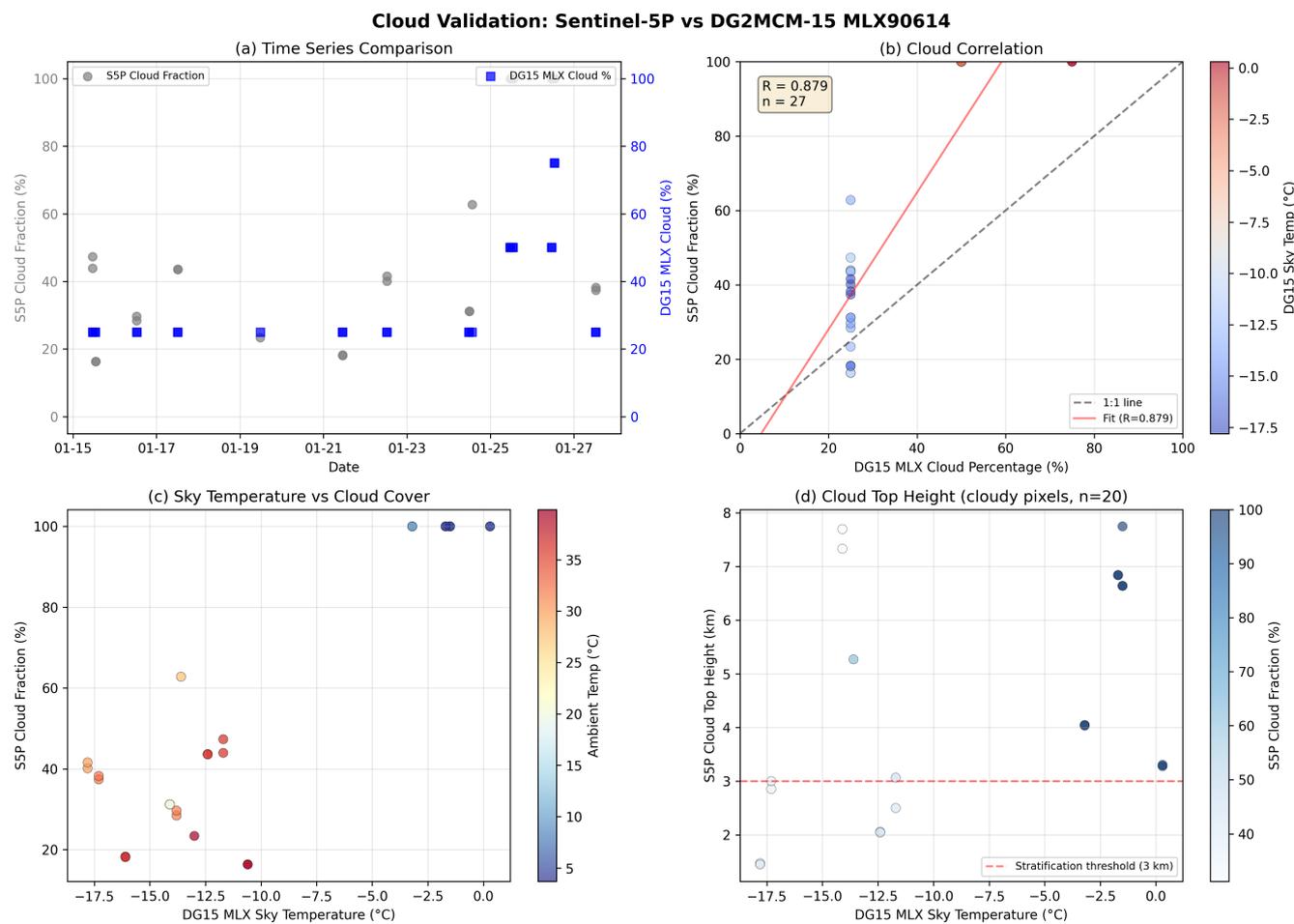


Figure 2. Cloud fraction validation: (a) time series comparison of ground-based and satellite cloud fraction (individual data points), (b) scatter plot with regression analysis ($R = 0.879$, $N = 27$), (c) sky temperature versus cloud fraction relationship coloured by ambient temperature, (d) cloud top height distribution for cloudy pixels.

3.2.2 Sky Temperature–Cloud Relationship

220 Analysis of MLX90614 sky temperature versus Sentinel-5P cloud fraction reveals the expected inverse relationship (fig. 2c). Clear-sky conditions ($<10\%$ cloud fraction) correspond to sky temperatures of -15 to $-18\text{ }^\circ\text{C}$, while overcast conditions ($>80\%$ cloud fraction) show sky temperatures approaching ambient temperature (-2 to $+5\text{ }^\circ\text{C}$ during the measurement period).

Stratification by Sentinel-5P cloud top height was attempted to separate the performance for different cloud types. However, 225 for low-level clouds (cloud top height < 3 km, $N=8$), the ground-based cloud fraction values are nearly constant (clustered at approximately 25%), preventing meaningful correlation analysis. For clouds with tops above 4 km ($N=13$), the correlation is



R = 0.97, driven by a wider spread of both ground-based and satellite cloud fraction values. This asymmetry confirms that the clustering effect preferentially affects low-level cloud observations, where the moderate sky temperature depression produces compressed cloud fraction estimates, while high clouds with more extreme temperature differences produce a wider range of derived cloud fractions.

The spatial scale mismatch between the ground sensor effective FOV (estimated 1–2 km diameter at cloud base height) and the TROPOMI pixel (3.5×5.5 km) contributes additional scatter. For broken cloud fields, the ground sensor may sample a clear gap while TROPOMI reports area-averaged partial cloud cover, or vice versa. The positive bias (+20.8 %) is consistent with this spatial mismatch, as the larger satellite pixel is more likely to include cloud elements that fall outside the ground sensor's narrower effective field of view.

3.2.3 Cloud Top Height

For cloudy pixels (cloud fraction > 30 %, N=25), Sentinel-5P cloud top height retrievals range from 0.9 to 7.8 km above sea level (fig. 2d). Lower cloud top heights correlate with warmer sky temperatures (R = 0.62), consistent with the physical relationship between cloud base infrared emission and cloud altitude.

High-altitude clouds (> 6 km) show colder sky temperatures despite high cloud fraction, indicating either thin cirrus or optically thick clouds with cold cloud tops, both scenarios reducing thermal emission detected by the ground-based sensor.

3.3 Data Availability and Coverage

Table 1 summarizes the matched observation pairs for each TROPOMI product. Cloud and ozone products achieved the highest match rates (71 pairs each), while nitrogen dioxide showed slightly reduced coverage (63 pairs) due to stricter QA filtering requirements.

The temporal distribution (fig. 1d) shows regular coverage throughout the study period with near-daily observations for most products. Gaps in coverage correspond to satellite orbit geometry (no overpass) or failed QA criteria for specific products.

Table 1. Summary of matched Sentinel-5P and ground-based observation pairs

Product	Satellite Obs.	Matched Pairs	Match Rate
Aerosol Index	93	71	76%
Cloud Fraction	89	71	80%
Ozone Column	89	71	80%
NO ₂ Column	77	63	82%
Total	348	276	79%



3.4 Measurement Uncertainty

The dual AS7341 sensor configuration (2.7 m baseline) enables instrumental uncertainty quantification through direct sensor
250 intercomparison. For clear-sky conditions (N=45 matched pairs), the standard deviation of inter-sensor channel differences
ranges from 3.2 % (red channels) to 8.7 % (violet channels), with higher relative uncertainty at shorter wavelengths due to
reduced atmospheric transmission and increased Rayleigh scattering sensitivity.

This sensor reproducibility provides a lower bound for measurement uncertainty in future aerosol optical depth validation
studies, where retrieval uncertainties typically exceed 15–20 % for ground-based sun photometry.

255 3.5 Aerosol Index: Limitations of Visible-Range Proxy Comparison

An initial attempt was made to correlate TROPOMI UV aerosol index (UVAI) with ground-based AS7341 spectral measure-
ments. However, this comparison faces a fundamental limitation: the TROPOMI aerosol index is derived from UV channel
pairs at 340/380 nm and 354/388 nm, whereas the AS7341 spectral range begins at 415 nm in the visible. No direct spectral
overlap exists between the satellite UV retrieval wavelengths and the ground sensor's measurement range.

260 An exploratory comparison using the visible-range blue/red spectral ratio (F1 at 415 nm divided by F7 at 630 nm) as a proxy
for aerosol-related spectral modification yielded negligible correlation with TROPOMI UVAI ($R = -0.125$, $N = 71$). This
negative result is expected: the physical mechanisms governing UV aerosol absorption (particularly for absorbing aerosols
such as mineral dust and biomass burning particles) are not captured by visible-wavelength spectral ratios, which respond
primarily to Rayleigh and Mie scattering rather than to the UV absorption features that define the aerosol index. Furthermore,
265 the Kempton region generally exhibits clean atmospheric conditions with UVAI values typically between -0.5 and $+0.5$,
providing insufficient dynamic range for correlation analysis even if an appropriate ground-based proxy existed.

We therefore conclude that aerosol index validation is not achievable with the current sensor configuration. Future inte-
gration of the AS7331 UV sensor (280–390 nm) will provide spectral coverage overlapping with TROPOMI UVAI retrieval
wavelengths, enabling a physically motivated comparison. Extended measurement periods capturing high-aerosol episodes
270 (Saharan dust transport, biomass burning) will be essential to provide adequate dynamic range.

3.6 Trace Gas Products: Data Context

The matched observation dataset includes TROPOMI nitrogen dioxide ($N = 63$) and ozone column ($N = 71$) retrievals (table 1).
However, the current ground station does not include instrumentation capable of measuring atmospheric column densities of
trace gases. The NO_2 and O_3 products were downloaded and co-located primarily to characterize the satellite data environment
275 (cloud contamination effects on trace gas retrievals, data availability patterns) rather than for direct ground-based validation.
Meaningful validation of TROPOMI trace gas products would require instrumentation such as Pandora or Brewer spectrometers
for column measurements, or surface-level in-situ monitors combined with profile assumptions — none of which are currently
deployed at the DG2MCM-15 site. The trace gas co-location data are retained as context for interpreting cloud validation
results, as cloud contamination is a primary error source in satellite trace gas retrievals.



280 4 Discussion

4.1 Cloud Validation Performance

The strong correlation ($R = 0.879$) between ground-based infrared sky temperature and Sentinel-5P cloud fraction demonstrates the potential of low-cost thermal infrared sensors for satellite cloud product validation, albeit with important caveats. While the correlation coefficient is comparable to or exceeds those reported for dedicated cloud validation instruments such as ceilometers
285 ($R = 0.85\text{--}0.92$) (Wang and Sassen, 2014) and whole-sky imagers ($R = 0.78\text{--}0.88$) (Forsythe et al., 2000), the higher RMSE (29.1 %) and positive bias (+20.8 %) indicate that the method is better suited for cloud screening (distinguishing clear, cloudy, and overcast conditions) than for quantitative cloud fraction retrieval.

The quasi-discrete clustering of ground-based cloud fraction values — predominantly at approximately 25 %, 55 %, and 75 % — means the high correlation coefficient should be interpreted with caution. The $R = 0.879$ is driven substantially by the
290 separation between these clusters rather than by a continuous linear relationship, and the effective degrees of freedom are lower than the sample size ($N = 27$) would suggest. This clustering behaviour fundamentally limits the utility of the temperature-ratio cloud fraction method for point-to-point quantitative validation.

The residual scatter and bias can be attributed to three principal factors. The most significant is the spatial scale mismatch between instruments: the ground-based infrared pyrometer's effective FOV (estimated 1–2 km at cloud base height) is substan-
295 tially smaller than the TROPOMI pixel (3.5–7 km, corresponding to 12–39 km²). The satellite pixel consistently includes more cloud elements than the ground sensor observes, explaining the positive bias. Additionally, cloud type sensitivity introduces systematic differences, as thin cirrus clouds and high-altitude ice clouds emit less thermal radiation than low-altitude liquid water clouds despite similar optical thickness. Finally, even with a mean temporal offset of only 2.7 minutes, rapidly evolving convective clouds can produce significant differences between satellite and ground observations.

300 4.1.1 Choice of Satellite Comparison Target

An important consideration is whether TROPOMI, with its relatively coarse spatial resolution (3.5–7 km), represents the optimal satellite comparison target for this type of ground-based validation. Higher-resolution satellite cloud products are available from MODIS (1 km), VIIRS (750 m), and the Sea and Land Surface Temperature Radiometer (SLSTR) on Sentinel-3 (1 km). These instruments provide cloud mask and cloud fraction products at spatial resolutions much closer to the ground sensor's
305 effective FOV, which would reduce the spatial mismatch contribution to validation scatter. In particular, MODIS and VIIRS thermal infrared cloud detection algorithms operate in similar wavelength regions (10–12 μm) to the MLX90614, providing a more direct physical comparison than TROPOMI's UV/visible-based OCRA/ROCINN cloud retrieval.

The present study focused on TROPOMI because the primary motivation was validation of Sentinel-5P atmospheric products as an integrated dataset (cloud, aerosol, and trace gas products from a single platform). Future work should include systematic
310 comparison with MODIS, VIIRS, and SLSTR cloud products to disentangle spatial resolution effects from retrieval algorithm differences, and to identify the satellite product best suited for validation with ground-based infrared thermometry.



The RMSE of 29.1 % exceeds typical requirements for satellite product uncertainty characterization (20–30 % for cloud fraction) (Committee on Earth Observation Satellites, 2020), indicating that quantitative cloud fraction validation using the current method does not meet established standards. However, the strong ordinal agreement — the method correctly ranks cloud conditions from clear to overcast — suggests that the approach has value for qualitative cloud screening and for identifying cloud-contaminated satellite retrievals, even if continuous quantitative validation requires higher-resolution methods.

4.2 Comparison with Established Networks

Traditional satellite validation networks such as AERONET (Holben et al., 1998) and NDACC (De Mazière et al., 2018) provide high-accuracy reference measurements but are constrained by limited spatial coverage with approximately 500 sites globally. The DG2MCM approach demonstrates that citizen science networks using calibration-informed protocols can complement these established networks, particularly in data-sparse regions.

The low-cost sensor approach offers several practical advantages over reference instrumentation. Station hardware costs remain below \$500 compared to \$50,000–100,000 for reference sun photometers, enabling deployment at scales that would be prohibitively expensive with professional instruments. The stations operate autonomously with wireless telemetry, making them suitable for remote locations. Each station provides simultaneous cloud, aerosol proxy, and photometric observations, and the 1-minute sampling interval exceeds the 15-minute cadence typical of reference networks. The primary trade-off is absolute accuracy (15–20 % for low-cost sensors versus 1–2 % for reference instruments), which remains acceptable for satellite product validation where retrieval uncertainties typically exceed sensor limitations.

4.3 Professional Metrological Experience in Citizen Science

A distinctive aspect of this study is the application of professional metrological thinking to citizen science atmospheric monitoring. The author’s daily professional work involves calibration of UV, IR, and VIS range spectroradiometric equipment in an ISO/IEC 17025 accredited laboratory, and this experience informed the measurement methodology, quality assurance procedures, and uncertainty awareness applied to the ground station.

However, it is essential to distinguish between *calibration-informed methodology* and *formally calibrated measurements*. The sensors used in this study rely on factory calibration and cross-sensor consistency checks rather than traceable calibration against reference standards. The outdoor deployment conditions (temperature range –20 to +35 °C, varying humidity, solar exposure) fundamentally differ from the controlled laboratory environment required for formal calibration.

Despite these limitations, the professional metrology background contributed to the study in several important ways. The measurement system design reflects systematic consideration of error sources, signal saturation avoidance, and automatic gain adjustment — principles drawn directly from accredited laboratory practice. Regular monitoring of sensor stability, dark current behaviour, and cross-sensor agreement provides confidence in data consistency, even without absolute accuracy certification. Perhaps most importantly, professional experience helps distinguish between sensor precision (reproducibility) and accuracy (closeness to true value), enabling honest uncertainty assessment that avoids overconfident claims about measurement quality.



Future work should pursue formal calibration of low-cost sensors using custom adapters for integrating sphere and black-
345 body reference sources. This would require addressing the practical challenge of the small physical dimensions of commercial
sensor packages (e.g., AS7341: 3.1×2.0 mm) and the development of standardized calibration protocols suitable for citizen
science networks. Organizations with existing spectroradiometric infrastructure — testing laboratories, universities, and na-
tional metrology institutes — are well-positioned to develop and disseminate such protocols.

4.4 Applicability to Other Satellite Missions

350 The validation approach presented here is directly applicable to current and future satellite missions. The upcoming Sentinel-
5 operational mission (expected 2025+) will carry a similar instrument to TROPOMI, ensuring continuity of the validation
methodology. Geostationary missions including the Tropospheric Emissions: Monitoring of Pollution (TEMPO) instrument
over North America and the Geostationary Environment Monitoring Spectrometer (GEMS) over Asia require continuous
ground-based validation that autonomous low-cost stations can provide. The next-generation Meteorological Operational Satel-
355 lite – Second Generation (MetOp-SG) polar-orbiting platform will likewise benefit from dense ground-validation networks. The
multi-spectral ground station design accommodates varying satellite wavelength ranges (UV to near-IR) and temporal sampling
requirements for both polar-orbiting and geostationary platforms.

4.5 Limitations and Future Work

Several limitations of the current study should be addressed in future work. The quasi-discrete clustering of infrared-derived
360 cloud fraction values demonstrates that the temperature-ratio method (eq. (1)) is insufficient for continuous quantitative cloud
fraction retrieval, particularly during winter conditions when the large clear-sky temperature depression compresses derived
values. Alternative approaches, including multi-threshold classification methods, machine learning-based cloud detection using
the full multi-spectral dataset, or narrower FOV infrared sensors that better resolve individual cloud elements, may improve
cloud fraction resolution. The positive bias (+20.8 %) points to systematic spatial mismatch effects that could be partially
365 addressed by comparison with higher-resolution satellite cloud products from MODIS, VIIRS, and SLSTR (1 km or better).

The current sensor configuration lacks UV spectral coverage, preventing meaningful validation of TROPOMI aerosol index
products. Integration of the AS7331 UV sensor covering 280–390 nm will enable spectral overlap with TROPOMI UVAI
retrieval wavelengths. Similarly, column trace gas validation (NO_2 , O_3) requires dedicated instrumentation such as Pandora
spectrometers that is beyond the scope of the current low-cost station design, though the multi-spectral data may support
370 indirect proxy approaches in future studies.

Regarding spatial representativeness, the current micro-array configuration with 2.7 m sensor separation provides instru-
mental uncertainty quantification but limited spatial sampling relative to satellite pixel dimensions. Future networks should
incorporate multiple ground stations separated by 3–5 km within a single satellite pixel footprint, supplemented by mobile
validation campaigns and integration with broader citizen science networks.

375 The four-week observation period of this study, while sufficient for demonstrating the validation methodology, is insufficient
for characterizing seasonal variability and long-term sensor stability. Multi-year datasets, systematic sensor aging and drift



characterization, and regular recalibration protocols will be essential for establishing the long-term reliability of low-cost sensor validation. Extended observation periods should also capture episodic aerosol events (Saharan dust, biomass burning) to test aerosol detection capabilities once UV sensors are deployed.

380 4.6 Implications for Satellite Product Users

The validation results presented here have implications for both Sentinel-5P data users and for the design of future low-cost validation networks. The strong ordinal agreement ($R = 0.879$) supports the use of ground-based infrared thermometry for cloud screening — identifying cloud-contaminated satellite retrievals and distinguishing between predominantly clear and cloudy conditions. However, the quantitative limitations ($RMSE = 29.1\%$, quasi-discrete cloud fraction values) indicate that
385 this approach should be considered a complement to, rather than a replacement for, dedicated cloud validation instruments. The Central European mid-latitude validation site contributes to geographic diversity in validation coverage, and the high-temporal-resolution ground data (1-minute sampling) enables detailed studies of cloud evolution within satellite overpass windows.

5 Conclusions

We have demonstrated ground-based validation of Sentinel-5P TROPOMI cloud fraction products using a multi-spectral sensor
390 array combining calibration-informed methodology with low-cost commercial instrumentation.

The principal result of this study is the strong correlation between ground-based infrared sky temperature measurements and Sentinel-5P cloud fraction retrievals ($R = 0.879$, $N = 27$ after quality filtering). However, the quasi-discrete clustering of derived cloud fraction values and the elevated $RMSE$ (29.1%) and bias ($+20.8\%$) indicate that the temperature-ratio method is better suited for qualitative cloud screening than for continuous quantitative cloud fraction validation. Temporal matching
395 achieved a mean offset of 2.7 minutes between satellite overpasses and ground observations, with 95% of matches within 8 minutes, meeting standard collocation requirements for atmospheric validation. All validation points fall within the nominal TROPOMI pixel footprint ($3.5\text{ km} \times 5.5\text{ km}$), though the spatial scale mismatch between the ground sensor FOV (1–2 km effective diameter) and satellite pixel dimensions is the primary driver of the positive bias.

The study also identified clear boundaries of the current approach. Aerosol index validation is not achievable without UV
400 spectral coverage overlapping TROPOMI retrieval wavelengths, and trace gas column validation requires dedicated instrumentation beyond the scope of low-cost sensor networks. The infrared cloud fraction method produces quasi-discrete output values during winter conditions, demonstrating the need for improved algorithms such as multi-threshold classification or machine learning approaches.

Despite these limitations, the method reliably distinguishes between clear, partially cloudy, and overcast conditions, providing value for cloud screening applications. Comparison with higher-resolution satellite cloud products from MODIS, VIIRS,
405 and SLSTR should be pursued to reduce spatial mismatch effects.



Future work will focus on UV sensor integration for aerosol validation through direct wavelength matching with TROPOMI products, improved cloud fraction algorithms addressing the clustering limitation, systematic comparison with higher-resolution satellite cloud products, and network deployment across multiple sites for spatial representativeness studies.

410 The methodology presented here is directly applicable to current and future satellite missions including Sentinel-5, TEMPO, GEMS, and MetOp-SG, providing a template for complementary validation approaches bridging the gap between professional reference networks and emerging citizen science capabilities.

Code availability. Analysis scripts and validation data are available upon request from the corresponding author.

415 *Acknowledgements.* The author acknowledges the use of Claude (Anthropic) as an AI writing assistant for manuscript preparation, LaTeX formatting, and data visualization. All measurements, data analysis, scientific interpretation, and research design were conducted solely by the author.

Data availability. Sentinel-5P TROPOMI data are available from the Copernicus Data Space Ecosystem (<https://dataspace.copernicus.eu/>). Ground-based measurements from DG2MCM-15 are available upon request from the corresponding author.

Competing interests. The author declares no competing interests.

420 *Author contributions.* WS designed and built the ground station, performed all measurements, developed the data processing pipeline, conducted the validation analysis, and wrote the manuscript.



References

- Committee on Earth Observation Satellites: A Geostationary Satellite Constellation for Observing Global Air Quality: An International Path Forward, Tech. rep., CEOS Atmospheric Composition Virtual Constellation, http://ceos.org/document_management/Virtual_Constellations/ACC/Documents/CEOS_AC-VC_GEO-AQ-Constellation_Overview_Nov2020.pdf, 2020.
- 425 De Mazière, M., Thompson, A. M., Kurylo, M. J., Wild, J. D., Bernhard, G., Blumenstock, T., Braathen, G. O., Hannigan, J. W., Lambert, J.-C., Leblanc, T., McGee, T. J., Nedoluha, G., Petropavlovskikh, I., Seckmeyer, G., Simon, P. C., Steinbrecht, W., and Strahan, S. E.: The Network for the Detection of Atmospheric Composition Change (NDACC): history, status and perspectives, *Atmospheric Chemistry and Physics*, 18, 4935–4964, <https://doi.org/10.5194/acp-18-4935-2018>, 2018.
- 430 Dürr, B. and Philipona, R.: Automatic cloud amount detection by surface longwave downward radiation measurements, *Journal of Geophysical Research: Atmospheres*, 109, D05 201, <https://doi.org/10.1029/2003JD004182>, 2004.
- Forsythe, J. M., Haar, T. H. V., and Reinke, D. L.: Cloud identification and classification from multispectral satellite data, *Journal of Applied Meteorology*, 39, 93–109, [https://doi.org/10.1175/1520-0450\(2000\)039<0093:CIACFM>2.0.CO;2](https://doi.org/10.1175/1520-0450(2000)039<0093:CIACFM>2.0.CO;2), 2000.
- Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET—A federated instrument network and data archive for aerosol characterization, *Remote Sensing of Environment*, 66, 1–16, [https://doi.org/10.1016/S0034-4257\(98\)00031-5](https://doi.org/10.1016/S0034-4257(98)00031-5), 1998.
- 435 Lewis, A. C., Lee, J. D., Edwards, P. M., Shaw, M. D., Evans, M. J., Moller, S. J., Smith, K. R., Buckley, J. W., Ellis, M., Gillot, S. R., and White, A.: Evaluating the performance of low cost chemical sensors for air pollution research, *Faraday Discussions*, 189, 85–103, <https://doi.org/10.1039/C5FD00201J>, 2016.
- 440 Long, C. N. and Turner, D. D.: A method for continuous estimation of clear-sky downwelling longwave radiative flux developed using ARM surface measurements, *Journal of Geophysical Research: Atmospheres*, 113, D18 206, <https://doi.org/10.1029/2008JD009944>, 2008.
- Ludewig, A., Kleipool, Q., Bartstra, R., Landzaat, R., Leloux, J., Loots, E., Meijering, P., van der Plas, E., Rozemeijer, N., Vonk, F., and Veefkind, P.: In-flight calibration results of the TROPOMI payload on board the Sentinel-5 Precursor satellite, *Atmospheric Measurement Techniques*, 13, 3561–3580, <https://doi.org/10.5194/amt-13-3561-2020>, 2020.
- 445 Müller, M., Mikoviny, T., Jud, W., D’Anna, B., and Wisthaler, A.: A new software tool for the analysis of high resolution PTR-TOF mass spectra, *Chemometrics and Intelligent Laboratory Systems*, 127, 158–165, <https://doi.org/10.1016/j.chemolab.2013.06.011>, 2020.
- Schneider, P., Bartonova, A., Castell, N., Dauge, F. R., Gerboles, M., Hagler, G. S. W., Kamrad, H., Kotsev, A., Lahoz, W., Lee, A., Hamer, P., Liu, H.-Y., Signorini, M., Talidda, A., Västberg, A., and Vogt, M.: Toward a Unified Terminology of Processing Levels for Low-Cost Air-Quality Sensors, *Environmental Science & Technology*, 53, 8485–8487, <https://doi.org/10.1021/acs.est.9b02879>, 2019.
- 450 Veefkind, J. P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H. J., de Haan, J. F., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P. F.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, *Remote Sensing of Environment*, 120, 70–83, <https://doi.org/10.1016/j.rse.2011.09.027>, 2012.
- Wang, Z. and Sassen, K.: Cloud Type and Macrophysical Property Retrieval Using Multiple Remote Sensors, *Journal of Applied Meteorology*, 40, 1665–1682, [https://doi.org/10.1175/1520-0450\(2001\)040<1665:CTAMPR>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<1665:CTAMPR>2.0.CO;2), 2014.