



Linear tomographic retrieval characterisation and performance estimation for the ESA Earth Explorer 11 candidate CAIRT (Changing Atmosphere Infra-Red Tomography explorer)

Bernd Funke¹, Michael Höpfner², Stefan Bender¹, Quentin Errera³, Udo Grabowski², Alex Hoffmann⁴, Han-Li Liu⁵, Nick M. Pedatella⁵, Piera Raspollini⁶, Jörn Ungermann⁷, and Björn-Martin Sinnhuber²

¹Instituto de Astrofísica de Andalucía, CSIC, Granada, Spain

²Karlsruhe Institute of Technology, Institute of Meteorology and Climate Research (IMKASF), Karlsruhe, Germany

³Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium

⁴European Space Research and Technology Centre, European Space Agency (ESA/ESTEC), Noordwijk, The Netherlands

⁵High Altitude Observatory, National Center for Atmospheric Research, Boulder, Colorado, USA

⁶Instituto di Fisica Applicata “N. Carrara” (IFAC) del Consiglio Nazionale delle Ricerche (CNR), Firenze, Italy

⁷Forschungszentrum Jülich, Institute of Climate and Energy Systems - Stratosphere (ICE-4), Jülich, Germany

Correspondence: Bernd Funke (bernd@iaa.es)

Abstract. The Changing-Atmosphere Infra-Red Tomography Explorer (CAIRT) was proposed as a candidate for the European Space Agency’s eleventh Earth Explorer mission to address critical knowledge gaps in middle atmosphere dynamics and composition. Utilizing an innovative Imaging Fourier Transform Spectrometer (IFTS) in a limb-sounding geometry, CAIRT is designed to provide high-resolution, three-dimensional (3D) tomographic observations of the atmosphere from the free troposphere (5 km) to the lower thermosphere (115 km). To support the scientific evaluation and requirement consolidation of the mission, this paper introduces two computationally efficient performance modeling tools: the Linear Performance Estimator (LPE) and the Fast L2 Simulator (FL2S).

The Linear Performance Estimator (LPE) serves as a diagnostic tool that applies linear retrieval theory to quantify spatial resolution and retrieval uncertainties for atmospheric targets like temperature and trace gas concentrations. It features a novel extension to 2D tomographic retrievals, enabling the assessment of how instrument characteristics and uncertainties propagate into final Level-2 (L2) geophysical data products. The Fast Level-2 Simulator (FL2S) complements this by utilizing LPE-derived metrics to project high-resolution atmospheric model data fields onto a simulated CAIRT measurement grid, generating synthetic L2 data products that mimic the full retrieval chain. This allows for a streamlined assessment of how CAIRT’s spatial resolution and uncertainties affect specific scientific use cases.

Applying these tools to the CAIRT instrument characteristics, we demonstrate the mission’s capability to meet specified L2 and science performance requirements. The results confirm that the modular end-to-end simulation framework effectively translates instrument specifications into robust scientific performance metrics, providing a mature basis for possible future implementations of limb-imaging space missions.



1 Introduction

20 The Changing-Atmosphere Infra-Red Tomography Explorer (CAIRT), recently evaluated as a candidate for the European Space Agency's (ESA) eleventh Earth Explorer (EE11) mission, aims to address critical gaps in our understanding of the middle atmosphere (Sinnhuber et al., 2026). By utilizing an innovative Imaging Fourier Transform Spectrometer (IFTS) in a limb-sounding geometry, CAIRT is designed to provide unprecedented three-dimensional (3D) tomographic observations of atmospheric temperature and composition in the vertical domain from the free troposphere at about 5 km altitude through to the
 25 lower thermosphere at about 115 km altitude. Although CAIRT was not selected for implementation in the EE11 competition, the mission concept was assessed as scientifically and technically mature, and could still be implemented within a different programmatic framework or through international collaboration.

To support science performance evaluation and requirement consolidation of the mission, robust as well as time-effective performance modeling is required to translate instrument specifications into scientific performance metrics and vice versa. This
 30 paper describes two pivotal tools developed for this purpose during the candidate mission's definition and feasibility study phases (Phases 0/A): the Linear Performance Estimator (LPE) and the Fast Level-2 (L2) Simulator (FL2S, where Level-2/L2 denotes the retrieved geophysical data product level). Furthermore, applying these tools on the basis of CAIRT's instrument characteristics demonstrates the mission's expected performance for a range of atmospheric parameters.

Infrared imaging Fourier-transform spectroscopy has been successfully employed in atmospheric temperature and compo-
 35 sition limb-sounding from aircraft and balloons for more than a decade (Friedl-Vallon et al., 2014; Riese et al., 2014) and is also the baseline for the nadir-viewing Infrared Sounder (IRS) on board the Meteosat Third Generation geostationary satellite. Due to the dense along-flight-track spacing of the limb-measurements with frequently overlapping lines-of-sight (Fig. 1), tomographic retrieval schemes are necessary to exploit the full content of information from the radiance spectra as well as for characterizing the instrument's performance. The complexity of tomographic retrievals – which involve the simultaneous
 40 inversion of thousands of spectral measurements to reconstruct the atmosphere's vertical and horizontal structure – presents a significant computational challenge (Krasauskas et al., 2019). Further, infrared emission sounding above the stratosphere requires the consideration of non-thermal local thermodynamic equilibrium (non-LTE) which implies the use of sophisticated and time-consuming non-LTE population models (Funke et al., 2012). During the early design and trade-off phases, traditional full-physics retrieval simulations are thus often too computationally expensive for large-scale sensitivity studies or global
 45 performance assessments. The LPE and FL2S were developed as high-efficiency alternatives to bridge this gap.

The LPE serves as the primary tool for rapid diagnostic evaluation. By applying linear retrieval theory (Rodgers, 2000), the LPE quantifies spatial resolution and retrieval uncertainties of the mission's target species. It allows for the immediate assessment of how changes in instrument parameters, such as noise-equivalent spectral radiance (NESR) or pointing knowl-
 50 edge, propagate into the final Level 2 (L2) products, i.e., vertical and horizontal distributions of atmospheric quantities like temperature and concentrations of trace gases and aerosols. Linear retrieval theory is also a standard approach for efficient L2 characterization and error reporting in remote sensing applications (von Clarmann et al., 2020). The novelty as used for CAIRT performance estimation resides in the extension to tomographic 2D retrievals.



Complementing the LPE, the Fast Level-2 Simulator (FL2S) provides a more comprehensive, yet computationally stream-
 lined, simulation environment. The FL2S utilizes the linear approximation of the L2 retrieval process to project “true” atmo-
 spheric states (e.g., from high-resolution global models) onto the CAIRT measurement grid. By incorporating the 2D averaging
 55 kernels and stochastic noise components, the FL2S generates synthetic L2 data products that mimic the expected output of the
 full retrieval chain. This supports the assessment of the extent to which CAIRT’s spatial resolution and L2 retrieval uncertain-
 ties affect specific scientific applications. Thanks to its computational efficiency, FL2S was used for generating extended time
 series of simulated observations that can be injected into Observation System Simulation Experiments to assess the scientific
 60 or application potential of future data (e.g., Errera et al., 2026).

The LPE and FL2S complement the development of the CAIRT End-to-End Performance Simulator (CEEPS, Kleinert et al.,
 in preparation) which, by simulating the entire physics of the instrument with all its internal dependencies and performing full
 iterative 2D retrievals, is computationally intensive, thus limiting its use for establishing comprehensive uncertainty budgets
 and for providing larger data volumes of simulated L2 products to be used in science impact assessments.

65 Section 2 of this paper details the mathematical frameworks of both tools. By application to the CAIRT concept, the mission’s
 capacity to meet L2 and science performance requirements is demonstrated in Section 3. A summary and conclusions are
 provided in Section 4.

2 Methodology

The CAIRT payload concept is based on an imaging FTIR spectrometer that measures limb-emitted spectral radiances over the
 70 range 718–2200 cm^{-1} (13.93–4.55 μm) at high spectral resolution. The instrument points backwards with respect to spacecraft
 flight direction, covering the altitude range from the upper troposphere to the lower thermosphere with a vertical sampling
 of about 1 km in the tangent plane. In the across-track direction, the instantaneous coverage extends up to approximately
 ± 250 km, with a sampling as fine as 25 km. Along track, a complete spectral image is acquired every 50 km (see also Fig. 1
 and Table 1). The Level-1b (L1b) product therefore consists of spectrally and radiometrically calibrated limb radiances resolved
 75 in the vertical, across-track, and along-track dimensions (Sinnhuber et al., 2026; European Space Agency, 2025).

The performance evaluation framework for the CAIRT mission relies on a modular, end-to-end simulation architecture
 designed to translate atmospheric variables observed with configurable instrument parameters into L2 product characteristics.
 This framework is structurally divided into four interconnected components: the climatological limb radiance Spectra and
 Jacobian Database (SJD), the L1b Simulator, and the evaluation modules comprising the Linear Performance Estimator (LPE)
 80 and the Fast Level-2 Simulator (FL2S). As illustrated in the structural scheme in Fig. 2, the SJD feeds forward-modeled
 radiances and Jacobians (i.e., the derivatives of the radiances with respect to the retrieved geophysical parameters) into the L1b
 Simulator, which imprints instrument-specific characteristics. L1b Simulator outputs are subsequently ingested by the LPE to
 analytically derive diagnostic metrics, which are then utilized by the FL2S to efficiently project high-resolution atmospheric
 model output data fields into simulated CAIRT L2 data products.

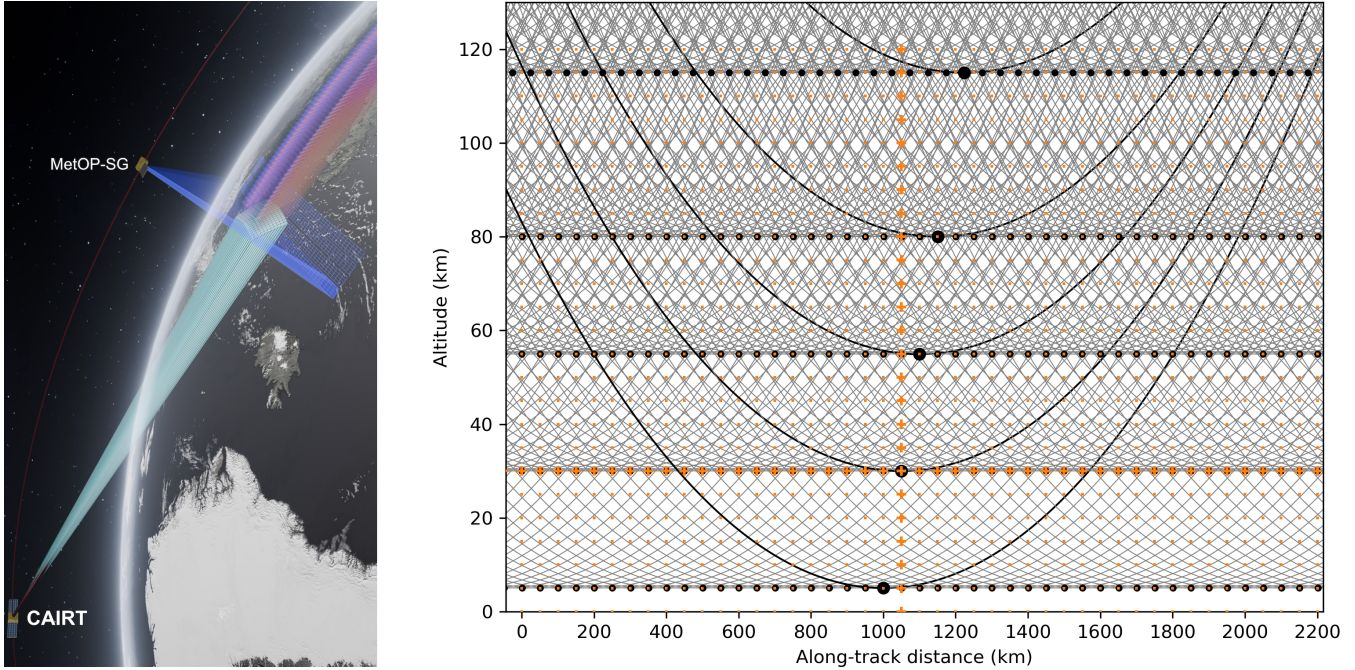


Figure 1. Left: Artist's impression of CAIRT observing concept, comprising the CAIRT limb sounder (rear-viewing) with its curved tangent point surface (cyan), and a simplified rendering of the IASI-NG across-track scanning pattern (blue). CAIRT was planned to be flown in loose formation with MetOp-SG. Right: 2D cross-section of the CAIRT limb-viewing geometry (grey and black) and retrieval grid (orange) at one across-track position. For clarity, only every fifth level of the vertical retrieval grid is shown. Similarly, only selected limb paths are displayed, with tangent altitudes of 5, 30, 55, 80, and 115 km. Grey and black lines denote limb lines-of-sight for successive acquisitions along-track for a rearward-looking instrument moving to the right. Black dots mark the tangent point positions of these limb paths. Large orange crosses in horizontal direction mark retrieval grid-points fixed to the tangent-point positions at 30 km altitude. In the vertical they indicated the retrieved profile fixed by the selected tangent point at 30 km (black dots and lines).

85 To mitigate the immense computational complexity inherent to full 3D atmospheric tomographic inversions, the retrieval strategy employs a dimensional decomposition. The across-track measurement swath is discretized into independent 2D vertical slices oriented along the instrument's line-of-sight track. Within each 2D slice, the inversion problem is further discretized over a custom retrieval grid defined in terms of n_h along-track positions and n_z vertical altitude levels (see Fig. 1, orange symbols). Note that the along-track and vertical spacing of the retrieval grid does not necessarily coincide with the measurement spacing

90 with m_z tangent points in the vertical (determined by the detector matrix) and m_h acquisitions (taken at subsequent times) defining the horizontal measurement grid and determined by the spacecraft's orbital velocity (black dots in Fig. 1). This 2D tomographic grid balances spatial resolution against processing overhead, ensuring the simultaneous inversion of overlapping lines of sight remains computationally tractable.

In order to further reduce complexity, the inversion problem is decomposed target-wise, similar to the processing carried

95 out for Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) on Envisat (von Clarmann et al., 2009). The

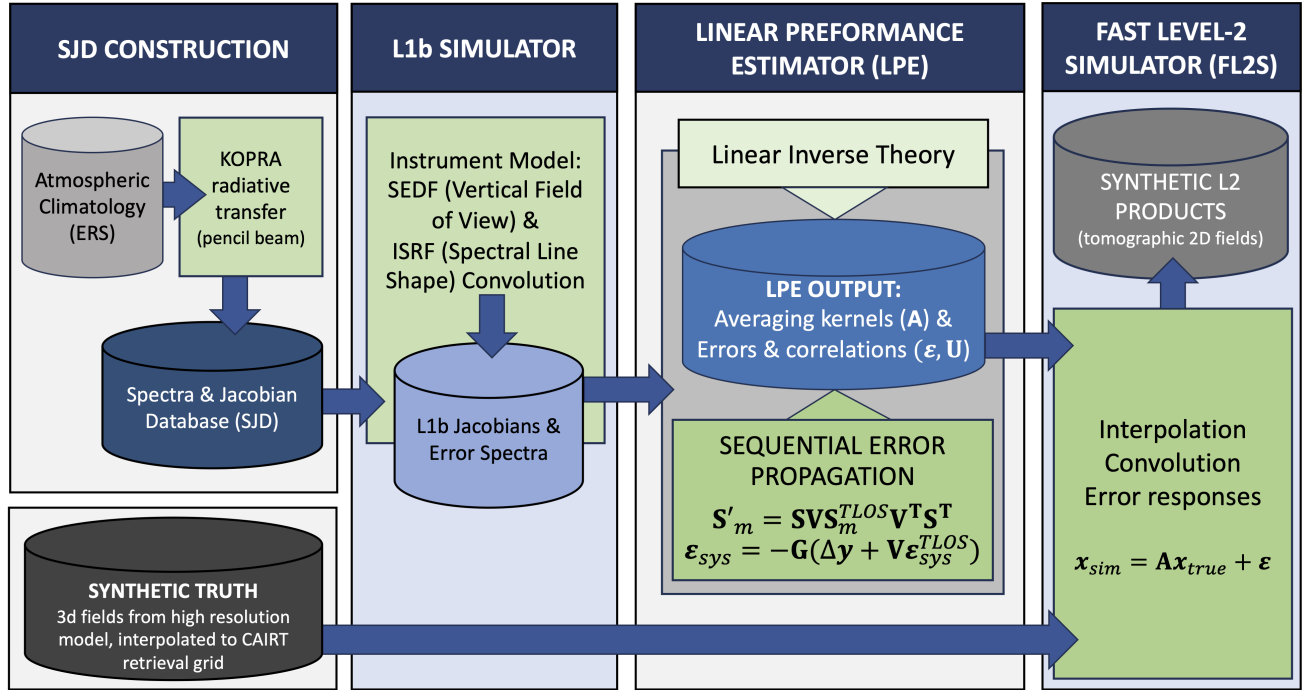


Figure 2. Structural scheme of CAIRT performance modeling. See text for more details.

sequential processing chain starts with a spectral calibration correction, followed by the joint retrieval of temperature and a line of sight (LOS) pointing correction. Afterwards, trace gas retrievals are performed, typically beginning with O_3 and H_2O . Error correlations are accounted for in the context of error propagation within the sequential retrieval chain. In this paper, we focus on temperature and LOS error propagation being most important for limb emission retrievals, as its impact cannot be avoided by dedicated spectral window selection.

2.1 Climatological limb radiance Spectra and Jacobian Database

For the generation of representative limb radiances, a climatology of temperature and composition of the middle atmosphere (the Extended Reference Scenarios, ERS) has been developed to describe the conditions that will be encountered by CAIRT (Er-rera et al., in prep.). This climatology contains all atmospheric parameters that affect spectral signatures within the measurement range (more than 30 chemical species). It includes vertical profiles of temperature and trace gases, as well as parameters needed for emitting states' population and radiance calculation in non-LTE conditions (see, e.g., García-Comas et al., 2023). The ERS was constructed for four representative months (January, April, July, October), five latitude bands (90°S – 70°S , 55°S – 35°S , 20°S – 20°N , 35°N – 55°N , 70°N – 90°N), two local overpass times (09:30 and 21:30), a vertical grid from 0 to 200 km (required



to simulate radiance contributions from nitric oxide in the lower thermosphere) in steps of 1 km, and considering the solar and volcanic activity for relevant species.

Based on this climatology, the Spectra and Jacobian Database (SJD) was generated using the line-by-line radiative transfer model KOPRA (Stiller, 2000) which allows for non-LTE radiance calculations by implicit calls of the Generic RAdiative traNsfer AnD non-LTE population algorithm (GRANADA, Funke et al., 2012). The SJD consists of sets of simulated broad-band infrared-emission radiance spectra $\tilde{\mathbf{y}}^{\text{SINC}}$ (i.e., arrays of $\tilde{m}_z$ spectra covering the instrument's vertical atmospheric coverage) as well as their Jacobians $\tilde{\mathbf{K}}^{\text{SINC}} = d\tilde{\mathbf{y}}^{\text{SINC}}/d\mathbf{x}$ of the radiances $\tilde{\mathbf{y}}^{\text{SINC}}$ with respect to retrieval parameters $\mathbf{x}$, for each of the atmospheric states of the ERS. The retrieval parameters are temperature and volume mixing ratios of trace gases at n_z different altitude and n_h along-track positions as basis for the 2D retrievals (or, optionally, at n_z different altitude levels to support 1D retrievals). Further, the SJD includes error disturbance spectra $\Delta\tilde{\mathbf{y}}^{\text{SINC}}$ for a variety of non-LTE parameters (García-Comas et al., 2023) as well as for uncertainties in the CO_2 -profile. All radiances are calculated in the spectral range between 707 cm^{-1} and 2203 cm^{-1} for an infinitesimal field of view, so-called 'pencil beam' spectra, for tangent altitudes with a spacing of 0.5 km from 0.5 km to 125 km to cover CAIRT's vertical observational range. This spacing is over-sampled compared to the instrument's vertical sampling in order to allow for the numerical simulation of field-of-view effects (see Sec. 2.2). In the spectral domain, the monochromatic radiances (calculated at $5 \times 10^{-4}\text{ cm}^{-1}$) are re-sampled at a wavenumber grid of 0.1 cm^{-1} using a sinc-function, corresponding to the goal spectral resolution of CAIRT.

The atmospheric parameters for which Jacobians are computed are discretized with a grid spacing of 0.5 km in the vertical and 50 km in the horizontal dimension. Note that the vertical grid spacing is finer than that of the retrieval grids used in the L2 processing. Jacobians are thus resampled in the LPE to the custom retrieval grid for each L2 target. Horizontal 2D-fields are constructed from the ERS by repetition of the same ERS altitude profiles over a horizontal region such that it contains the entire limb-paths for all viewing directions. The ERS atmospheric scenarios have been used to map a large part of the atmospheric variability. 2D Jacobians are calculated for the following parameters in non-LTE: temperature, H_2O , O_3 , CO , CH_4 , NO , NO_2 , gray-body extinction, and in LTE for targets in the 'lower' atmosphere where non-LTE effects can be neglected: C_2H_2 , C_2H_6 , CFC-11 , CFC-12 , HCFC-22 , ClONO_2 , HCN , HCOOH , HNO_3 , N_2O , N_2O_5 , NH_3 , PAN , SF_6 , SO_2 , H_2SO_4 -liquid, NH_4NO_3 -solid.

The Jacobians for the spectral shift, i.e. a shift in the spectral axis, are also determined already at this stage because the reduction of the spectra by the sinc-function convolution impedes a determination of spectral shift-derivatives within the L1b simulator. These Jacobians are required to simulate spectral calibration errors and to assess the performance of the spectral calibration correction as first step of the L2 retrieval chain.

2.2 The L1b simulator

Instrumental characteristics are imprinted to the contents of the SJD through the Level-1b (L1b, i.e. calibrated measurement) simulator module. This module applies vertical convolution to account for the vertical sampling and vertical System Energy Distribution Function (SEDF) as well as spectral convolution to account for the spectral sampling and apodized instrumental line shape (Apodized Instrument Spectral Response Function, AISRF). This is done for all the Jacobians and error spectra



Table 1. Basic instrument characteristics.

Parameter	Set-up / value
Altitude coverage of tangent points	5–115 km (unrefracted)
LOS ^a elevation angles	115 limb-views with constant elevation angle increments, corresponding to 1 km at 5 km tangent point (unrefracted)
SEDF ^b	3 km wide Gaussian with FWHM ^c 1.5 km (at 5 km tangent altitude)
Spectral sampling	0.2 cm ⁻¹
Maximum interferometric optical path difference	2.5 cm
AISRF ^d	Norton-Beer strong (Norton and Beer, 1976, 1977)

^aLine of sight, ^bSystem Energy Distribution Function, ^cFull Width at Half Maximum, ^dApodized Instrument Spectral Response Function.

already contained in the SJD. Furthermore, the instrument simulator determines new Jacobians with respect to vertical pointing, as well as additional error spectra regarding knowledge of the SEDF and knowledge of the AISRF. Table 1 summarizes the basic instrument characteristics used in the simulations described below.

As mentioned above, the ‘pencil beam’ limb spectral radiances, their Jacobians and error spectra in the SJD are stored as pure SINC-spectra $\tilde{\mathbf{y}}(\alpha, \tilde{\nu})$. Here α denotes the vertical elevation angle and $\tilde{\nu}$ the wavenumber. The L1b simulator convolves these spectra with the Norton-Beer strong (NBS) apodization function in spectral space (Norton and Beer, 1976, 1977) adapting and resampling either to the goal (0.1 cm⁻¹) or the threshold spectral resolution (0.2 cm⁻¹) resulting in $\tilde{\mathbf{y}}^{\text{NBS}}(\alpha, \tilde{\nu})$.

After that, a vertical spectrally and spatially dependent SEDF ($\alpha - \alpha_p, \alpha_p, \tilde{\nu}$) is applied to all spectra ($\tilde{\mathbf{y}}^{\text{NBS}} \rightarrow \mathbf{y}$), Jacobian spectra ($\tilde{\mathbf{K}}^{\text{NBS}} \rightarrow \mathbf{K} = d\mathbf{y}/d\mathbf{x}$), and error spectra ($\Delta\tilde{\mathbf{y}}^{\text{NBS}} \rightarrow \Delta\mathbf{y}$), as a convolution over the viewing angle range of the SEDF:

$$\mathbf{y}(\alpha_p, \tilde{\nu}) = \int_{-\alpha_{\max}}^{\alpha_{\max}} \tilde{\mathbf{y}}^{\text{NBS}}(\alpha, \tilde{\nu}) \text{SEDF}(\alpha - \alpha_p, \alpha_p, \tilde{\nu}) d\alpha. \quad (1)$$

Here α_p denotes the central viewing angle of each vertical detector pixel p with $p = 1 \dots m_z$ and $\pm\alpha_{\max}$ the angle range covered by the SEDF. The SEDF-shape is flexibly implemented such that it can depend on wavenumber $\tilde{\nu}$ as well as on vertical pixel position α_p .

Despite the application of the vertical SEDF and spectral apodisation to the contents of the SJD, the L1b simulator module also calculates numerically additional perturbation spectra for the spectral line shape by changing the width of the AISRF, and for the SEDF by either changing the SEDF width in the near-field or by perturbing the far-field SEDF in different regions depending on the distance to the SEDF center (for more information see Table 5).

Further, Jacobians with respect to the LOS pointing are computed by application of a vertically shifted SEDF to the reference spectra. These Jacobians are used in the LPE for the LOS (absolute pointing, i.e. the pointing of the entire detector) correction retrieval and to simulate relative pointing errors (i.e. distortions between pixel positions within the detector). Perturbation



spectra related to radiometric calibration errors are not computed by the L1b simulator but are simulated within the LPE (see Sec. 2.3).

165 2.3 Linear performance evaluation

A key element of CAIRT's L2 performance evaluation is the Linear Performance Estimator (LPE) which ingests the simulated Jacobians and perturbation spectra from the L1b Simulator. Based on linear retrieval theory (Rodgers, 2000), the LPE computes the 2D retrieval covariance matrix $\mathbf{S}$ and the 2D averaging kernel $\mathbf{A}$ as

$$\mathbf{S} = \left[\left(\sum_i \mathbf{K}_i^\top \mathbf{S}_y^{-1} \mathbf{K}_i \right) + \mathbf{R} \right]^{-1} \quad (2)$$

$$170 \quad \mathbf{A} = \mathbf{S} \left(\sum_i \mathbf{K}_i^\top \mathbf{S}_y^{-1} \mathbf{K}_i \right), \quad (3)$$

where $\mathbf{K}_i$ are the resampled 2D Jacobians provided by the L1b simulator, $\mathbf{R}$ describes the assumed regularization, $\mathbf{S}_y$ is the measurement noise covariance matrix. The sum in Equations (2) and (3) extends over all m_h along-track measurements i included in the 2D retrieval. Each measurement i comprises m_s spectra recorded at elevations corresponding to the vertical pixels of the image, yielding a total of m_s spectral points. The decomposition into a sum over individual measurements i conveniently
 175 avoids the handling of extremely large matrices of dimension $m = m_h \times m_s$. The measurement noise covariance matrix $\mathbf{S}_y$ (dimension $m_s \times m_s$) is computed from the wavenumber-dependent noise-equivalent spectral radiance (NESR) estimates and is independent on the along-track position i . Note that $\mathbf{S}_y$ is non-diagonal because of the applied NBS apodisation.

For the CAIRT 2D tomographic retrieval simulations, the chunk size m_h is set to 77 (corresponding to an along-track coverage of the LOS extension passing through the entire atmosphere at 200 km altitude). The dimension of the parameter
 180 space is $n = n_h \times n_z$ with n_z being the number of vertical profile points. The number of along-track retrieval gridpoints is equal to the number of along-track observations, $n_h = m_h$. The matrix elements of $\mathbf{K}_i$ with respect to the along-track parameters j (note that for each j there are n_z vertical parameters) are computed as $\mathbf{K}_{i,j} = \mathbf{K}_{i',j'}$ where i' corresponds to the along-track center measurement (for which the Jacobian is actually calculated in the L1b Simulator) and $j' = j + i - i'$. That is, the Jacobians corresponding to any measurement i are constructed from the pre-calculated Jacobian by shifting the elements of
 185 the latter in dependence of the along-track displacement of measurement i with respect to the along-track center measurement i' . It also implies that $\mathbf{K}_i$ is assumed to be independent of along-track variations in the atmospheric state. This is a reasonable assumption for sensitivity assessments conducted under climatological mean-state conditions.

To characterize the retrieval response, the LPE provides averaging kernel sub-matrices $\mathbf{A}'$ (dimension $n_z \times n$) built from the rows that correspond to the retrieval vector at the along-track center position j' . Spatial resolution is diagnosed from the
 190 reciprocal data density, which is the local grid width divided by the respective diagonal value of the averaging kernel matrix (Purser and Huang, 1993). Vertical resolution v is estimated from the horizontally integrated $\mathbf{A}'$ sub-matrix:

$$v(z_k) = \frac{\Delta z_k}{\sum_j A'_{k;j}} \quad (4)$$



where the summation is taken over all along-track positions j of the retrieval vector and Δz_k is the grid width at vertical level k . The horizontal (along-track) resolution h is then given by

$$195 \quad h(z_k) = \frac{\Delta z_k \Delta x}{v(z_k) A'_{k;k,j'}} \quad (5)$$

with Δx being the along-track grid width.

L2 uncertainty estimation follows the methodology used in MIPAS error assessment (von Clarmann et al., 2022), here applied to 2D tomographic retrievals. The mapping of measurement noise onto L2 products, ϵ_{noise} , is assessed by means of Gaussian error propagation. We compute the estimated standard deviation (ESD) of L2 noise errors from the noise error
 200 covariance matrix $\mathbf{S}_m$:

$$\mathbf{S}_m = \mathbf{A}\mathbf{S}, \quad ESD = \sqrt{\text{diag}(\mathbf{S}_m)}. \quad (6)$$

ESD is reported for the parameters j' (a vertical profile of dimension n_z) at the along-track center.

Propagation of measurement noise errors through the temperature and LOS (TLOS) retrieval onto constituent retrievals is considered by the additional term

$$205 \quad \mathbf{S}'_m = \mathbf{S}\mathbf{V}\mathbf{S}_m^{\text{TLOS}}\mathbf{V}^\top\mathbf{S}^\top, \quad (7)$$

where $\mathbf{S}_m^{\text{TLOS}}$ is the noise error covariance matrix of the preceding temperature and LOS joint retrieval, and

$$\mathbf{V} = \sum_i \mathbf{K}_i^\top \mathbf{S}_y^{-1} \mathbf{K}_i^{\text{TLOS}} \quad (8)$$

is an auxiliary matrix in the parameter space with $\mathbf{K}_i^{\text{TLOS}}$ being the Jacobians with respect to temperature and LOS for the spectral windows used in the constituent retrieval. $\mathbf{S}'_m$ is then added to $\mathbf{S}_m$ in Eq. 6 in order to quantify the ESD of the total
 210 retrieval noise error. However, the ESD of the TLOS-propagated contribution is also reported separately.

Realistic emulation of L2 data in the FL2S requires not only information about the ESD of L2 measurement noise errors, but also about their spatial and temporal correlations. The noise error response ϵ_{noise} , including correlations, can be emulated via factorisation of the retrieval measurement covariance matrix $\mathbf{S}_m$ (including also the TLOS-propagated contribution $\mathbf{S}'_m$ for trace gas retrievals):

$$215 \quad \epsilon_{\text{noise}} = \mathbf{U}\mathbf{r} \quad \text{with} \quad \mathbf{S}_m = \mathbf{U}^\top\mathbf{U}. \quad (9)$$

$\mathbf{r}$ is a normalized random vector spanning the parameter space, and $\mathbf{U}$ is computed from $\mathbf{S}_m$ by means of Cholesky decomposition of Eq. (9). The LPE provides the sub-matrix $\mathbf{U}'$ formed by the rows corresponding to the parameters j' at the along-track center. The computation of ϵ_{noise} is performed in the FL2S.

Error propagation from other sources than measurement noise (e.g., calibration uncertainties, systematic instrumental un-
 220 certainties, non-LTE uncertainties) is estimated via the Gain matrix $\mathbf{G}$ as

$$\epsilon_{\text{sys}} = -\mathbf{S} \left(\sum_i \mathbf{K}_i^\top \mathbf{S}_y^{-1} \right) \Delta \mathbf{y} = -\mathbf{G} \Delta \mathbf{y} \quad \text{with} \quad \mathbf{G} = \mathbf{S} \left(\sum_i \mathbf{K}_i^\top \mathbf{S}_y^{-1} \right) \quad (10)$$



from pre-calculated spectral perturbations $\Delta \mathbf{y}$. These are either provided by the L1b simulator or computed internally in the case of calibration errors (radiometric scaling and additive uncertainties). Note that Eq. (10) assumes that the spectral perturbations are the same for all along-track measurements i . This assumption is justified for all temporally correlated errors (including calibration errors if the measurements i correspond to the same calibration period). The LPE computes ϵ_{sys} for the parameters j' (a vertical profile of dimension n_z) at the along-track center. Assuming that all error contributions are uncorrelated among themselves, the total systematic uncertainty is estimated as the square root of the quadratic sum of all systematic uncertainty contributions. Note that the total systematic error in the LPE output netcdf file also accounts for calibration uncertainties, while these are considered independently in the performance summary figures discussed in Section 3.

In the case of spatially and/or spectrally uncorrelated (or partly correlated) uncertainties, as well as for calibration errors, Monte Carlo (MC) simulations are performed with random spectral perturbations $\Delta \mathbf{y}$. For radiometric calibration errors, these perturbations are computed by multiplicatively or additively perturbing the reference spectrum with random vectors reflecting the spatial and spectral correlations to be considered. The spectral correlation width (typically 10 cm^{-1}) and the MC ensemble size (typically 100) for these types of errors are configurable. For pointing (both relative and absolute) and spectral position errors, $\Delta \mathbf{y}$ is computed from the LOS and spectral shift Jacobians provided by the L1b simulator. Also in this case, perturbation spectra – building the MC ensemble – are constructed randomly considering the desired degree of spatial or spectral correlation.

Systematic uncertainties that are evaluated based on MC simulations are reported as standard deviations of the error responses ϵ_{sys} corresponding to all MC members for each error source. The LPE also provides the total systematic error response $\epsilon_{\text{sys_tot}}^l$ (i.e., the sum of all error sources), corresponding to each individual MC member l , which is required by the FL2S in order to simulate the impact of different calibration periods. For temporally correlated error contributions (being independent on calibration periods), only one randomly selected MC member is considered in $\epsilon_{\text{sys_tot}}^l$. In order to properly account for correlations between different across-track 2D retrieval slices, $\epsilon_{\text{sys_tot}}^l$ is further separated into a component considering only spatially uncorrelated error sources $\epsilon_{\text{sys_tot_uncorr}}^l$ and another component including spatially fully correlated errors $\epsilon_{\text{sys_tot_corr}}^l$. The latter includes also fully correlated errors in all domains such as ISRF errors.

In trace gas retrievals, the errors from other uncertainty sources than measurement noise always include the impact of temperature and LOS error propagation associated to the same uncertainty source. That is, Eq. (10) is modified to

$$\epsilon_{\text{sys}} = -\mathbf{G}(\Delta \mathbf{y} + \mathbf{V} \epsilon_{\text{sys_TLOS}}), \quad (11)$$

with $\mathbf{V}$ from Eq. (8) and $\epsilon_{\text{sys_TLOS}}$ being the temperature and LOS error introduced by the same uncertainty source. Note that the consideration of temperature and LOS error propagation can lead to error compensation in case of spectrally correlated uncertainties, or amplification in the case of spectrally uncorrelated uncertainties (see von Clarmann et al., 2022, for a detailed discussion).

In an extended analysis framework, the LPE concept has been adapted to assess the performance of co-retrievals of atmospheric parameters using measurements from CAIRT and selected nadir-viewing instruments onboard the METOP-SG satellite (e.g., IASI-NG and Sentinel-5), with which CAIRT was planned to operate in loose formation. This extended framework builds on synergistic retrievals using the Complete Data Fusion (CDF) method (Ceccherini et al., 2015, 2022), extended to 2D re-



trieval by (Tirelli et al., 2025), and uses CAIRT Jacobians for all ERS atmospheric conditions generated by the L1b Simulator. Generated output can be ingested into the FL2S to evaluate scientific performance of combined limb-nadir products.

2.4 The Fast Level-2 Simulator

The Fast Level-2 Simulator (FL2S) applies the 2D averaging kernels and estimated error terms to each of the grid points of the atmospheric scene, i.e., the synthetic “truth” taken from model data output and interpolated to the CAIRT retrieval grid, and executes the respective convolution resulting in the parametrically retrieved simulated view of the scene. The FL2S allows to simulate the CAIRT L2 products for large data volumes in a computationally efficient way. These simulations are used in science impact studies (e.g., Errera et al., 2026; Sellitto et al., 2026). The FL2S selects or interpolates 2D averaging kernel row sub-matrices $\mathbf{A}'$ and error responses corresponding to the retrieval vector at the along-track center position j' , from the 2D linear retrieval simulations performed with the LPE for each of the ERS scenarios (defined by season, latitude, illumination, and volcanic and solar activity conditions) to the actual atmospheric conditions of the model field at the scene grid points. The model field around each grid point in along-track direction is then convolved with averaging kernel sub-matrices in order to emulate the CAIRT retrieval response. A priori knowledge can optionally be considered. Finally, the simulated error responses are added. The emulated CAIRT L2 profile $x_{j,k}$ at along-track position j of the 2D retrieval slice corresponding to across-track position k is computed as

$$x_{j,k} = \mathbf{A}' \left(x_{j,k}^{truth} - x_{j,k}^{apriori} \right) + x_{j,k}^{apriori} + \mathbf{U}' \mathbf{r}_{j,k} + \epsilon_{\text{sys_tot_corr}}^l + \epsilon_{\text{sys_tot_uncorr}}^{l'} \quad (12)$$

In Eq. (12), $x_{j,k}^{apriori}$ is the corresponding a priori profile, whereas $x_{j,k}^{truth}$, $x_{j,k}^{apriori}$, and $\mathbf{r}_{j,k}$ are along-track segments (corresponding to the chunk size of the 2D retrieval window) around j at across-track position k of the model input field, the a priori field, and a random number field, respectively, covering all along- and across-track positions of the emulated L2 dataset. The MC member l of the total spatially correlated systematic error response is chosen randomly for each simulated calibration period. The MC member l' for the total spatially uncorrelated systematic error response varies further with across-track position k . The computations are performed in an inner loop over all along-track positions j of a given track, which is repeated in an outer loop over all across-track positions k associated to the center tangent point position corresponding to the detector row elements which are binned together as measure of noise reduction. By default, FL2S uses 100 km bins (i.e., the minimum bin size applied in the retrievals, see Section 3.2) for all altitudes, resulting in 4 tracks for a 400 km swath width. If across-track binning of 300 km is applied, the first 3 and the last 3 tracks are averaged to two overlapping 300 km tracks, which are replicated as two 200 km tracks to be consistent with the four-track setup. Since the noise and the spatially uncorrelated systematic errors have been estimated for 300 km binning for these cases, and since random errors computed by the FL2S are spatially uncorrelated for each 100 km wide track, the averaging procedure would underestimate those errors by a factor of $\sqrt{3}$. To counteract that unwanted error reduction by averaging, the noise part and the spatially uncorrelated part of the systematic error are multiplied by the same factor in places where the tracks are averaged to emulate a coarser across-track binning.

To implement a realistic cloud masking strategy in the FL2S simulator, an approach was adopted that mirrors the treatment of clouds in real L2 data processing. In operational retrievals, spectra strongly affected by optically thick clouds are routinely

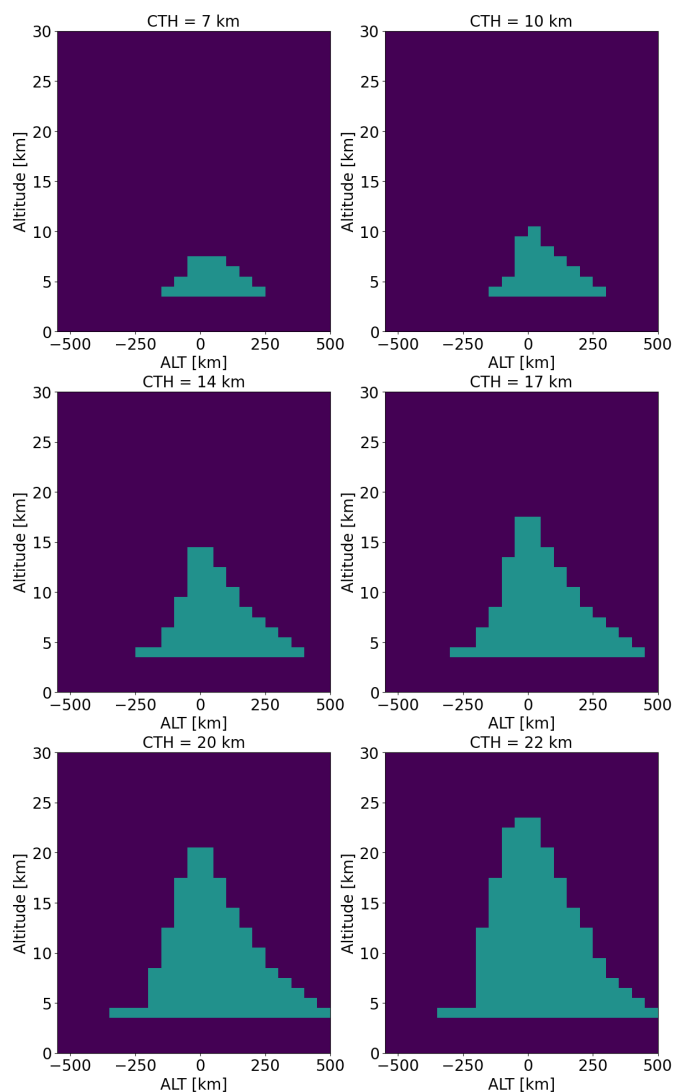


Figure 3. Extension of the 2D (along-track position $\times$ altitude) area of L2 data points classified as unreliable (with averaging kernel diagonal element < 0.003), indicated by the green shading, due to the presence of a column cloud at along-track (ALT) center position extending from 4 km to different cloud top heights (CTH) as indicated in the title of each map.



excluded prior to inversion to prevent instabilities and systematic errors in the retrieved trace-gas concentrations. Accordingly,
 290 the spatial extent of the retrieval response degraded by the presence of thick clouds was assessed by evaluating the diagonal
 elements of the 2D averaging kernel matrix corresponding to the grid point around the cloud position after the removal of
 those spectra whose lines of sight intersect the cloud volume. To represent the effect of spatially extended cloud fields along
 the satellite orbit, the FL2S implementation decomposes the cloud masking problem into a superposition of localized cloud
 contributions at individual grid points. This decomposition is made possible by assuming that clouds extend vertically from the
 295 lowest CAIRT tangent altitude to the cloud top altitude — referred to hereafter as columnar clouds. This assumption provides a
 conservative representation of cloud masking, since stratified clouds with regional horizontal extent would, in principle, permit
 cloud-free observations below their base. Based on this analysis, a set of lookup-table masks was derived for discrete, spatially
 localized clouds — restricted to a single horizontal grid point — at specific cloud top heights (see Fig. 3). These lookup tables
 were constructed from the averaging kernel response to retrievals performed after removing all spectra whose lines of sight
 300 intersect the cloud volume. They define the spatial region in which retrieved profile points exhibit an averaging kernel diagonal
 value below 0.003 and must therefore be flagged as unreliable. A grid point is classified as cloudy when the cloud fraction
 from the model dataset, interpolated to the corresponding L2 product geolocation, exceeds a given threshold (typically 0.2).
 The simulated L2 profiles are filtered a posteriori using the masks derived from the cloud-impact analysis described above.

Input to the FL2S are the scene data, e.g, the model data representing the “synthetic truth” interpolated to the CAIRT 2D
 305 retrieval grid, see Sec. 3.4, and the LPE output files for the L2 product under consideration. The latter contain the entire ERS
 climatology of 2D averaging kernel row sub-matrices, decomposed correlated noise error terms $\mathbf{U}$ (see Eq. (9)), and the total
 systematic error responses corresponding to individual MC members.

2.5 Validation against End-to-End performance simulations

To validate the LPE results, we cross-compare LPE-predicted temperature and ozone uncertainties due to measurement noise
 310 (ESD) with a full end-to-end mission performance simulation performed with CEEPS (Kleinert et al., in preparation). This
 simulation uses as model truth a horizontally homogeneous atmosphere, corresponding to the mid-latitude daytime atmosphere
 of ERS. The simulation extends over hundred CAIRT acquisitions and 2D retrievals are performed over the across-track center
 slice. Here, the CEEPS L1b radiance data were generated to be representative for NESR performance as indicated in Table 4.
 In this simulation, the a priori profile that constrains the retrieval was set to the same mid-latitude daytime profile as used to
 315 construct the model synthetic truth. This configuration avoids smoothing errors in the retrievals, such that differences between
 retrieval and truth can be directly assessed as propagated errors. Standard deviation across these differences, computed along
 the horizontal dimension and as a function of altitude, provides estimates of the noise error of the retrieved L2 product.

Figure 4 compares the CEEPS-derived estimates of the noise error for temperature and O_3 with the LPE predictions. For the
 temperature retrieval, this standard deviation and the LPE-derived noise error (grey and red solid lines in panel a) agree well.
 320 Moreover, the noise error estimated with the LPE agrees very well with the CEEPS-diagnosed noise error (black solid line) for
 the same atmospheric conditions.

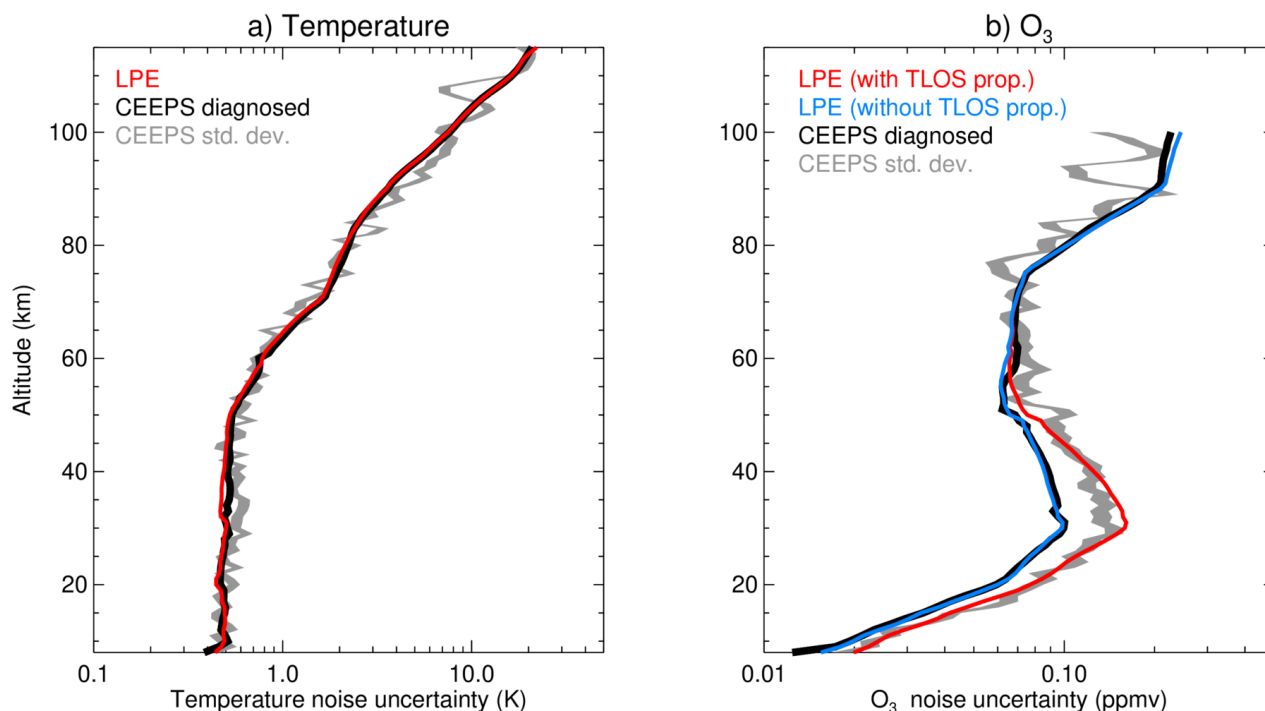


Figure 4. Comparison of LPE-predicted and CEEPS-derived measurement noise uncertainties for (a) temperature and (b) ozone. Red: LPE predictions (ESD), blue: LPE predictions excluding the TLOS-propagated contribution (only for ozone), black: CEEPS (diagnosed), grey: CEEPS (computed from standard deviation of the retrieved profiles, note that the thickness of the grey line corresponds to the associated statistical uncertainty).

For the O_3 retrieval (see panel b), conversely, the standard deviation (grey) exceeds the diagnosed noise error (black) and the LPE direct noise error (blue) estimates in the stratosphere. This is expected, since the diagnosed O_3 noise error does not account for error propagation from the preceding temperature and LOS retrieval. LPE estimates of the total noise error (red line) do consider this additional contribution and align very well with the standard deviation across the differences between retrieved O_3 and model truth from the CEEPS simulations. This excellent match provides evidence and confidence that the covariant error propagation in the LPE has been implemented adequately.

3 Characterization of CAIRT retrievals

3.1 L2 geophysical data requirements and performance metrics

Geophysical data requirements for CAIRT L2 products have been established during the mission definition and feasibility studies (Phase 0/A) and are summarized for primary L2 targets in Table 4.2 of the Report for Mission Selection (European



Table 2. L2 requirements for secondary L2 targets (G: goal; T: threshold requirement).

Target	vertical resolution (G/T)	horizontal resolution (G/T)	total uncertainty (G/T)
HCN	1/2 km	100/200 km	0.05/0.2 ppbv
HCOOH	1/2 km	100/200 km	0.1/1.0 ppbv
C ₂ H ₂	1/2 km	100/200 km	0.02/0.1 ppbv
C ₂ H ₆	1/2 km	100/200 km	0.05/0.5 ppbv
NH ₃	1/2 km	100/200 km	0.1/0.5 ppbv
NH ₄ NO ₃ —solid	1/2 km	100/200 km	0.2/0.6 $\mu\text{m}^3 \text{ cm}^{-3}$

Space Agency, 2025). These requirements are set up as threshold (T) and goal (G) requirements for horizontal and vertical resolution, as well as total uncertainty, where threshold values represent the minimum performance that the mission must achieve, and goal values indicate a desired function or performance. The L2 requirements for the secondary L2 targets (HCN, HCOOH, C₂H₂, C₂H₆, NH₃, and NH₄NO₃—solid) are listed in Table 2. Uncertainty requirements for these targets have been defined based on detectability of enhanced plumes with respect to background concentrations.

In consonance with the L2 requirements, L2 performance metrics are defined in terms of spatial resolution and total uncertainty. To extract scalar retrieval spatial resolution metrics, the concept of combined spatial resolution for a volume element is used here (see, e.g., Fig. 18, bottom right). The latter is determined as the volume around a retrieval grid-point that would contain one piece of independent information (one degree of freedom), which can be determined from grid configuration and the averaging kernel matrix (see also Section 2.3). To assess compliance, this metric can then be compared against equivalent volumes calculated from the required horizontal and vertical resolutions. Similar metrics for vertical and horizontal resolution can be derived. This approach also facilitates the trade between horizontal and vertical resolution by fine-tuning retrieval regularisation. Estimated spatial resolution and total uncertainties have been divided by the goal and threshold requirement values at each altitude level of each ERS atmosphere. Then, the mean of all these ratios is calculated over all ERS atmospheres and altitude levels within the altitude regions as per the L2a requirements. These average ratios are finally subdivided into four categories for L2 requirement verification (see Table 3).

3.2 Simulation procedure baseline

LPE-simulations for all ERS scenarios have been performed sequentially for spectral shift, temperature and line of sight (T+LOS), O₃, NO, NO₂, CO, H₂O, CH₄, N₂O, HNO₃, N₂O₅, ClONO₂, CFC–11, CFC–12, HCFC–22, SF₆, HCN, HCOOH, SO₂, H₂SO₄—liquid, C₂H₂, C₂H₆, NH₃, NH₄NO₃—solid, and PAN. All simulations are performed for a spectral sampling of 0.2 cm^{−1} and an along-track sampling distance of 50 km corresponding to about 7 s acquisition time. Horizontal across-track binning of spectra is applied as a measure of noise reduction, depending on retrieval target and altitude range. 100 km binning (in terms of tangent point coverage in across-track direction) is applied to all targets with negligible concen-



Table 3. L2 requirement compliance categories applied to L2 performance metrics (G: goal; T: threshold requirement). The color coding corresponds to the shading used in performance assessment figures in Section 3.5.

Category	Color code	Performance evaluation	Assessment (example: predicted [p] vs required [G/T] uncertainty u)
1	dark green	Exceeds (better than) goal	$u_p/u_G \leq 1$
2	light green	Between goal and threshold	$u_p/u_G > 1$ and $u_p/u_T \leq 0.8$
3	yellow	Threshold compliant	$0.8 < u_p/u_T \leq 1.2$
4	red	Non-compliant	$u_p/u_T > 1.2$

trations above the mid-stratosphere, while 300 km binning is applied to all reactive nitrogen species and CO. Temperature is retrieved with 100 km binning below 60 km and 300 km binning above. For H₂O, CH₄, N₂O, and O₃, the transition to 300 km binning is at 50 km.

All retrievals except spectral shift correction are performed as 2D retrievals of the geophysical target parameters in the vertical and along-track domain. Besides the target quantity, the retrievals include an altitude- and spectral window-dependent radiometric offset correction, which is assumed to be temporally constant (within a calibration period), and a spectral window-dependent extinction profile up to 60 km. The spectral shift is retrieved in a first step, jointly with a microwindow- and tangent height-dependent radiometric scale (gain) and offset, both not being used further in the retrieval sequence. A spatially constant but temporally variable LOS correction is then retrieved jointly with temperature. HNO₃, PAN, and CFC-11 retrievals also include a H₂O ln(vmr) profile joint fit to reduce systematic error propagation in the free troposphere and UTLS region.

The horizontal (along-track) retrieval-grid spacing was set to $\Delta x = 50$ km. The vertical retrieval grid is variable and defined according to the target parameter. For NO and CO, it spans 0–125 km at 1 km spacing, 130–150 km at 10 km spacing, with additional grid points at 160, 170, 180, and 200 km. For T+LOS, O₃, H₂O, CH₄, and N₂O, the grid spans 0–120 km at 1 km spacing. For HNO₃, it spans 0–80 km at 1 km spacing, with additional grid points at 85, 90, 100, and 120 km. For all remaining targets, it spans 0–70 km at 1 km spacing, with additional grid points at 75, 80, 90, 100, and 120 km.

Tikhonov first order regularization (Tikhonov, 1963; Twomey, 1963; Phillips, 1962) is applied in both horizontal and vertical dimensions for the target quantities and for H₂O in the vertical direction when jointly fitted. The regularization strength profiles were adjusted in accordance with the L2 geophysical requirements for the CAIRT project, based on empirical testing. For the radiometric offset, a diagonal (optimal estimation type) constraint was employed. Its strength was adjusted according to the radiometric offset knowledge requirement with the exception of CO, HNO₃, and PAN where the strength has been modified to optimize the trade-off between residual radiometric offset errors and the increase of measurement noise errors. We use the same spectral windows as used in MIPAS IMK-IAA V8 processing (Kiefer et al., 2021) in all retrievals, with slight modifications for temperature and LOS, O₃, NO₂, CFC–11, SF₆, HCN, and PAN.



Table 4. Noise-Equivalent Spectral Radiance (NESR) specification values considered in the error budgets. Linear interpolation has been applied for wavenumbers in-between.

Wavenumber (cm^{-1})	NESR ($\text{nWcm}^{-2}\text{sr}^{-1}\text{cm}$)
718	42.0
750	25.5
800	13.5
1500	25.5
1550	51.0
1590	6.0
1650	3.2
2150	4.2
2200	8.3

3.3 Sources of uncertainty

Observing system effects are represented through the instrument-specific measurement performance, which includes knowl-
 380 edge uncertainties of the spectral, radiometric, geometric, and calibration characteristics for which measurement requirements
 have been defined. For all these error sources, propagation through the temperature and LOS retrieval into the subsequent
 trace-gas retrievals is fully taken into account.

One of the most important error sources is the NESR which was treated according to CAIRT's threshold (breakthrough in
 the MWIR channel, see below) requirement for a spectral sampling of 0.2 cm^{-1} and a spatial sample of 1 km in the vertical
 385 and 50 km in across-track direction (Table 4). Across-track binning in the retrievals is accounted for by NESR reduction
 proportional to the square root of the number of 50 km samples binned together. The remaining observing system effects were
 implemented as follows.

The spectrally and spatially uncorrelated components of the radiometric scaling and additive errors were assumed to have
 a spectral correlation length of 10 cm^{-1} . In the MC simulations, random perturbations were therefore applied at each tangent
 390 height every 10 wavenumber samples, with linear interpolation between adjacent samples. The MC ensemble consisted of 100
 realizations (see also Sect. 2.3).

CAIRT's focal plane array layout utilizes a distinct two-channel optical design, splitting the incoming infrared radiation
 into separate spectral bands, namely the LWIR band ($718\text{--}1550 \text{ cm}^{-1}$) and the MWIR band ($1550\text{--}2200 \text{ cm}^{-1}$), directed
 onto separate detectors. While this dual-band approach maximizes the target species coverage, it introduces the potential for
 395 inter-band pointing misalignments due to mechanical or thermal deformations between the channels. We consider inter-band
 relative pointing errors only for spectral windows above 1550 cm^{-1} (corresponding the MWIR band) by applying independent
 random elevation-angle displacements at the upper and lower boundaries of the vertical field of view, followed by linear
 interpolation between them.



Table 5. Error sources considered in the error budgets. SEDF ranges are provided as vertical distances in the tangent plane.

Error source	Uncertainty (1- σ)	Spatial corr.	Spectral corr.	Time corr.	Combined in legends Figs. 5–17
NESR	see Tab. 4	uncorr	uncorr	uncorr	random/ NESR
Spectral accuracy	uncertainty of spectral shift retrieval (0.1–0.6 ppm)	full	linear	uncorr.	random/ spectral shift
Pointing stability (vertical)	25 m	full	full	uncorr	random/ pointing stability
Radiometric scaling	1%	full	full	calibration period	calibration/ radiometric scaling
Radiometric scaling	0.25%	uncorr	10 cm ⁻¹	calibration period	calibration/ radiometric scaling
Radiometric additive	3 nWcm ⁻² sr ⁻¹ cm (<1550 cm ⁻¹) 0.3 nWcm ⁻² sr ⁻¹ cm (>1550 cm ⁻¹)	uncorr	10 cm ⁻¹	calibration period	calibration/ radiometric additive
Rel. pointing (intraband)	100 m	linear displacement	full	full	systematic/ rel. point. knowledge
Rel. pointing (interband)	100 m	linear displacement	full	full	systematic/ rel. point. knowledge
ISRF	1% of ISRF width	full	full	full	systematic/ ISRF knowledge
SEDF center (0–1.5 km)	2% FWHM	full	full	full	systematic/ SEDF knowledge
SEDF near field (1.5–2.5 km)	1%	full	full	full	systematic/ SEDF knowledge
SEDF far field (2.5–20.5 km)	5.3 × 10 ⁻⁴ (2.5–5.5 km) 6.1 × 10 ⁻⁵ (5.5–20.5 km)	full	full	full	systematic/ SEDF knowledge
SEDF far field (20.5–115 km)	1.2 × 10 ⁻⁵ (20.5–60.5 km) 5 × 10 ⁻⁶ (60.5–115 km)	full	full	full	systematic/ SEDF knowledge
CO ₂ uncertainty	height-dependent	full	full	full	systematic/ CO ₂ uncertainty
Various non-LTE error sources	García-Comas et al. (2023)	full	full	full	systematic/ NLTE uncertainties

Intra-band relative pointing errors were applied over the full spectral range covered by both channels, assuming a linear correlation in the vertical dimension. In this case, the perturbation was applied only at the upper boundary, because displacements



at the lower boundary are fully represented by the LOS retrieval. The effect of both intra- and inter-band pointing errors on the L2 products was quantified by MC simulations with 100 ensemble members.

The effect of vertical pointing instability was evaluated by perturbing simulated limb interferogram series with elevation-jitter frequencies ranging from 0.1 to 500 Hz. The corresponding elevation time series was generated assuming Gaussian-distributed jitter amplitudes with a standard deviation of 25 m displacement of the tangent point altitude. Perturbation spectra derived for all ERS atmospheres were then used as input for a one-dimensional linear error propagation analysis. Because perturbations induced by pointing jitter are uncorrelated in the along-track dimension, the two-dimensional LPE uncertainty analysis cannot be applied in this case. Therefore, for the performance assessment presented here, the jitter-induced uncertainties obtained from the 1D analysis were scaled by a factor of $\sqrt{50 \text{ km}/h}$ where h denotes the horizontal resolution derived from the LPE retrieval diagnostics. This scaling accounts for the fact that the 1D analysis refers to a single along-track acquisition, whereas the 2D retrieval combines several uncorrelated spectra in the along-track dimension, corresponding to approximately $h/50 \text{ km}$.

Spectral accuracy was estimated from the linear spectral shift retrievals using 11 spectral windows within $750\text{--}1950 \text{ cm}^{-1}$, a slightly extended version of the spectral window selection used in MIPAS IMK-IAA retrievals. Spectra from all spatial samples with tangent heights above 25 km recorded within a single acquisition are used. Both systematic and random error components were estimated. Here, the systematic component also includes the effect of absolute pointing accuracy, because spectral shift is retrieved prior to temperature and line-of-sight correction. For the tangent heights included in the spectral shift retrieval ($>25 \text{ km}$), however, the error is dominated by measurement noise.

Uncertainties in the external input data were considered for the atmospheric parameters that are not retrieved explicitly but still affect the measured radiances. In particular, uncertainties in CO_2 influence the temperature retrieval, especially in the mesosphere and lower thermosphere, and propagate into the trace-gas retrievals. Following Kiefer et al. (2021), the assumed uncertainties in CO_2 volume mixing ratio were 0.2% below 30 km, 0.5% at 40 km, 1% from 60 to 80 km, 10% from 90 to 100 km, and 20% at 110 km. Below 60 km, these values are based on uncertainties reported in the IPCC Fifth Assessment Report and include the contribution from seasonal variability. Above 60 km, they are derived from comparisons of WACCM CO_2 fields with MIPAS/Envisat, Atmospheric Chemistry Experiment - Fourier Transform Spectrometer (ACE-FTS) on SciSat, and Sounding of the Atmosphere using Broadband Emission Radiometry (SABER) observations (López-Puertas et al., 2017).

Further atmospheric input parameters, including atomic oxygen and nitrogen in different electronic states (O , N^4S , and N^2D), are required for modelling non-LTE populations of CO_2 , O_3 , H_2O , NO_2 , NO , and CO . The associated uncertainties, together with those of non-LTE-specific auxiliary parameters such as collisional rate constants and photolysis rates, were adopted from a comprehensive assessment based on MIPAS data analysis (García-Comas et al., 2023). It should be noted that uncertainties in non-LTE processes involving CO_2 vibrational states affect not only the temperature retrieval but also the retrieved volume mixing ratios of other species through error propagation.

Uncertainties caused by emissions of other trace gases can be efficiently mitigated by appropriate selection of spectral windows in the retrieval or by fitting them jointly with the target L2 parameter (Dinelli et al., 2004; Schwarz et al., 2002). They were therefore not considered in the LPE-based performance assessment.



Spectroscopic contributions, including uncertainties in line intensities, broadening coefficients, tabulated cross sections of heavy molecules, and aerosol optical constants, are not included in the current CAIRT uncertainty budget. A major limitation is that correlations among spectroscopic parameter errors are generally unknown and are not provided in standard spectroscopic databases, although such correlations are likely, in particular when laboratory measurements are affected by systematic errors.

440 Moreover, for some species, the reported spectroscopic uncertainties appear to be inconsistent with validation results. For example, a CH₄ volume mixing ratio profile retrieved from a Gimbalbed Limb Observer for Radiance Imaging of the Atmosphere (GLORIA-B) flight over Kiruna in August 2021 agrees with in situ measurements to within 10%, whereas uncertainties inferred from line intensities and broadening coefficients would suggest a substantially larger error (e.g. Wetzel et al., 2025). In addition, future improvements in spectroscopic datasets may significantly reduce related uncertainties.

445 Table 5 summarizes all the considered error sources together with their assumed magnitudes and their vertical, spectral, and temporal correlation characteristics, as well as their combinations as shown in the selected example Figures 5–17.

3.4 Test scenario for FL2S simulations

FL2S simulations offer the opportunity to showcase scientific performance on realistic scenes by emulating L2 products. Here we choose a test scenario corresponding to the dynamically perturbed 2019 Northern hemisphere (NH) winter. That winter was characterized by a major stratospheric warming event in January, followed by strongly enhanced polar mesospheric descent, and the formation of an elevated stratopause in the recovery phase (Okui et al., 2021). This scenario is particularly suited to demonstrate the ability of CAIRT to quantify dynamical variability and vertical coupling throughout the atmosphere. Our test scenario is based on a extended Whole Atmosphere Community Climate Model (WACCM-X) simulation with a vertical extension from the ground to the exobase, configured similarly to that described by Liu et al. (2024) and generated by nudging

455 WACCM-X to Modern-Era Retrospective analysis for Research and Applications Version 2 (MERRA-2) reanalysis data up to about 45 km altitude. The specific scenario dataset is based on interpolating from the model’s native cubed-sphere grid (25 km resolution) to a $0.25^\circ \times 0.25^\circ$ horizontal grid and interpolated from model levels to equidistant 1 km vertical grid in the altitude range of 0 to 200 km. We used a reduced dataset of 1.5 hours duration on 18 February 2019, 0:00–1:30 UT (sufficient to cover an orbit); it contains all CAIRT L2 targets and has the necessary high time resolution (5 min) to resolve atmospheric variability

460 down to minute scale. The model data has been interpolated to the CAIRT retrieval grid using the scene generator module of CEEPS (Kleinert et al., in preparation).

3.5 Performance results for selected atmospheric parameters

Figures 5–13 (left panels) illustrate the LPE-estimated global average performance for selected CAIRT L2 products with significant contributions at least up to the middle stratosphere, evaluated against the resolution and uncertainty metrics defined

465 in Sec. 3.1. Resolution and uncertainty budgets for CFC–11, ClONO₂, PAN, OCS, SO₂, H₂SO₂–liq, HCN, C₂H₂, C₂H₆, HCOOH, NH₃, and NH₄NO₃–sld are provided in the Supplement Figures S1–S12.

Beyond the total uncertainty, the error budgets distinguish between three distinct contributors: (i) random uncertainties driven by spatially and temporally uncorrelated errors, which can be mitigated via averaging; (ii) calibration uncertainties that

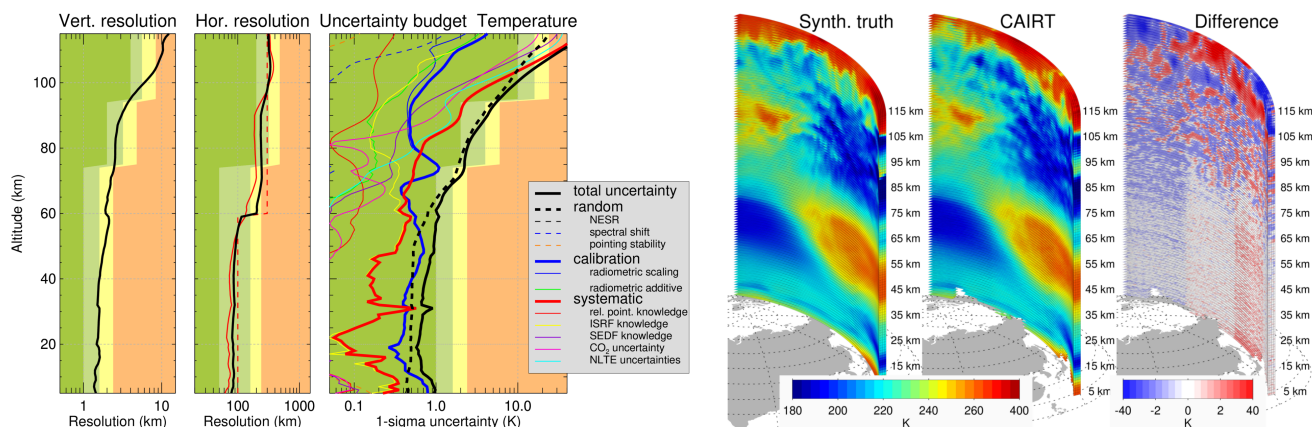


Figure 5. Left: L2 performance estimation for temperature generated with the LPE. Background shading corresponds to the L2 requirement compliance categories of Table 3, following the sequence of category 1 to 4 from left to right. The total error is the 1- σ equivalent width of an error distribution, calculated as the square root of the quadratic sum of all random and systematic components assessed. In the horizontal resolution sub-plot, red solid lines indicate along-track resolution, and dashed red lines denote across-track binning. Right: “True” temperature and FL2S-emulated CAIRT data, alongside their differences, for an orbit segment across the test scenario.

cause short-term systematic effects over calibration periods but behave randomly on larger timescales; and (iii) systematic
 470 uncertainties from correlated errors, resulting in persistent biases in the L2 products.

On average, the threshold L2 requirements are predicted to be successfully met across all altitude regions for the majority of the targets; specific regional deviations, exceptions, and error budget characteristics are detailed below.

Temperature: While compliance is maintained on average (Fig. 5), the vertical resolution threshold is slightly exceeded in the lower thermosphere under polar summer conditions (not shown). Random errors (NESR) dictate the overall uncertainty.
 475 However, calibration errors, chiefly radiometric scaling errors, contribute up to 60% in the upper troposphere and stratosphere. In the mesosphere, systematic errors are driven primarily by ISRF knowledge (up to 50%), whereas non-LTE errors dominate at higher altitudes, led by uncertainties of the CO₂-O quenching rate.

Ozone (O₃): Performance generally complies with requirements (Fig. 6), except for a minor vertical resolution exceedance in the lower stratosphere during Southern Hemisphere (SH) polar winter nights (not shown). Random noise remains the primary
 480 uncertainty contributor. Radiometric scaling errors dictate the calibration uncertainty, acting as the most relevant source below the stratopause and transitioning to additive errors above it. Systematic errors are governed by ISRF knowledge up to the mesosphere (up to 30% of the total budget), giving way to non-LTE processes (O₃-O-vt) above (up to 50% around 90 km.).

Water Vapor (H₂O): Requirements are met across most domains, with a notable exception for vertical resolution in the stratosphere and UTLS region (Fig. 7). Because the spatial resolution threshold is satisfied there (see Fig. 18), this localized
 485 non-compliance can be mitigated by optimizing the applied regularization. The horizontal resolution thresholds are occasionally exceeded in the upper mesosphere. NESR strongly dominates the budget in the stratosphere and mesosphere, though

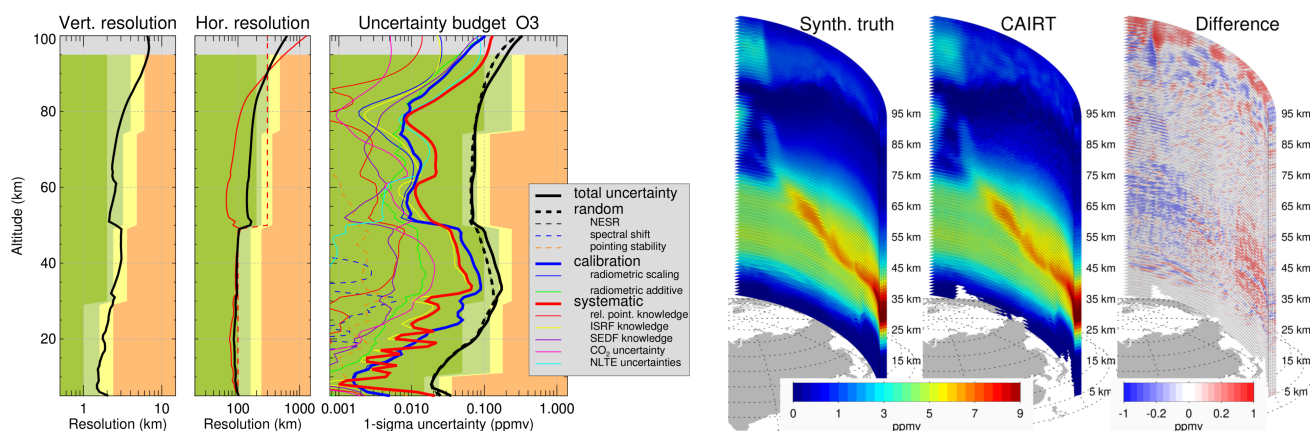


Figure 6. Same as Fig. 5, but showing L2 performance estimation for O_3 .

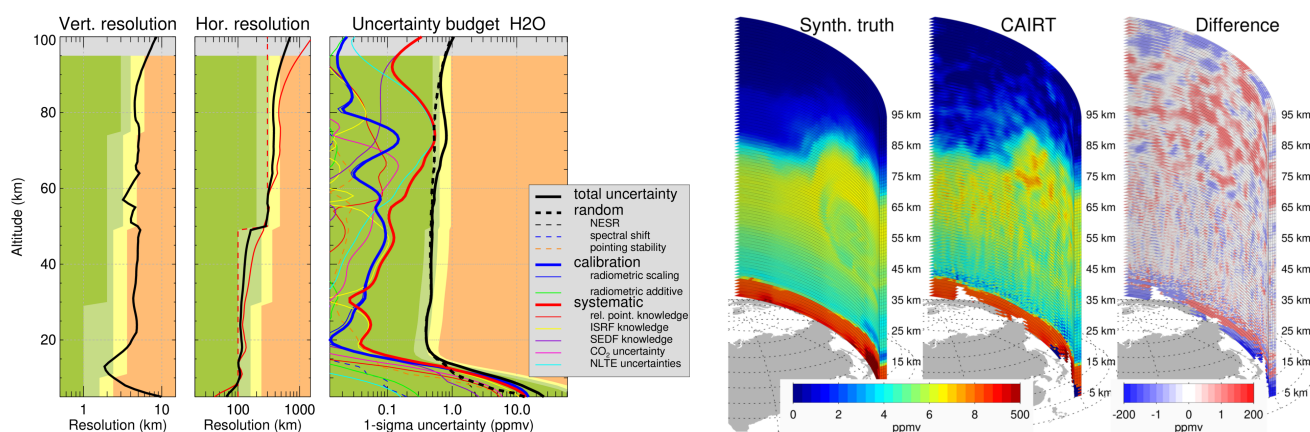


Figure 7. Same as Fig. 5, but showing L2 performance estimation for H_2O .

calibration errors (particularly radiometric scaling errors) are the major contributors in the troposphere, and non-LTE systematic errors become relevant in the upper mesosphere.

Methane (CH_4): Full threshold compliance is achieved globally (Fig. 8), except for isolated vertical resolution non-compliance in the lower stratosphere during SH polar winter (not shown). NESR governs the total uncertainty; calibration errors remain below 10%, while systematic errors, dominated by ISRF knowledge, contribute roughly 10–30%.

Nitrous Oxide (N_2O): Spatial resolution and total uncertainty remain well below threshold in the targeted vertical range (Fig. 9). NESR is the main contributor to total uncertainty. Systematic errors are dominated by ISRF knowledge, while calibration errors do not exceed 10%.

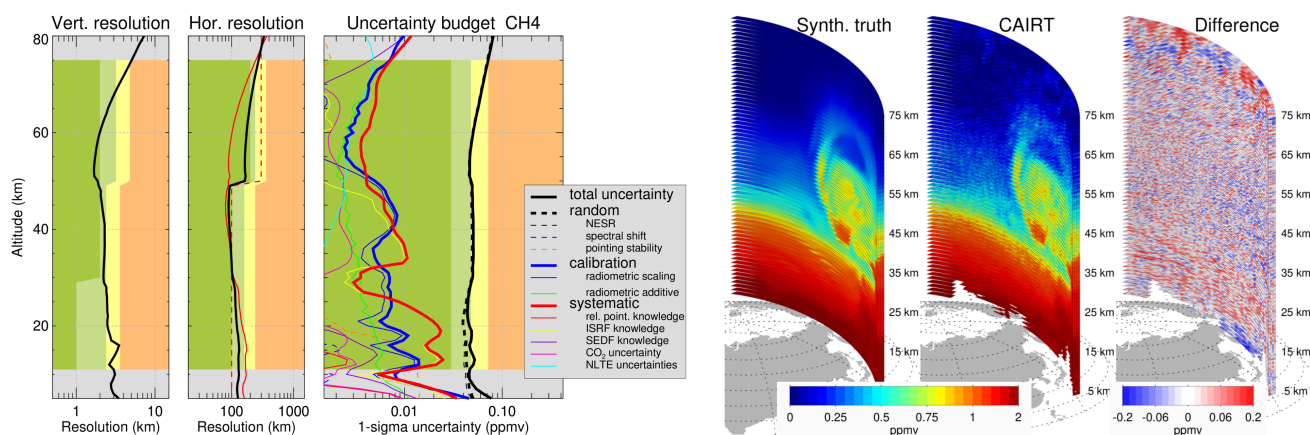


Figure 8. Same as Fig. 5, but showing L2 performance estimation for CH₄.

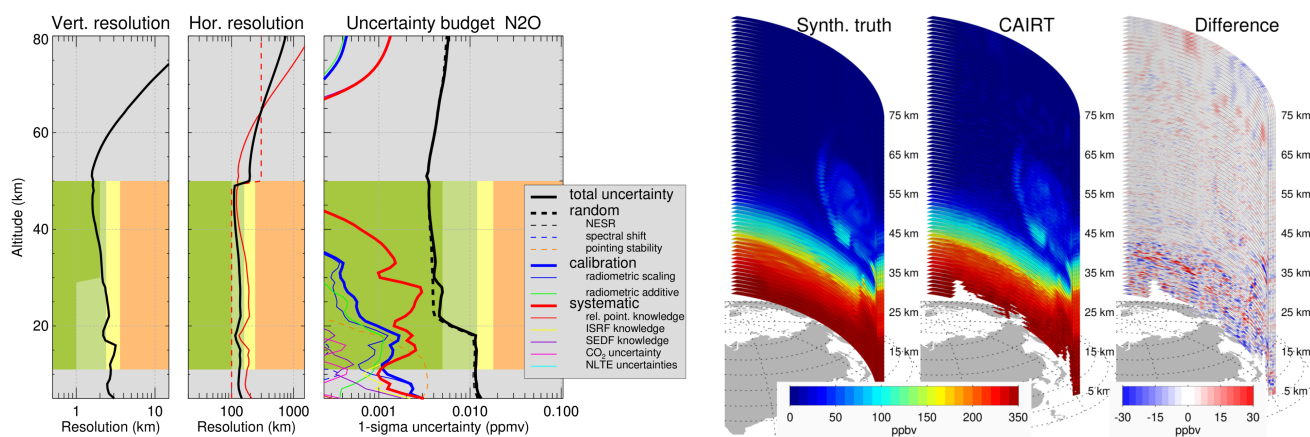


Figure 9. Same as Fig. 5, but showing L2 performance estimation for N₂O.

495 *Chlorodifluoromethane* (HCFC–22): L2 threshold compliance is well achieved in all domains (Fig. 10). Again, NESR is dominating the total uncertainty, followed by calibration errors contributing with 20–30% to the uncertainty budget. ISRF knowledge is the dominant systematic error source with contributions of about 10%.

Dichlorodifluoromethane (CFC–12) : Threshold L2 requirements are met on average in all altitude regions (Fig. 11). The uncertainty is dominated by random errors NESR). Calibration and systematic errors contribute about 20% in the troposphere
 500 and lower stratosphere. Systematic errors are dominated by ISRF knowledge.

Sulfur Hexafluoride (SF₆) : While spatial resolution is on average L2 compliant throughout the vertical range, individual measurements exceed the uncertainty threshold at all levels and conditions (Fig. 12), primarily due to NESR. Daily zonal

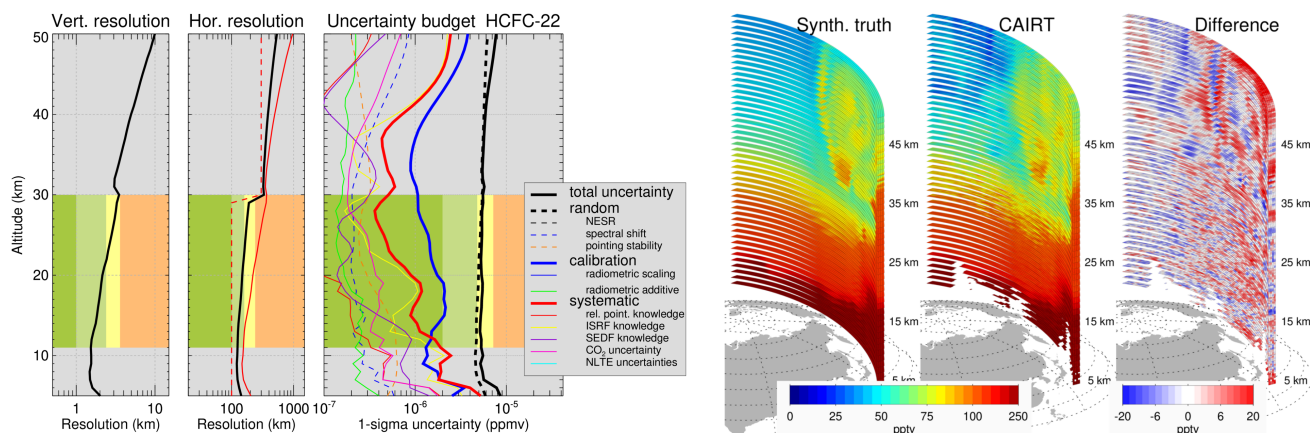


Figure 10. Same as Fig. 5, but showing L2 performance estimation for HCFC–22.

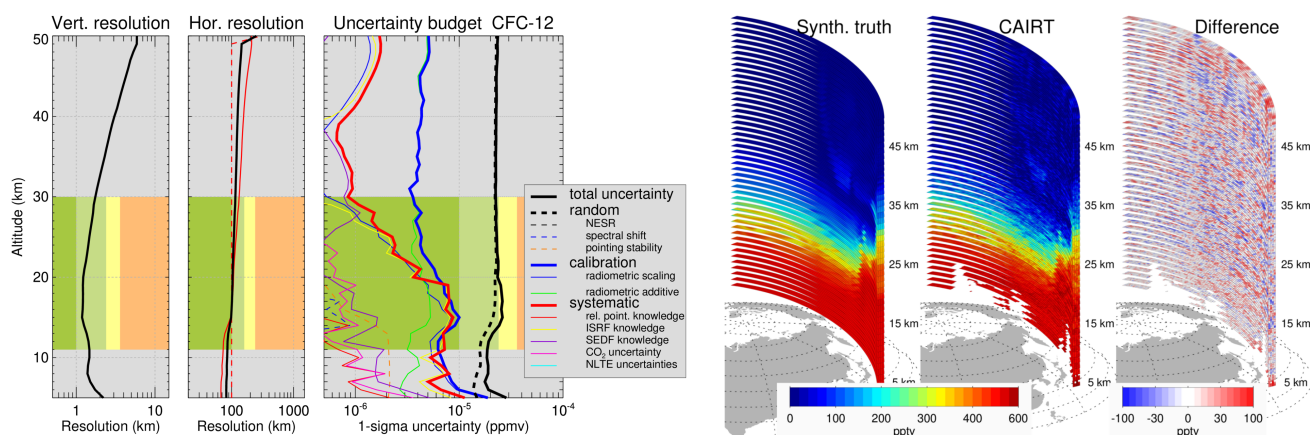


Figure 11. Same as Fig. 5, but showing L2 performance estimation for CFC–12.

averaging is therefore required, reducing the random uncertainty by a factor of 5. Other relevant error sources are calibration errors (mainly radiometric additive errors) and systematic errors dominated by ISRF knowledge.

505 *Carbon Monoxide (CO):* Daytime retrievals satisfy all L2 thresholds except for vertical resolution above 105 km (Fig. 13), night-time conditions introduce spatial resolution non-compliance in the upper mesosphere and lower thermosphere. This behavior stems from the missing solar excitation of the CO vibrationally excited state at night, which significantly weakens non-LTE limb radiances compared to daytime. This limitation in sounding upper-atmospheric CO via infrared limb emissions is a well-documented phenomenon known from MIPAS-ENVISAT observations (Funke et al., 2009). At lower altitudes,
 510 requirements are only occasionally exceeded, primarily regarding vertical resolution in the tropical lower stratosphere (not

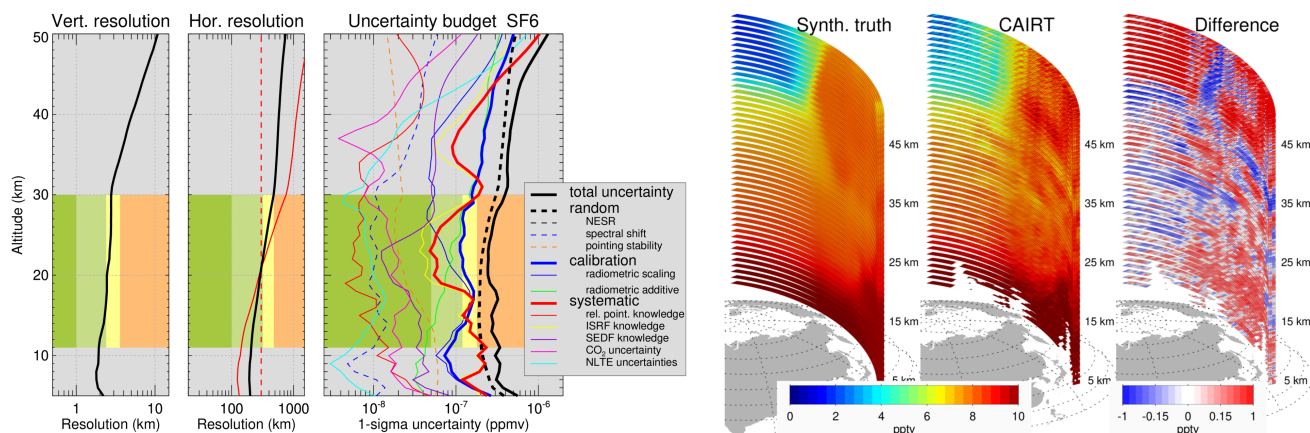


Figure 12. Same as Fig. 5, but showing L2 performance estimation for SF₆.

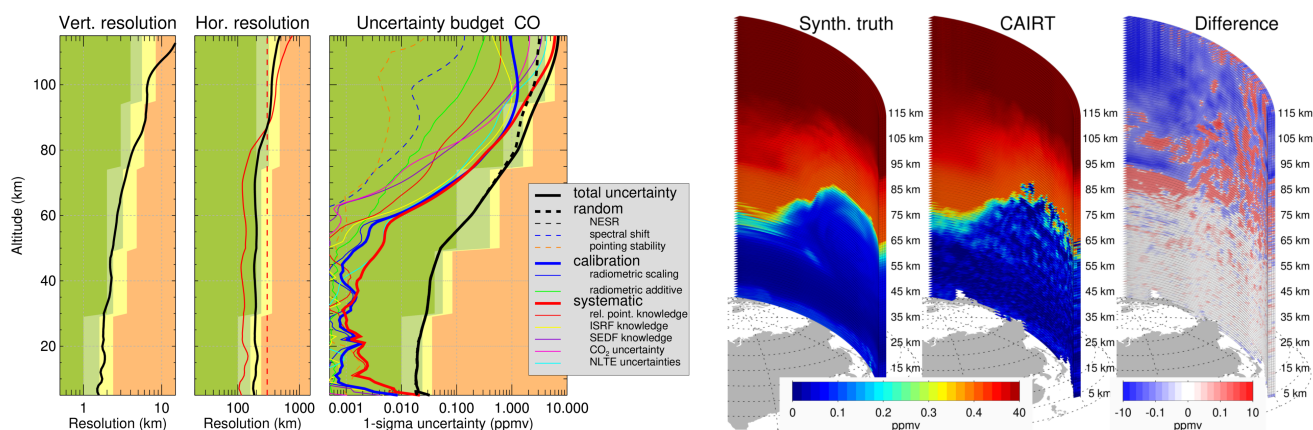


Figure 13. Same as Fig. 5, but showing L2 performance estimation for CO during daytime conditions (left).

shown). The uncertainty is heavily dominated by NESR, with a non-negligible lower-thermospheric contribution from non-LTE systematic errors.

Nitric acid (HNO₃): All threshold L2 requirements are met in the entire vertical range except vertical resolution above approximately 70 km (Fig. 14). The uncertainty is dominated by random errors (NESR) while calibration errors (mainly radiometric additive error) contribute about 20%. Systematic errors are dominated by ISRF mainly in the lower stratosphere with a contribution up to 40%.

Dinitrogen pentoxide (N₂O₅): For this trace gas, even goal L2 requirements are met in all altitude regions (Fig. 15). The uncertainty is dominated by NESR up to the lower stratosphere while above calibration errors (mainly radiometric additive

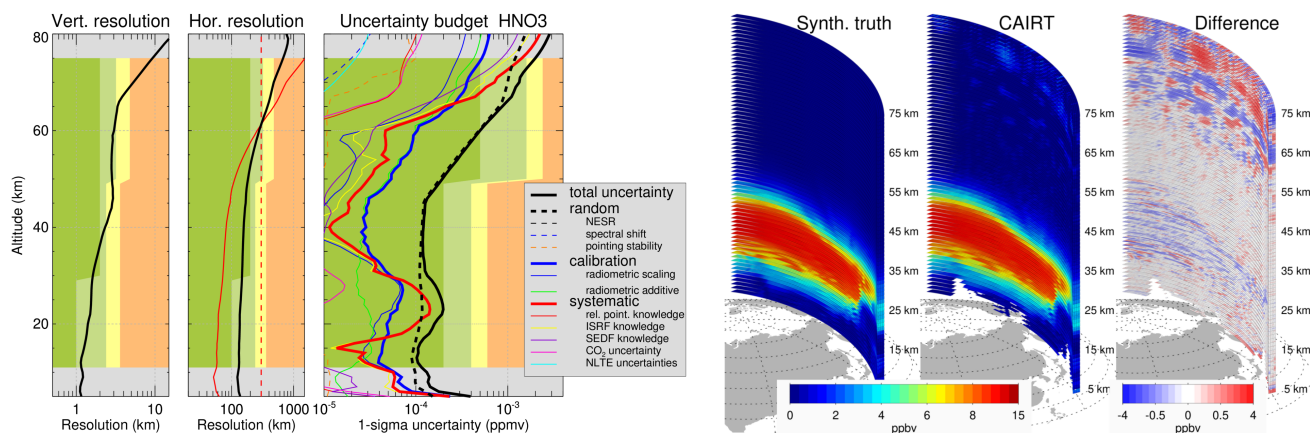


Figure 14. Same as Fig. 5, but showing L2 performance estimation for HNO_3 .

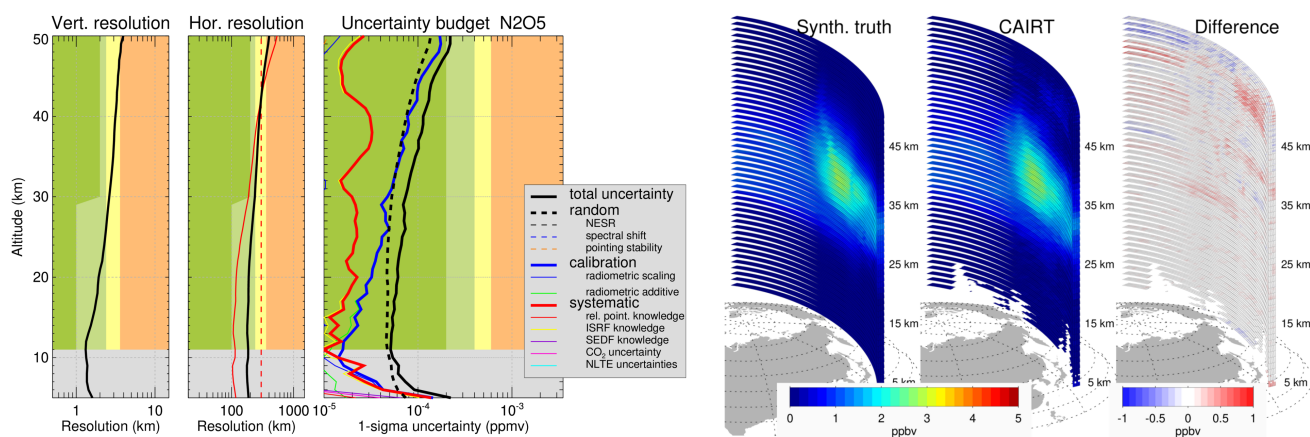


Figure 15. Same as Fig. 5, but showing L2 performance estimation for N_2O_5 .

errors) contribute equally. This is explained by the broad spectral signature of N_2O_5 which cannot be easily distinguished from a radiometric offset within the retrieval.

Nitrogen Dioxide (NO_2): Requirements are generally met (Fig. 16), though the total uncertainty threshold is slightly exceeded in the tropical night-time upper stratosphere (not shown). Random instrument noise dominates, but calibration and systematic errors (primarily ISRF and inter-band pointing knowledge) become locally competitive.

Nitric Oxide (NO): Performance falls within the required limits on average (Fig. 17). However, vertical resolution thresholds are systematically exceeded in the tropical daytime lower stratosphere and sporadically above the stratopause, while horizontal resolution limits are surpassed exclusively in the SH polar winter lower stratosphere (not shown). Random errors (NESR) dictate the uncertainty profile, though non-LTE systematic errors contribute up to 50% in the lower thermosphere.

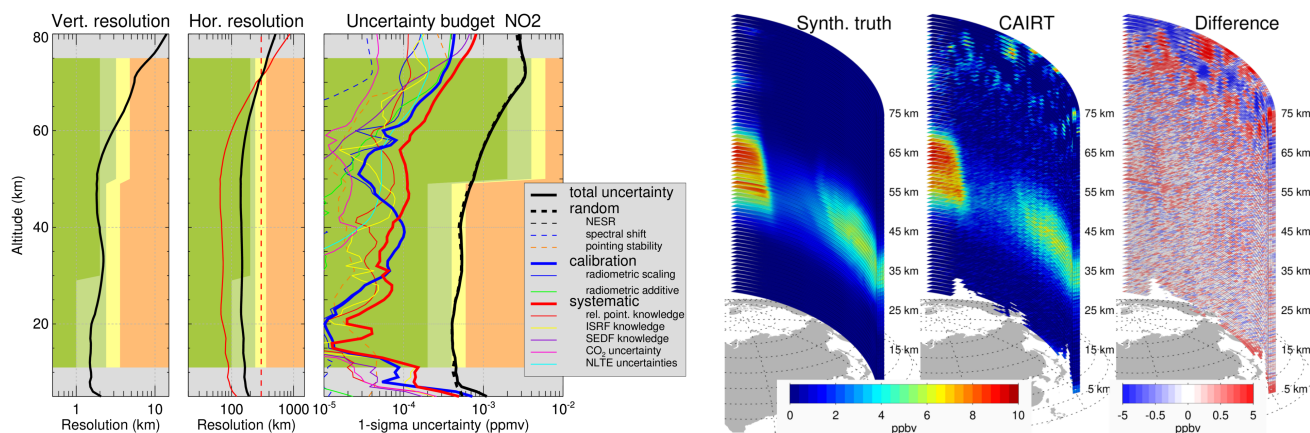


Figure 16. Same as Fig. 5, but showing L2 performance estimation for NO₂.

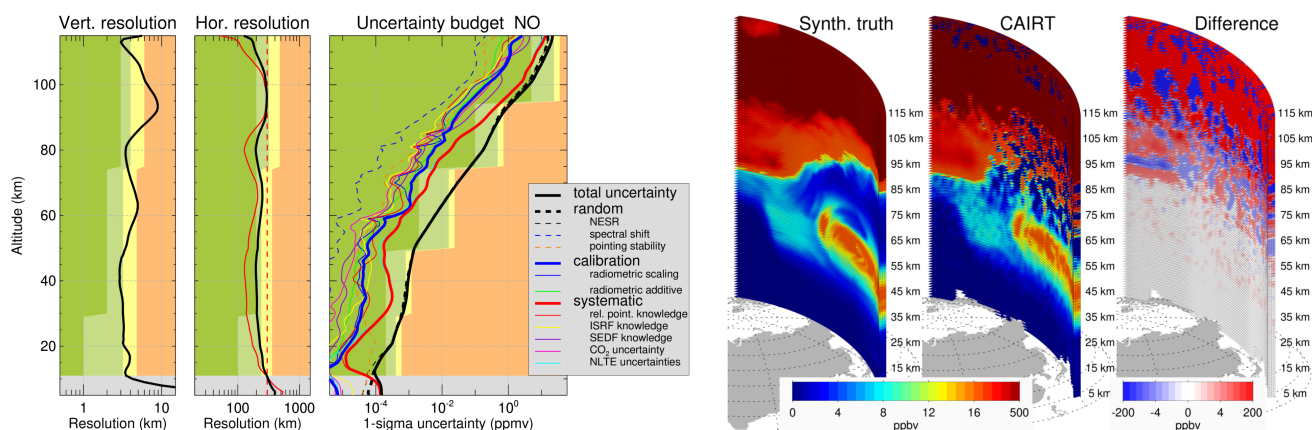


Figure 17. Same as Fig. 5, but showing L2 performance estimation for NO.

In addition, FL2S simulations were conducted for the atmospheric scene described in Sec. 3.4, utilizing the LPE performance outputs. Synthetic CAIRT L2 data are compared directly to the model “truth” in the right-hand panels of Figs. 5–13. This comparison highlights CAIRT’s ability to reproduce wave-induced small-scale temperature features (mostly visible at mesospheric altitudes in Figure 5) and intricate trace-gas filaments (see, e.g., Figure 8 over the Japanese sector) with high fidelity. Crucially, this capability persists even under the highly perturbed NH polar winter conditions of the test scenario, which features an elevated stratopause over the North Pole alongside record-low underlying temperatures. Discrepancies between the simulated data and the synthetic truth occasionally exceed the estimated absolute uncertainty; this is due to “smoothing” errors, as averaging kernels were intentionally not applied to the model truth to preserve a direct comparison.



3.6 Overview of CAIRT retrieval performance

A summary of the estimated performance for both primary and secondary target species is provided in Fig. 18. The color scale is defined as in Table 3 and Sect. 3.1 and is identical to that used in the previous section. The left column of Fig. 18 separates the total retrieval uncertainty into the contributions arising from instrumental spectral noise (NESR) and from all remaining error sources. The right column shows the individual horizontal and vertical contributions to the combined spatial resolution.

Overall, the total uncertainty and combined spatial resolution indicate compliance with the Level-2 requirements for nearly all geophysical target parameters. One exception is the spatial resolution for CO in the mesosphere/lower thermosphere during night-time conditions. This limitation is not considered critical at mission level, as the corresponding requirements are met during daytime. For the secondary target species HCN (not discussed in Section 3.5, but shown in Supplement Figure S7), the results suggest a systematic contribution from ISRF uncertainty which leads to reduced retrieval performance in the upper troposphere.

More generally, the uncertainty contribution due to NESR is typically larger than that from all other evaluated error sources combined. However, this random component could be reduced, if needed for specific scientific objectives, by adjusting the retrieval regularization in cases where maximum spatial resolution is not required, or by applying temporal averaging (zonal averaging across multiple orbit passes). Likewise, the balance between horizontal and vertical spatial resolution can be tuned through the choice of regularization, allowing the retrieval setup to be adapted to the requirements of individual target parameters.

4 Conclusions

We have presented a computationally efficient mission performance modeling framework developed for the Changing-Atmosphere Infra-Red Tomography Explorer (CAIRT), an innovative imaging limb sounder proposed as a candidate for ESA's eleventh Earth Explorer mission. To overcome the significant computational burden associated with full-physics 3D tomographic inversions and non-LTE radiative transfer modeling during early design and trade-off phases, two highly optimized diagnostic tools were established: the Linear Performance Estimator (LPE) and the Fast L2 Simulator (FL2S). Together, these tools bridge the gap between instrumental specifications and scientific utility.

The application of this linear framework to the CAIRT mission concept yielded several key conclusions regarding its technical feasibility and scientific potential:

1. The performance assessments demonstrate that the CAIRT instrument configuration securely satisfies the previously established threshold geophysical L2 requirements on average across all designated altitude regions for the majority of the targeted atmospheric parameters. Where minor, localized exceptions occur, such as night-time spatial resolution limitations for CO, the underlying physical and algorithmic mechanisms are well understood and can be traced back to established non-LTE phenomena already observed by preceding instruments like MIPAS-ENVISAT.

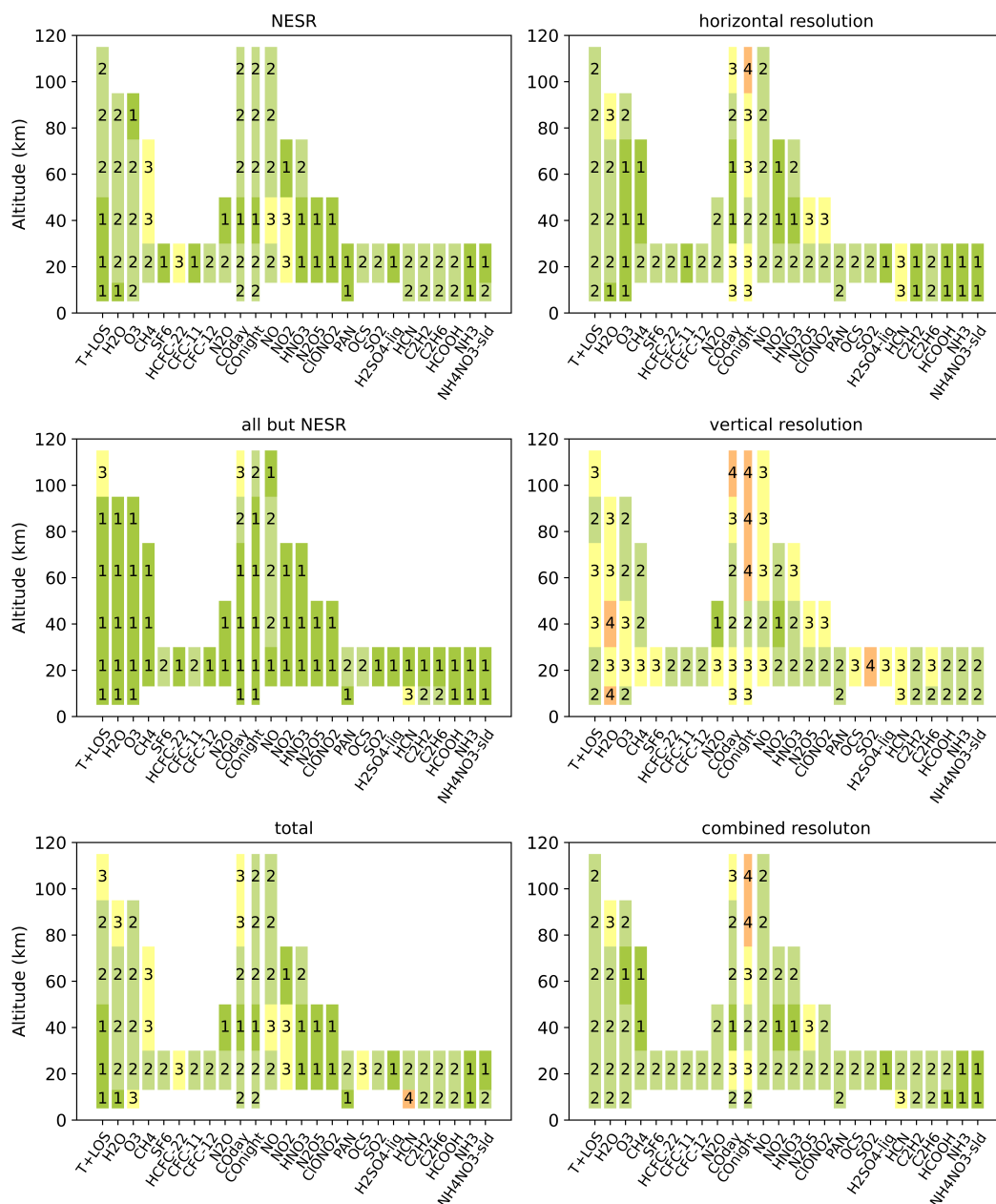


Figure 18. Overview of the linear retrieval performance for the primary and secondary CAIRT target species, evaluated against the Level-2 performance requirements for uncertainty and spatial resolution. Colours and numbers indicate the mean performance relative to the goal and threshold geophysical requirements (performance categories of Table 3). Left: estimated retrieval uncertainty due to NESR only (top), all error sources except NESR (middle), and their quadratic combination (bottom). Right: horizontal spatial resolution (top), vertical spatial resolution (middle), and combined spatial resolution as defined in Sect. 3.1 (bottom).



2. Through the formulation of 2D tomographic averaging kernels and automated error propagation chains, the LPE enables a comprehensive quantification of spatial resolution alongside a detailed accounting of the L2 error budget, distinguishing between random, calibration-induced, and systematic uncertainties. By applying these analytical metrics to high-resolution atmospheric models, the FL2S efficiently emulates large data volumes of synthetic L2 products which can be used for scientific impact assessments or Observation System Simulation Experiments.
3. The underlying mathematical framework remains highly flexible with respect to the underlying retrieval setups. By adjusting horizontal/vertical discretization and tuning regularization matrices (such as Tikhonov-Phillips constraints), the framework acts as an efficient testbed for trade-off analyses, allowing developers to balance spatial resolution against noise and to optimize the overall retrieval chain prior to full operational deployment. Although not tested here, the method further supports retrievals operating in the logarithmic volume mixing ratio space.
4. Although developed to mature the CAIRT concept during Phases 0/A, the modular architecture of the database, the L1b Simulator, the LPE, and the FL2S is fundamentally instrument-agnostic. The framework can be easily adapted to simulate other multi-dimensional remote sensing geometries, spectral resolutions, or prospective satellite missions, offering a computationally streamlined blueprint for future spaceborne instrument design and target characterization.

In summary, while CAIRT was not selected for immediate implementation in the EE11 competition, the mission concept remains scientifically and technically mature. The modeling framework detailed in this study provides a powerful and computationally lightweight suite of tools that stands ready to support future implementation opportunities for next-generation limb-sounding missions.

Code and data availability. The LPE is available as open source IDL code at <https://gitlab.kit.edu/kit/cairt/lpe>. The Python-based L1b Simulator is available as open source at https://gitlab.kit.edu/kit/cairt/open_instsim. The FL2S is available as open source Python code at https://jugit.fz-juelich.de/cairt-scirec/cairt_fl2s. The KOPRA and GRANADA non-LTE radiative transfer packages, utilized for the SJD generation, as well as the RCP retrieval package integrated into the LPE, are available as open source Fortran 90 code at https://gitlab.kit.edu/kit/cairt/open_kopra_granada_rcp.git. A tool for computing tomographic 3-D CAIRT retrieval grids is available at https://jugit.fz-juelich.de/cairt-scirec/sim_CAIRT_grid. The LPE and FL2S output data used to generate the results shown in Figs. 5–18 are archived and available via the Zenodo repository at <https://doi.org/10.5281/zenodo.20814017> and <https://doi.org/10.5281/zenodo.21072779>, respectively.

Author contributions. BF had the original idea and is responsible for the conceptual design of the LPE and FL2S-based performance evaluation, coded the LPE and the first version of the FL2S, and generated Figures 2, 4–16. MH contributed to the conceptual design, coded the L1b simulator, set up the SJD database, and generated Figures 1 and 17. BF and MH prepared the manuscript. QE compiled the ERS climatologies. PR developed the cloud-masking algorithm implemented in the FL2S. MH, SB and QE generated the python-based FL2S code. NP and HL generated and provided the WACCM-X high-resolution simulation used in this study as test scenario. All authors reviewed the full manuscript, contributed to writing and discussions, and supported the analysis of results.



Competing interests. At least one of the (co-)authors is a member of the editorial board of Atmospheric Measurement Techniques. The authors have no other competing interests to declare.

600 *Acknowledgements.* This work was supported by ESA via the CAIRT Performance and Requirement Consolidation study (PerReC) under the ESA/ESTEC contract 4000136480/21/NL/LF CCM1. The IAA team (BF and SB) acknowledges financial support from the Agencia Estatal de Investigación, MCIN/AEI/10.13039/501100011033, through grant nos. PID2022-141216NB-I00 and CEX2021-001131-S. The WACCM-X generated test scenario is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the U.S. National Science Foundation under Cooperative Agreement 1852977.



605 References

- Ceccherini, S., Carli, B., and Raspollini, P.: Equivalence of data fusion and simultaneous retrieval, *Optics Express*, 23, 8476, <https://doi.org/10.1364/OE.23.008476>, 2015.
- Ceccherini, S., Zopetti, N., and Carli, B.: An improved formula for the complete data fusion, *Atmospheric Measurement Techniques*, 15, 7039–7048, <https://doi.org/10.5194/amt-15-7039-2022>, 2022.
- 610 Errera, Q., Op de beeck, M., Bender, S., Flemming, J., Funke, B., Hoffmann, A., Höpfner, M., Livesey, N., Poli, G., Pieroux, D., Raspollini, P., and Sinnhuber, B.-M.: On the Capability of the Changing Atmosphere Infra-Red Tomography Explorer (CAIRT) Candidate Mission to Constrain O₃ and H₂O in the Upper Troposphere and Lower Stratosphere, 19, 2601–2620, <https://doi.org/10.5194/amt-19-2601-2026>, 2026.
- Errera, Q., Funke, B., Hoffmann, A., Höpfner, M., Raspollini, P., Sinnhuber, B.-M., and Walker, K.: Extended Reference Scenarios (ERS) for the ESA Earth Explorer 11 candidate CAIRT (Changing Atmosphere Infra-Red Tomography explorer), *Earth Syst. Sci. Data*, in prep.
- 615 European Space Agency: Report for Mission Selection: Earth Explorer 11 Candidate Mission CAIRT, <https://doi.org/10.5281/ZENODO.15606819>, 2025.
- Friedl-Vallon, F., Gulde, T., Hase, F., Kleinert, A., Kulesa, T., Maucher, G., Neubert, T., Olschewski, F., Piesch, C., Preusse, P., Rongen, H., Sartorius, C., Schneider, H., Schönfeld, A., Tan, V., Bayer, N., Blank, J., Dapp, R., Ebersoldt, A., Fischer, H., Graf, F., Guggenmoser, T., Höpfner, M., Kaufmann, M., Kretschmer, E., Latzko, T., Nordmeyer, H., Oelhaf, H., Orphal, J., Riese, M., Schardt, G., Schillings, J., Sha, M. K., Suminska-Ebersoldt, O., and Ungermann, J.: Instrument concept of the imaging Fourier transform spectrometer GLORIA, *Atmospheric Measurement Techniques*, 7, 3565–3577, <https://doi.org/10.5194/amt-7-3565-2014>, 2014.
- 620 Funke, B., López-Puertas, M., García-Comas, M., Stiller, G. P., von Clarmann, T., Höpfner, M., Glatthor, N., Grabowski, U., Kellmann, S., and Linden, A.: Carbon monoxide distributions from the upper troposphere to the mesosphere inferred from 4.7 μ m non-local thermal equilibrium emissions measured by MIPAS on Envisat, *Atmospheric Chemistry and Physics*, 9, 2387–2411, <https://doi.org/10.5194/acp-9-2387-2009>, 2009.
- Funke, B., López-Puertas, M., García-Comas, M., Kaufmann, M., Höpfner, M., and Stiller, G. P.: GRANADA: a Generic RAdiative traNsfer AnD non-LTE population Algorithm, *J. Quant. Spectrosc. Radiat. Transfer*, 113, 1771–1817, <https://doi.org/10.1016/j.jqsrt.2012.05.001>, 2012.
- 630 García-Comas, M., Funke, B., López-Puertas, M., and Höpfner, M.: Non-LTE error assessment for CAIRT, <https://doi.org/10.5281/zenodo.19106615>, 2023.
- Kiefer, M., von Clarmann, T., Funke, B., García-Comas, M., Glatthor, N., Grabowski, U., Kellmann, S., Kleinert, A., Laeng, A., Linden, A., López-Puertas, M., Marsh, D. R., and Stiller, G. P.: IMK/IAA MIPAS temperature retrieval version 8: nominal measurements, *Atmospheric Measurement Techniques*, 14, 4111–4138, <https://doi.org/10.5194/amt-14-4111-2021>, 2021.
- 635 Krasauskas, L., Ungermann, J., Ensmann, S., Krisch, I., Kretschmer, E., Preusse, P., and Riese, M.: 3-D tomographic limb sounder retrieval techniques: irregular grids and Laplacian regularisation, *Atmos. Meas. Tech.*, 12, 853–872, <https://doi.org/10.5194/amt-12-853-2019>, 2019.
- Liu, H.-L., Lauritzen, P. H., and Vitt, F.: Impacts of Gravity Waves on the Thermospheric Circulation and Composition, *Geophysical Research Letters*, 51, e2023GL107453, <https://doi.org/10.1029/2023GL107453>, _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1029/2023GL107453>, 2024.
- 640



- López-Puertas, M., Funke, B., Jurado-Navarro, A. A., García-Comas, M., Gardini, A., Boone, C. D., Rezac, L., and Garcia, R. R.: Validation of the MIPAS CO₂ volume mixing ratio in the mesosphere and lower thermosphere and comparison with WACCM simulations, *Journal of Geophysical Research: Atmospheres*, 122, 8345–8366, <https://doi.org/10.1002/2017JD026805>, <https://onlinelibrary.wiley.com/doi/pdf/10.1002/2017JD026805>, 2017.
- 645 Norton, R. H. and Beer, R.: New apodizing functions for Fourier spectrometry, *JOSA*, 66, 259–264, <https://doi.org/10.1364/JOSA.66.000259>, publisher: Optica Publishing Group, 1976.
- Norton, R. H. and Beer, R.: Errata: New Apodizing Functions For Fourier Spectrometry, *JOSA*, 67, 419–419, <https://doi.org/10.1364/JOSA.67.000419>, publisher: Optica Publishing Group, 1977.
- Okui, H., Sato, K., Koshin, D., and Watanabe, S.: Formation of a Mesospheric Inversion Layer and the Subsequent Elevated Stratopause Associated With the Major Stratospheric Sudden Warming in 2018/19, *Journal of Geophysical Research: Atmospheres*, 126, e2021JD034681, <https://doi.org/10.1029/2021JD034681>, 2021.
- 650 Phillips, D.: A Technique for the numerical solution of certain integral equations of first kind, *J. Ass. Comput. Mat.*, 9, 84–97, 1962.
- Purser, R. J. and Huang, H.-L.: Estimating Effective Data Density in a Satellite Retrieval or an Objective Analysis, *Journal of Applied Meteorology and Climatology*, 32, 1092–1107, [https://doi.org/10.1175/1520-0450\(1993\)032<1092:EEDDIA>2.0.CO;2](https://doi.org/10.1175/1520-0450(1993)032<1092:EEDDIA>2.0.CO;2), 1993.
- 655 Riese, M., Oelhaf, H., Preusse, P., Blank, J., Ern, M., Friedl-Vallon, F., Fischer, H., Guggenmoser, T., Höpfner, M., Hoor, P., Kaufmann, M., Orphal, J., Plöger, F., Spang, R., Suminska-Ebersoldt, O., Ungermann, J., Vogel, B., and Woiwode, W.: Gimballed Limb Observer for Radiance Imaging of the Atmosphere (GLORIA) scientific objectives, *Atmospheric Measurement Techniques*, 7, 1915–1928, <https://doi.org/10.5194/amt-7-1915-2014>, 2014.
- Rodgers, C. D.: Inverse methods for atmospheric sounding: theory and practice, no. 2 in Series on atmospheric, oceanic and planetary physics, World Scientific, Singapore, ISBN 978-981-02-2740-1, 2000.
- 660 Sellitto, P., Kosary, M., Höpfner, M., Funke, B., Hoffmann, A., Ungermann, J., Errera, Q., Tilmes, S., and Sinnhuber, B.-M.: The Enhanced Capabilities of Mid-Infrared Limb Emission Sounding to Observe Stratospheric Aerosol Injection Geoengineering Interventions, *EGU-sphere*, pp. 1–22, <https://doi.org/10.5194/egusphere-2026-919>, 2026.
- Sinnhuber, B.-M., Friedl-Vallon, F., Chipperfield, M. P., Preusse, P., Errera, Q., Funke, B., Godin-Beekmann, S., García-Comas, M., Höpfner, M., Plöger, F., Riese, M., Stiller, G., von Clarmann, T., and Walker, K.: The need for a new comprehensive view of the middle atmosphere – the potential of infrared limb imaging tomography, *Bull. Amer. Meteor. Soc.*, <https://ametsoc.org>, accepted Author Manuscript, In Press, 2026.
- 665 Stiller, G. P.: The Karlsruhe Optimized and Precise Radiative Transfer Algorithm (KOPRA), Tech. rep., Forschungszentrum Karlsruhe, <https://doi.org/10.5445/IR/270048971>, 22.02.11; LK 01; Wissenschaftliche Berichte, FZKA-6487 (Dezember 2000), 2000.
- 670 Tikhonov, A. N.: On the solution of ill-posed problems and the method of regularization, *Dokl. Akad. Nauk SSSR*, 151, 501–504, 1963.
- Tirelli, C., Ceccherini, S., Del Bianco, S., Funke, B., Höpfner, M., Cortesi, U., and Raspollini, P.: Extension of the Complete Data Fusion algorithm to tomographic retrieval products, *Atmospheric Measurement Techniques*, 18, 5619–5636, <https://doi.org/10.5194/amt-18-5619-2025>, 2025.
- Twomey, S.: On the Numerical Solution of Fredholm Integral Equations of the First Kind by the Inversion of the Linear System Produced by Quadrature, *Journal of the ACM*, 10, 97–101, <https://doi.org/10.1145/321150.321157>, 1963.
- 675 von Clarmann, T., Höpfner, M., Kellmann, S., Linden, A., Chauhan, S., Funke, B., Grabowski, U., Glatthor, N., Kiefer, M., Schieferdecker, T., Stiller, G. P., and Versick, S.: Retrieval of temperature, H₂O, O₃, HNO₃, CH₄, N₂O, ClONO₂ and ClO from MIPAS reduced resolution



- nominal mode limb emission measurements, *Atmospheric Measurement Techniques*, 2, 159–175, <https://doi.org/10.5194/amt-2-159-2009>, 2009.
- 680 von Clarmann, T., Degenstein, D. A., Livesey, N. J., Bender, S., Braverman, A., Butz, A., Compernelle, S., Damadeo, R., Dueck, S., Eriksson, P., Funke, B., Johnson, M. C., Kasai, Y., Keppens, A., Kleinert, A., Kramarova, N. A., Laeng, A., Langerock, B., Payne, V. H., Rozanov, A., Sato, T. O., Schneider, M., Sheese, P., Sofieva, V., Stiller, G. P., von Savigny, C., and Zawada, D.: Overview: Estimating and reporting uncertainties in remotely sensed atmospheric composition and temperature, *Atmospheric Measurement Techniques*, 13, 4393–4436, <https://doi.org/10.5194/amt-13-4393-2020>, 2020.
- 685 von Clarmann, T., Glatthor, N., Grabowski, U., Funke, B., Kiefer, M., Kleinert, A., Stiller, G. P., Linden, A., and Kellmann, S.: TUNER-compliant error estimation for MIPAS: methodology, *Atmospheric Measurement Techniques*, 15, 6991–7018, <https://doi.org/10.5194/amt-15-6991-2022>, 2022.
- Wetzel, G., Kleinert, A., Johansson, S., Friedl-Vallon, F., Höpfner, M., Ungermann, J., Neubert, T., Catoire, V., Crevoisier, C., Engel, A., Gulde, T., Jacquet, P., Kirner, O., Kretschmer, E., Kulesa, T., Laube, J. C., Maucher, G., Nordmeyer, H., Piesch, C., Preusse, P., Retzlaff, 690 M., Schardt, G., Schillings, J., Schneider, H., Schönfeld, A., Schuck, T., Woiwode, W., Riese, M., and Braesicke, P.: Intercomparison and validation of first GLORIA-B measurements of stratospheric and upper tropospheric long-lived tracers and photochemically active species, *Atmospheric Measurement Techniques*, 18, 5873–5894, <https://doi.org/10.5194/amt-18-5873-2025>, 2025.