Standardized Definition and Reporting of Vertical Resolution and Uncertainty in the NDACC Lidar Ozone and Temperature Algorithms

Part 2: Uncertainties

T. Leblanc [1], R. Sica [2], A. van Gijsel [3], S. Godin-Beekmann [4], A. Haefele [5], T. Trickl [6], G. Payen [7], and G. Liberti [8]

[1] Jet Propulsion Laboratory, California Institute of Technology, Wrightwood, CA 92397, USA
[2] The niversity of Western Ontario, London, Canada
[3] Royal Netherlands Meteorological Institute (KNMI), Netherlands
[4] LATMOS-IPSL, CNRS-INSU, Paris, France
[5] Meteoswiss, Payerne, Switzerland
[6] Karlsruher Institut für Technologie, IMK-IFU, Garmisch-Partenkirchen, Germany
[7] Observatoire des Sciences de l'Univers de La Réunion, CNRS and Université de la Réunion (UMS3365), Saint Denis de la Réunion, France
[8] ISAC-CNR, Via Fosso del Cavaliere 100, I-00133, Rome, Italy

Table of Contents

1 Introduction

Not all lidar techniques are concerned by the present results on uncertainty. For ozone, the technique covered in this report is the DIAL (Differential Absorption Lidar) technique (Mégie et al., 1977), using Rayleigh or vibrational Raman backscatter. DIAL is the only known lidar technique to date to measure ozone. For temperature, the only technique covered in this report is the traditional density integration method, as described by Hauchecorne and Chanin (1980), using Rayleigh or vibrational Raman backscatter. The less common rotational-Raman technique (Archinov et al., 1983), and the recent density integration technique using an Optimal Estimation Method (Sica and Haefele, 2014) are not concerned by this report. In the rest of this report, every single mention of "temperature lidar" refers to the traditional density integration technique.

Uncertainties in lidar measurements have been discussed since the early development of the lidar technique for the measurement of atmospheric parameters. In the case of the DIAL technique, earlier publications dealt with the optimization of the wavelengths pairs for tropospheric and stratospheric ozone measurements taking into account the measurement's error budget (e.g. Mégie and Menzies, 1980; Pelon and Mégie, 1982). In the frame of the Network for the Detection of Atmospheric Composition Changes (NDACC), various groups have set up lidar instruments for the measurements of ozone and temperature vertical distribution. They have generally described their lidar systems with a detailed assessment of the measurement errors (e.g. Uchino et al., 1991; McDermid et al., 1990; Papayannis et al., 1990; McGee et al., 1991; Godin-Beckmann et al., 2003). In addition, inter-comparison campaigns set up in the frame of NDACC have assessed the evaluation of lidar measurement uncertainties (Keckhut et al., 2004; see also http://ndacc-lidar.org/ for more information on NDACC lidars).

The treatment of uncertainty in the ozone and temperature lidar retrievals depends on the choice of the theoretical equations used as well as their implementation to the real world, i.e., after considering all the caveats associated with the design, setup, and operation of an actual lidar instrument. To retrieve an ozone number density profile in the troposphere or stratosphere using the DIAL technique and a temperature profile using the density integration technique, we start from the same initial theoretical model, namely the *Lidar Equation* (e.g., Hinkley, 1976). This equation in its most compressed form describes the emission of light by a laser source, its backscatter at altitude z, its extinction and scattering along its path up and back, and its collection back on a detector:

$$P(z,\lambda_E,\lambda_R) = P_L(\lambda_E)\frac{\eta(\lambda_R,z)\delta z}{(z-z_L)^2}\tau_{UP}(z,\lambda_E)\beta(z,\lambda_E,\lambda_R)\tau_{DOWN}(z,\lambda_R) \quad \textbf{(1.1)}$$

λ_E is the laser emission wavelength and λ_R is the receiver detection wavelength

P is the total number of photons collected at wavelength λ_R on the lidar detector surface

δz is the thickness of the backscattering layer sounded during the time interval δt ($\delta z = c\delta t/2$, where c is the speed of light)

P_L is the number of photons emitted at the emission wavelength λ_E

η is the optical efficiency of the receiving channel, including optical and spectral transmittance and geometric obstruction

z is the altitude of the backscattering layer

z_L is the altitude of the lidar (laser and receiver assumed to be at the same altitude)

β is the total backscatter coefficient (including particulate β_P and molecular β_M backscatter)

τ_{UP} is the optical thickness integrated along the outgoing beam path between the lidar and the scattering altitude z, and is defined as:

$$\tau_{UP}(z) = \exp\left[-\int_{z_L}^{z} \left(\sigma_M(\lambda_E) N_a(z') + \alpha_P(z', \lambda_E) + \sum_i \sigma_i(z', \lambda_E) N_i(z') \right) dz' \right] \quad \textbf{(1.2)}$$

τ_{DOWN} is the optical thickness integrated along the returning beam path between the scattering altitude z and the lidar, and is defined as:

$$\tau_{DOWN}(z) = \exp\left[-\int_{z_L}^{z} \left(\sigma_M(\lambda_R) N_a(z') + \alpha_P(z', \lambda_R) + \sum_i \sigma_i(z', \lambda_R) N_i(z') \right) dz' \right] \quad \textbf{(1.3)}$$

σ_M is the Rayleigh extinction cross-section, N_a is the air number density, α_P is the particulate extinction coefficient, σ_i is the absorption cross-section of absorbing constituent i, and N_i is the number density of absorbing constituent i. For the altitude range of interest of the ozone DIAL and temperature lidar measurements, the Rayleigh cross-sections can be considered constant with altitude, and therefore depend only on wavelength. The absorption cross-sections however are temperature-dependent, and should therefore be considered a function of both altitude and wavelength.

Ozone number density is retrieved by reverting **Eq. (1.1)** with respect to the absorption term $\sigma_i N_i$, while temperature is retrieved by reverting **Eq. (1.1)** with respect to the backscatter term β.

1.1 Ozone DIAL retrieval

For the ozone DIAL retrieval, **Eq. (1.1)** is solved for the ozone absorption contribution of $\sigma_i N_i$ by taking the vertical derivative of the logarithm of the lidar signals measured at two different wavelengths, the light at one wavelength being more absorbed by ozone than the light at the other wavelength (Mégie et al., 1977). Using the notation "*ON*" for the most absorbed wavelength, and "*OFF*" for the least absorbed wavelength, inverted **Eq. (1.1)** takes the form:

$$N_{O3}(z) = \frac{1}{\Delta\sigma_{O3}(z)} \left[\frac{\partial}{\partial z} \left(\ln \frac{P_{OFF}(z)}{P_{ON}(z)} \right) - \Delta\sigma_M N_a(z) - \left(\sum_{ig} \Delta\sigma_{ig}(z) N_{ig}(z) \right) - \Delta\alpha_P(z) + \Lambda\eta(z) + \Lambda\beta(z) \right] \quad \textbf{(1.4)}$$

$N_{O3}(z)$ is the ozone number density retrieved at altitude z. The ozone absorption cross-section differential $\Delta\sigma_{O3}$ is given by:

$$\Delta\sigma_{O3}(z) = \sigma_{O3}(z, \lambda_1) + \sigma_{O3}(z, \lambda_3) - \sigma_{O3}(z, \lambda_2) + \sigma_{O3}(z, \lambda_4) \quad \textbf{(1.5)}$$

The following wavelength definitions have been used:

λ_1 is the emitted "*ON*" wavelength

λ_2 is the emitted "*OFF*" wavelength

λ_3 is the received "*ON*" wavelength

λ_4 is the received "*OFF*" wavelength

For elastic (Rayleigh) scattering, the emitted and received wavelengths are identical yielding $\lambda_1 = \lambda_3$ and $\lambda_2 = \lambda_4$. For inelastic scattering, the emitted and received wavelengths are different, and all 4 terms (2 terms up and 2 terms down) are different (McGee et al., 1993). A list of most commonly used DIAL wavelength pairs for the measurement of stratospheric and tropospheric ozone is presented in **Table 1.1.**

Table 1.1 List of most commonly used ozone DIAL wavelength pairs

λ_1 "ON" (nm)	λ_2 "OFF" (nm)	λ_3 "ON" (nm)	λ_4 "OFF" (nm)	Backscatter technique	Application	Light source details (λ_1)	Light source details (λ_2)
266	289	266	289	Rayleigh	Troposphere	Nd:YAG quadrupled 266 nm non-shifted	Quadrupled Nd:YAG 266 nm Raman-shifted
277	291	277	291	Rayleigh	Troposphere	Excimer KrFl 248 nm Raman-shifted	Excimer KrFl 248 nm Raman-shifted
277	313	277	313	Rayleigh	Troposphere	Excimer KrFl 248 nm Raman-shifted	Excimer KrFl 248 nm Raman-shifted
287	294	287	294	Rayleigh	Troposphere	Ce:LiCAF tunable 263 nm tuned	Ce:LiCAF tunable 263 nm tuned
289	299	289	299	Rayleigh	Troposphere	Nd:YAG quadrupled 266 nm Raman-shifted	Quadrupled Nd:YAG 266 nm Raman-shifted
299	316	299	316	Rayleigh	Troposphere	Nd:YAG quadrupled 266 nm Raman-shifted	Quadrupled Nd:YAG 266 nm Raman-shifted
308	353	308	353	Rayleigh	Stratosphere	Excimer XeCl 308 nm non-shifted	Excimer XeCl 308 nm Raman-shifted
308	355	308	355	Rayleigh	Stratosphere	Excimer XeCl 308 nm non-shifted	Nd:YAG tripled 355 nm non-shifted
308	353	332	385	N_2 Raman	Stratosphere	Excimer XeCl 308 nm non-shifted	Excimer XeCl 308 nm Raman-shifted
308	355	332	387	N_2 Raman	Stratosphere	Excimer XeCl 308 nm non-shifted	Nd:YAG tripled 355 nm non-shifted

P_{ON} and P_{OFF} are the number of photons collected on the detectors of the "*ON*" and "*OFF*" channels respectively. $N_a(z)$ is the air number density and $\Delta\sigma_M$ is the Rayleigh extinction cross-section differential between the "ON" and "OFF" wavelengths computed along the beam path up to altitude z and back:

$$\Delta\sigma_M = \sigma_M(\lambda_1) + \sigma_M(\lambda_3) - \sigma_M(\lambda_2) - \sigma_M(\lambda_4) \quad \textbf{(1.6)}$$

N_{ig} is the number density of absorbing constituent *ig*, and $\Delta\sigma_{ig}$ is the absorption cross-section differential of constituent *ig* along the beam path up to altitude z and back, and defined as:

$$\Delta\sigma_{ig}(z) = \sigma_{ig}(z,\lambda_1) + \sigma_{ig}(z,\lambda_3) - \sigma_{ig}(z,\lambda_2) - \sigma_{ig}(z,\lambda_4) \quad \textbf{(1.7)}$$

$\Delta\alpha_P$ is the extinction differential due to particles and computed along the beam path up to altitude z and back:

$$\Delta\alpha_P(z) = \alpha_P(z,\lambda_1) + \alpha_P(z,\lambda_3) - \alpha_P(z,\lambda_2) + \alpha_P(z,\lambda_4) \quad \textbf{(1.8)}$$

This term depends strongly on the type of particulate matter, and is difficult to estimate for typical ozone and temperature lidar instruments that do not have dedicated aerosol channels (multi-wavelength, polarization, etc.).

Finally, $\Lambda\eta$ and $\Lambda\beta$ are defined as:

$$\Lambda\eta(z) = \frac{\partial}{\partial z}\left(\ln\frac{\eta_{ON}(z)}{\eta_{OFF}(z)}\right) \quad \textbf{(1.9)}$$

$$\Lambda\beta(z) = \frac{\partial}{\partial z}\left(\ln\frac{\beta_{ON}(z)}{\beta_{OFF}(z)}\right) \quad \textbf{(1.10)}$$

η_{ON} and η_{OFF} are the optical efficiencies of the "ON" and "OFF" channels respectively, including optical and spectral transmittance and geometric obstruction, and β_{ON} and β_{OFF} are the total backscatter coefficients at the "ON" and "OFF" wavelength respectively.

1.2 Temperature retrieval (density integration technique)

For the temperature retrieval (integration technique), **Eq. (1.1)** is solved for the molecular contribution of the backscatter coefficient β (Hauchecorne and Chanin, 1980). In the absence of particulate backscatter, the backscatter coefficient, and therefore the lidar signal collected on the detector, is proportional to air number density. **Eq. (1.1)** can then be inverted and a temperature profile can be calculated by vertically integrating air number density from the top of the profile assuming hydrostatic balance, and that the air is an ideal gas with a constant mean molecular mass:

$$T(z-\delta z) = \frac{P_C(z)}{P_C(z-\delta z)}T(z) + \frac{M_a}{R_a P_C(z-\delta z)}\overline{P_C}(z)\overline{g}(z)\delta z \quad \textbf{(1.11)}$$

T is the retrieved temperature, M_a is the molecular mass of dry air, R_a is the ideal gas constant, and g is the acceleration of gravity. The horizontal bar above P_C and g refers to the average value of P_C and g between z and z-δz.

P_C relates to the number of photons collected on the lidar detector P by the equation:

$$P_C(z) = \frac{P(z)\left(z - z_L\right)^2}{\eta(\lambda_R, z)\tau_{UP}(z,\lambda_E)\tau_{DOWN}(z,\lambda_R)} \quad \textbf{(1.12)}$$

The emitted and received wavelengths are identical for elastic (Rayleigh) scattering, and different for inelastic (vibrational Raman) scattering (Strauch et al., 1971; Gross et al., 1997). A

list of most commonly used wavelengths for the measurement of temperature is presented in **Table 1.2**.

Table 1.2 List of most commonly used backscatter temperature lidar wavelengths

λ_E (nm)	λ_R (nm)	Backscatter technique	Application	Light source details (λ_E)
353	353	Rayleigh	30 < z < 100 km	Excimer XeCl 308 nm Raman-shifted
353	385	N_2 Raman	10 < z < 40 km	Excimer XeCl 308 nm Raman-shifted
355	355	Rayleigh	30 < z < 100 km	Nd:YAG tripled 355 nm non-shifted
355	387	N_2 Raman	10 < z < 40 km	Nd:YAG tripled 355 nm non-shifted
532	532	Rayleigh	30 < z < 110 km	Nd:YAG doubled 532 nm non-shifted
532	608	N_2 Raman	10 < z < 40 km	Nd:YAG doubled 532 nm non-shifted

1.3 From theory to actual measurements with uncertainty

Eqs. (1.4)-(1.12) relate to the number of photons collected on the lidar detectors rather than the raw lidar signals recorded in the data files. Their practical implementation for the retrieval of ozone and temperature therefore implies the inclusion of several signal correction procedures and numerical transformations related to the instrumentation. The detectors' quantum efficiency and the effects of the data recorders, namely the sky and electronic background noise and the signal saturation, must be taken into account. Due to the diversity of lidar instrumentation, it is not possible to provide a single expression for the parameterization of these effects and obtain a unique, real-world version of **Eqs. (1.4)-(1.12)**. However, a few commonly-used expressions are provided for reference in this report.

Using the metrology vocabulary (introduced in **chapter 2**), **Eq. (1.4)** modified for instrumental effects is called the "measurement model" for the retrieval of ozone using the DIAL technique. Similarly, **Eqs. (1.11)-(1.12)** modified for instrumental effects form the "measurement model" for the retrieval of temperature using the traditional density integration technique. These measurement models constitute the mathematical architecture around which ozone and temperature uncertainty budgets should be built. However it is shown later in this report that, due to their complexity, these measurement models should be split into multiple, more simple sub-models through which uncertainty can be more easily propagated. **Chapter 2** reviews the fundamentals of the expression and propagation of uncertainty. **Chapter 3** provides specific recommendations on how to apply these fundamentals to the case of the NDACC lidars. Based on the definitions and recommendations of **chapters 2-3**, the expression of uncertainty and its step-by-step propagation through the ozone and temperature data processing chains are described in details in **chapters 4-6**. All aspects of data processing common to both ozone and temperature retrievals are described in **chapter 4**. All aspects of data processing specific to the exclusive retrieval of ozone are described in **chapter 5**, and all aspects of data processing specific to the exclusive retrieval of temperature are described in **chapter 6**. The validity of the approach and correctness of expressions provided in **chapters 4-6** were quantitatively verified using simulated

lidar signals and Monte Carlo experiments (introduced in **appendix A**). The results of these experiments are described in details in **appendix B** for ozone and **appendix C** for temperature.

Several other aspects closely related to the ozone and temperature lidar uncertainty budgets were reviewed by the ISSI Team. The results are reported in seven appendices. **Appendix D** briefly reviews approximations of the molecular scattering expressions. **Appendix E** is an assessment of the latest published ozone cross-section datasets. **Appendix F** provides suggested ways to compute uncertainty in the ancillary measurements based on their correlative characteristics. **Appendix G** illustrates with a simple example the complex problem of the propagation of uncertainty owed to systematic effects. **Appendix H** provide insights on the computation of saturation uncertainty when using a paralyzable system. **Appendix I** briefly reviews the computation of fitting coefficient uncertainty for the least-squares and Singular Value Decomposition methods. Finally, **Appendix J** provides a possible pathway to compute uncertainty owed to incomplete overlap in a standardized manner.

2 Reference definitions from the BIPM

The definitions of uncertainty used in this work and recommended to be used for all NDACC lidar measurements originate in the two internationally recognized reference documents endorsed by the Bureau International des Poids et Mesures (BIPM), namely the International Vocabulary of Basic and General Terms in Metrology (commonly abbreviated "VIM") (JCGM 200, 2012), and the Guide to the Expression of Uncertainty in Measurement (commonly abbreviated "GUM") (JCGM 100, 2008). These two documents and their supplements provide a complete framework to the treatment of uncertainty. In the present chapter, we describe how this framework is adapted to our case of ozone and temperature lidar measurements, and what the resulting recommendations are for a practical implementation within all NDACC ozone and temperature lidars.

2.1 Standard uncertainty

Starting from the source definition the word "uncertainty" means doubt, and thus in its broadest sense, "uncertainty of measurement" means doubt about the validity of the result of a measurement. In metrological sense (article 2.26 of the VIM) (JCGM 200, 2012), it is a "*non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand based on the information used*". Measurement uncertainty includes components arising from systematic effects, as well as the definitional (or "intrinsic") uncertainty, i.e., the practical minimum uncertainty achievable in any measurement. It may be a standard deviation or the half-width of an interval with a stated coverage probability. The particular case of "standard uncertainty" u is defined in article 2.30 of the VIM (JCGM 200, 2012), as "*the measurement uncertainty expressed as a standard deviation*".

Some measurement uncertainty components can be evaluated by a so-called "Type A evaluation", i.e., from the statistical distribution of a series of measurements of the quantity, and can be characterized by standard deviations. Other components may be evaluated by a so-called "Type B evaluation", i.e., from probability density functions based on experience or other information, and can also be characterized by standard deviations. This latter type includes all evaluation means other than Type A. Both types of evaluation are based on probability distributions and the uncertainty components resulting from either type are quantified by variances or standard deviations (JCGM 100, 2008). The estimated variance u^2 characterizing an uncertainty component obtained from a Type A evaluation is calculated from series of repeated observations and is the statistically estimated variance σ^2. The estimated standard deviation u, the positive square root of u^2, is thus $u = \sigma$ and for convenience is often called a Type A standard uncertainty. For an uncertainty component obtained from a Type B evaluation, the estimated variance u^2 is evaluated using available knowledge, and the estimated standard deviation u is called a Type B standard uncertainty. Thus a Type A standard uncertainty is obtained from a probability density function derived from an observed frequency distribution, while a Type B standard uncertainty is obtained from an assumed probability density function based on the degree of belief that an event will occur, using best available knowledge. In the present work (ozone and temperature lidar measurements) both types of evaluation have been identified and will be discussed in the detailed list of sources presented in chapter 4.

As explained in art. 3.2 of the GUM (JCGM 100, 2008) it is essential to distinguish between the terms "error" and "uncertainty". "*They are not synonyms, but represent completely different*

concepts; they should not be confused with one another or misused". Error is an idealized concept and cannot be known exactly. The experimental standard deviation of the average of a series of observations is not the random error of the mean. It is instead a measure of the uncertainty of the mean due to random effects. The exact value of the error in the mean arising from these effects "*cannot be known*" (JCGM 100, 2008).

It is also important to avoid using the term "systematic uncertainty". Instead, the term "uncertainty arising from systematic effects" should be used. One significant advantage of adopting the vocabulary of the VIM (JCGM 200, 2012) is the ability to treat mathematically uncertainties arising from both systematic and random effects in a similar way. The only distinction in this treatment is made when dealing with either correlated or uncorrelated variables. On a practical point of view however, keeping a clear distinction between uncertainty components arising from systematic effects and those arising from random effects is crucial because it provides the physical or experimental basis for the consideration of either correlated and uncorrelated components in the implementation of an uncertainty budget, which eventually paves the path towards a successful analytical propagation of uncertainties.

2.2 Combined standard uncertainty

The combined standard uncertainty is defined in article art. 2.31 of the VIM (JCGM 200, 2012) as the "*standard measurement uncertainty that is obtained using the individual standard measurement uncertainties associated with the input quantities in a measurement model*". If some of these input quantities are correlated, covariances must be taken into account. A measurement model is defined as a "*mathematical relation among all quantities known to be involved in a measurement*" (VIM art. 2.48 (JCGM 200, 2012)). The measurement model can be written:

$$Y = f(X_1, X_2, ..., X_N,) \tag{2.1}$$

where Y is the output quantity in the measurement model (the measurand), and the X_n (n=1,2,..,N) are the input quantities in this model. If there is more than one output quantity, then the covariances of the output quantities must be provided together with their uncertainty.

Referring to our case of ozone DIAL introduced in **chapter 1**, **Eq. (1.4)** modified for instrumental effects is called the "measurement model" for the retrieval of ozone using the DIAL technique. The "output quantity" is ozone number density (left-hand side of **Eq. (1.4)**), while the input quantities are all the variables introduced on the right-hand side of the equation, as well as the instrumental parameters impacting the recorded signals. Similarly, **Eqs. (1.11)-(1.12)** modified for instrumental effects is the "measurement model" for the retrieval of temperature using the density integration technique. The "output quantity" is temperature (left-hand side of **Eq. (1.11)**), while the input quantities are all the variables introduced on the right-hand side of **Eqs. (1.11)-(1.12)** as well as the instrumental parameters impacting the recorded signals.

The true values of a model input quantities X_n are unknown. These quantities are characterized by probability distributions and should be treated mathematically as random variables (art. 3.3.5 of the GUM (JCGM 100, 2008)). These distributions describe the respective probabilities of their true values lying in different intervals, and are assigned based on available knowledge. Some, or all of them may occasionally be interrelated and the relevant distributions, known as joint-

distributions, apply to these quantities taken together (JCGM 104, 2009). For the sake of practicability, in the present work we have made specific efforts to ensure that none of the input quantities introduced in the ozone and temperature models are interrelated (see **chapters 4-6**).

Knowledge about an input quantity X_n is inferred from repeated indication values (Type A evaluation) or scientific judgment or other information concerning the possible values of the quantity (Type B evaluation). In Type A evaluations the default assumption is made that the distribution best describing an input quantity X_n is a Gaussian distribution. When the uncertainty is evaluated from a small number of indication values the corresponding distribution can be taken as a t-distribution. For a Type B evaluation, the only available information is that X_n lies in a specified interval [a,b]. In such a case, knowledge of the quantity can be characterized by a rectangular probability distribution with limits a and b. If different information is available, a probability distribution consistent with that information should be used (JCGM 104, 2009).

The function f describing the measurement model can be re-written for individual values y of the quantity Y in a Taylor-expanded form:

$$y = f(x_1, x_2, ..., x_N) = y_0 + \sum_{n=1}^{N} \frac{\partial y}{\partial x_n} x_n + \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \frac{\partial^2 y}{\partial x_n \partial x_m} x_n x_m + \frac{1}{3!} \sum_{n=1}^{N} \sum_{m=1}^{N} \sum_{l=1}^{N} \frac{\partial^3 y}{\partial x_n \partial x_m \partial x_l} x_n x_m x_l + ... \quad \textbf{(2.2)}$$

In the case of small-disturbance approximation such as the estimation of measurement and retrieval uncertainty, the non-linearity of the function Y is generally considered small enough so that the terms of order 2 and higher in the Taylor expansion can be neglected. This will be our assumption in the rest of this work, which leads to the typical first order expression of the measurement model $Y = f(X_1, X_2, ..., X_N,)$:

$$y = f(x_1, x_2, ..., x_N) = y_0 + \sum_{n=1}^{N} \frac{\partial y}{\partial x_n} x_n \quad \textbf{(2.3)}$$

In these conditions, the standard uncertainty of the output quantity, namely the "combined standard uncertainty" u_y, is the estimated standard deviation associated with the result, and is equal to the positive square root of the combined variance obtained from all variance and covariance components using the "law of propagation of uncertainty" (art. 5.2 of the GUM (JCGM 100, 2008)):

$$u_y^2 = \sum_{n=1}^{N} \sum_{m=1}^{N} \frac{\partial y}{\partial x_n} \frac{\partial y}{\partial x_m} \mathrm{cov}(x_n, x_m) = \sum_{n=1}^{N} \left(\frac{\partial y}{\partial x_n} \right)^2 u_n^2 + 2 \sum_{m=1}^{N-1} \sum_{n=m+1}^{N} \frac{\partial y}{\partial x_n} \frac{\partial y}{\partial x_m} \mathrm{cov}(x_n, x_m) \quad \textbf{(2.4)}$$

The covariance between two random variables X_1 and X_2 with estimates x_1 and x_2 is defined by:

$$\mathrm{cov}(x_1, x_2) = \overline{(x_1 - \overline{x_1})(x_2 - \overline{x_2})} \quad \textbf{(2.5)}$$

The horizontal bar symbolizing the mean:

$$\overline{x} = \frac{1}{N_P} \sum_{i_P=1}^{N_P} x(i_P)$$

(2.6)

Eq. (2.4) can also be expressed in terms of correlation coefficient r_{nm} instead of covariance:

$$u_y^2 = \sum_{n=1}^{N}\left(\frac{\partial y}{\partial x_n}\right)^2 u_n^2 + \sum_{m=1}^{N}\left(\sum_{n=1(n\neq m)}^{N}\frac{\partial y}{\partial x_n}\frac{\partial y}{\partial x_m} r_{nm} u_n u_m\right)$$

(2.7)

with the correlation coefficients r_{nm} defined as:

$$r_{nm} = \frac{\mathrm{cov}(x_n, x_m)}{u_n u_m}$$

(2.8)

Correlations between input quantities cannot be ignored if present and significant. The associated covariances should be evaluated experimentally if feasible by varying the correlated input quantities (Type A evaluation of covariance), or by using the pool of available information on the correlated variability of the quantities in question (Type B evaluation of covariance). As stated in JCGM 104 (2009), "*the use of the GUM uncertainty framework becomes difficult when forming partial derivatives for a measurement model that is complicated, as needed by the law of propagation of uncertainty. A valid and sometimes more readily applicable treatment is obtained by applying a suitable Monte Carlo implementation of the propagation of distributions*". Monte Carlo experiments results such as those presented in **appendix A-C** for the ozone and temperature lidar models provide critical insight in the quantitative estimation of these correlations.

Combined standard uncertainty is a reference uncertainty value that can be used as is or to determine the so-called "expanded uncertainty" U, which defines "*an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand*" (art. 2.3.5 of JCGM 100 (2008)). Expanded uncertainty is expressed as a multiple of the standard uncertainty, the scaling factor being the "coverage factor" k so that $U=ku$. The value of k that produces an interval corresponding to a specified level of confidence requires detailed knowledge of the probability distribution characterized by the measurement result and its combined standard uncertainty. In measurement situations where the probability distribution characterized by a measurement and its uncertainty u is approximately normal, and the effective degrees of freedom of u is of significant size (typically greater than 10), taking k = 2 produces an interval having an approximate level of confidence p = 95.5%. Similarly, taking k = 3 produces an interval having an approximate level of confidence p = 99.7%. Correspondence between several key values of k and p for the normal and rectangular probability distributions is reported in **Table 2.1**.

Table 2.1 Correspondence between key values of coverage factor and level of confidence for two common probability distributions

Level of Confidence p (%)		*Coverage factor k*
Rectangular distribution	*Normal distribution*	
57.74	68.27	1
	90	1.645
95		1.65

	95	1.96
	95.45	2
	99	2.576
99		1.71
	99.73	3

Because of its essential meaning, the recommended definition and use of uncertainty for all NDACC lidars is the standard uncertainty, i.e., $k = 1$ (also commonly expressed as "1σ uncertainty"), which for a normal probability distribution, corresponds to an interval of confidence of approximately 68%.

2.3 Systematic and random effects, correlated and uncorrelated variables

The terms "systematic uncertainties" and "systematic errors", widely used in the literature, are mathematically too ambiguous to easily be assimilated in the analytical expressions described in the GUM (JCGM 100, 2008) for the propagation of uncertainty. This terminology should be avoided and will therefore not be used in this work unless it explicitly refers to the terminology used in specific cited works.

"Systematic", as used in those cited works, refers to components known to be present consistently in multiple samples of the same sampling population and owing to one or several well-identified systematic effects. For this reason a significant degree of correlation between measured samples is implied. It is only after reported systematic effects have been characterized by a randomized uncertainty component for each sample, and by a well-defined correlation matrix within the sampling population, that they can contribute to the analytical implementation of the combined uncertainty budget. The term "randomize" here consists of computing the value of an uncertainty component arising from a systematic effect using a probability distribution obtained from a Type-B evaluation. If an uncertainty component arising from a systematic effect cannot be randomized or if the covariance matrix within the sampling population cannot be computed, then this systematic effect cannot be accounted for in the uncertainty budget and it must be removed before measurement is made. If a systematic effect is reported as a non-zero (positive or negative) bias with the assumption that the value of this bias is known, then the measured samples must be corrected for this value before a combined uncertainty can be computed, and an uncertainty component characterizing the correction procedure must be introduced in the combined uncertainty budget. In order to preserve the full independence of a measurement, corrections for systematic effects must rely on the physical processes altering the measurement, and must apply to the input quantities X_i rather than the output quantity Y.

Often, when the term "systematic" is used in the literature, the term "statistical" or the term "random" is also used, this time to describe components that apply to one measured sample at a time, and that are independent of each other within the same sampling population. These components are typically taken as the experimental standard deviation of a given sample and treated as uncorrelated components.

Appendix G illustrates the complexity of propagating uncertainty components arising from systematic effects, but also provides a practical solution to the propagation process. It presents two simple examples of the different impact of systematic effects on combined uncertainty. In the first example, neglecting uncertainty components arising from systematic effects, or treating

the input variables as if they were uncorrelated results in the underestimation of the combined uncertainty. In the second example, treating the input variables as if they were uncorrelated results in the overestimation of the combined uncertainty. Unfortunately measurement models are rarely as simple as they are in the two examples of **appendix G**, and a thorough and possibly tedious determination of the covariance matrix or correlation coefficients should be expected.

Fortunately, there are existing numerical methods that yield an estimate of covariance terms. The Monte Carlo experiments introduced in **appendix A** show that combining the approach described in **appendix G** with the strict application of propagation of variance (**Eq. (2.4))** allows for a standardized and practical estimation of combined uncertainty in the presence of correlated variables. This approach prescribes the replacement of a single, complex measurement model by the successive application of multiple, simpler measurement sub-models. At each sub-model level, the standard uncertainty is evaluated in parallel for every independent uncertainty source introduced at the current or at a previous sub-model level. A key aspect of this approach is to ensure that the input quantities for which a correlation exists and is well-known are introduced together in the same sub-model, so that their covariance can be easily estimated. As explained in **appendix A**, this estimation can be done by creating a normally-distributed population of samples with a mean that equals the estimated (or most likely) value of the input quantity, and a standard deviation equaling its estimated standard uncertainty. We can then propagate in a similar manner each uncorrelated component in parallel until the final sub-model is reached, and combine all uncorrelated components together to obtain the ozone combined standard uncertainty and temperature combined standard uncertainty.

3 Practical treatment of uncertainty for the NDACC lidars

In this chapter we define a standardized approach agreed to by the ISSI Team for the practical treatment of measurement and retrieval uncertainty of the NDACC ozone and temperature lidars, taking into account the reality of the complex measurement models introduced in **chapter 1** and the strict application of variance propagation rules described in **chapter 2**. We provide generic, physically and metrologically sound recommendations and guidelines for the introduction of uncertainty into the ozone and temperature data processing chains, and its propagation until a combined standard uncertainty in ozone or temperature is derived. The generic nature of the present recommendations allows the proposed standardized treatment to be used not only by NDACC lidars, but also lidars contributing to other networks such as TOLNet (Tropospheric Ozone Lidar Network), GRUAN (GCOS Reference Upper Air Network), and other networks with similar instrumentation and/or scientific objectives.

3.1 Proposed approach for the sources of uncertainty

As explained in **chapter 2**, before the contribution of multiple uncertainty sources can be propagated and combined, it is critical to introduce sources adequately into the data processing chain. Proper knowledge of the input quantities, i.e., knowledge of their expected values and standard uncertainties, is necessary. Furthermore, the choice of measurement model or sub-model (as defined in **chapter 2**), is equally important to the choice of the ancillary datasets used a priori for the input quantities. All these aspects were considered in order to produce the set of recommendations and guidelines presented below.

3.1.1 On the choice of input quantities

The input quantities of potential significance to the ozone and temperature measurement models were first identified. From **chapter 1**, it is clear that the parameters impacting the retrievals relate to the propagation and backscattering of the laser light emitted by the lidar, and therefore include a number of atmospheric species and their scattering and/or absorption properties. Furthermore, several signal correction procedures and numerical transformations related to the instrumentation were identified and therefore must also be considered.

Figure 3.1-Figure 3.3 illustrate the impact on ozone and temperature of ignoring a number of standard lidar signal corrections, or the impact of changing the values of some input quantities used in these corrections. The curves show the magnitude of the difference between ozone (or temperature) retrieved without applying a specific correction, and ozone (or temperature) retrieved with all corrections applied ("control run"). This magnitude was computed for multiple intensity (i.e., altitude) ranges, when available. For the stratospheric ozone lidar for example, two separate red curves and two separate blue curves can be observed, each one corresponding to a specific altitude range. Though the correction procedures of the various NDACC ozone and temperature lidars can be different from an instrument to another, the order of magnitude of the observed differences is typically the same and **Figure 3.1-Figure 3.3** are representative of this order of magnitude. For tropospheric and stratospheric ozone, only the corrections and/or input quantities having an impact greater than 0.1% have been included. For temperature, the corrections and/or input quantities having an impact approaching 0.1 K have been included. Real measurements from the JPL stratospheric ozone and temperature lidar at Mauna Loa

Observatory, Hawaii, and the Univ. of Alabama RO3QET tropospheric ozone lidar in Huntsville, Alabama have been used (M. Newchurch and S. Kuang, personal communication). At 355 nm (a detection wavelength of the JPL lidar), the effect of ozone absorption on temperature is negligible (orange curves). At other wavelengths, this effect is not negligible, and for completeness, we also plotted this effect as if the detection wavelength was 532 nm (pink curves). We also included the effect at 532 nm of a change in ozone cross-section dataset (namely, using the so-called "Serdyuchenko dataset" instead of the "DMB dataset", see details later in this report). For tropospheric ozone, the impact of a missing SO_2 absorption correction is estimated for highly polluted boundary layer conditions, with SO_2 concentrations of the order of the part-per-billion (worst case scenario). Because of either low concentrations or low cross-section values, no other atmospheric species is known to interfere with the ozone DIAL and temperature lidar measurement techniques.

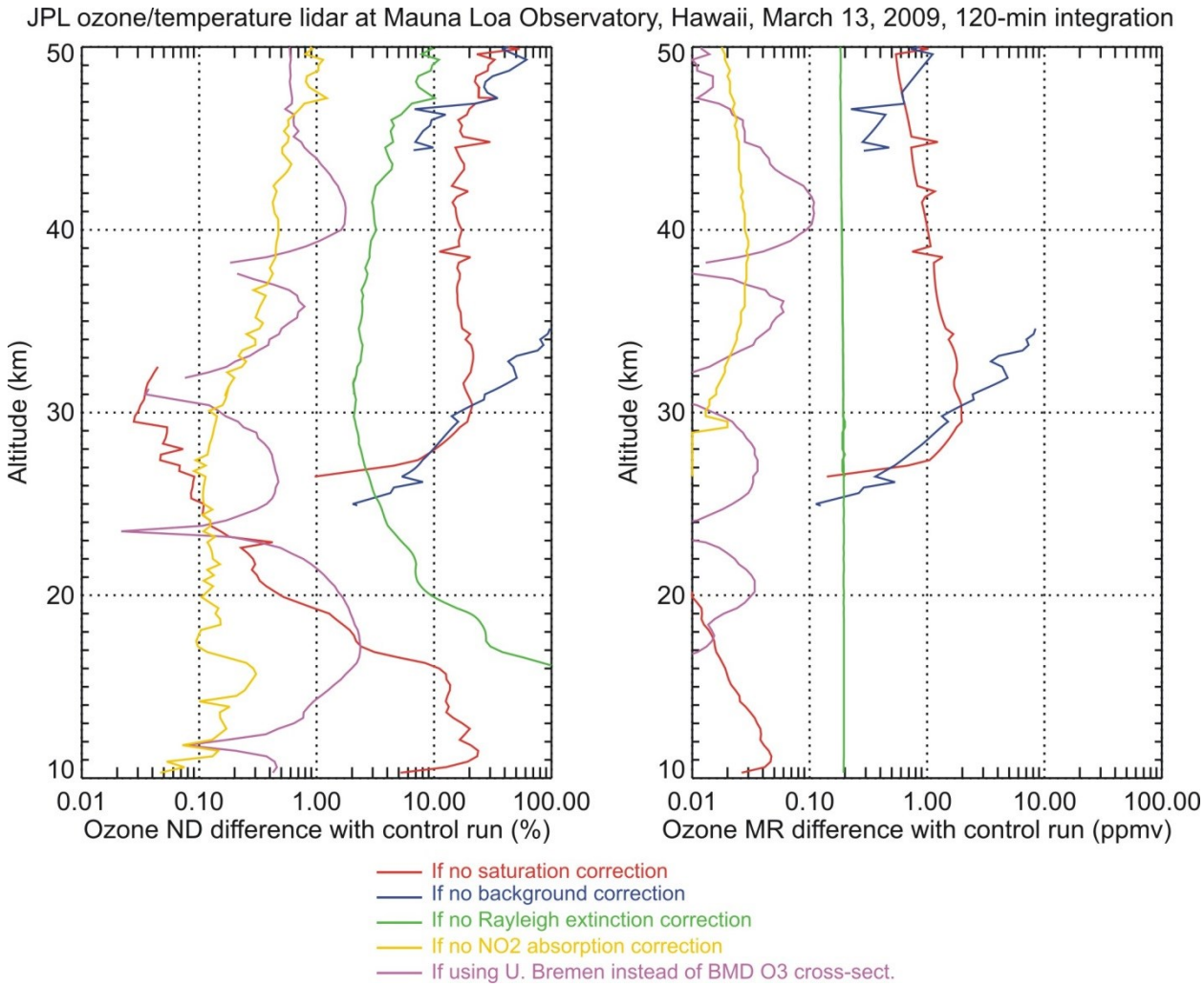


Figure 3.1 Stratospheric ozone error resulting from the removal of various signal corrections from the normal DIAL retrieval, or from changing the values of some input quantities used in these corrections. Each curve shows the magnitude of the difference between ozone retrieved without a specific correction and ozone retrieved with all corrections applied ("control run"). The actual differences can either be negative or positive, depending on the correction/dataset considered. An actual measurement from the JPL stratospheric ozone and temperature lidar at Mauna Loa Observatory, Hawaii was used to plot these differences.

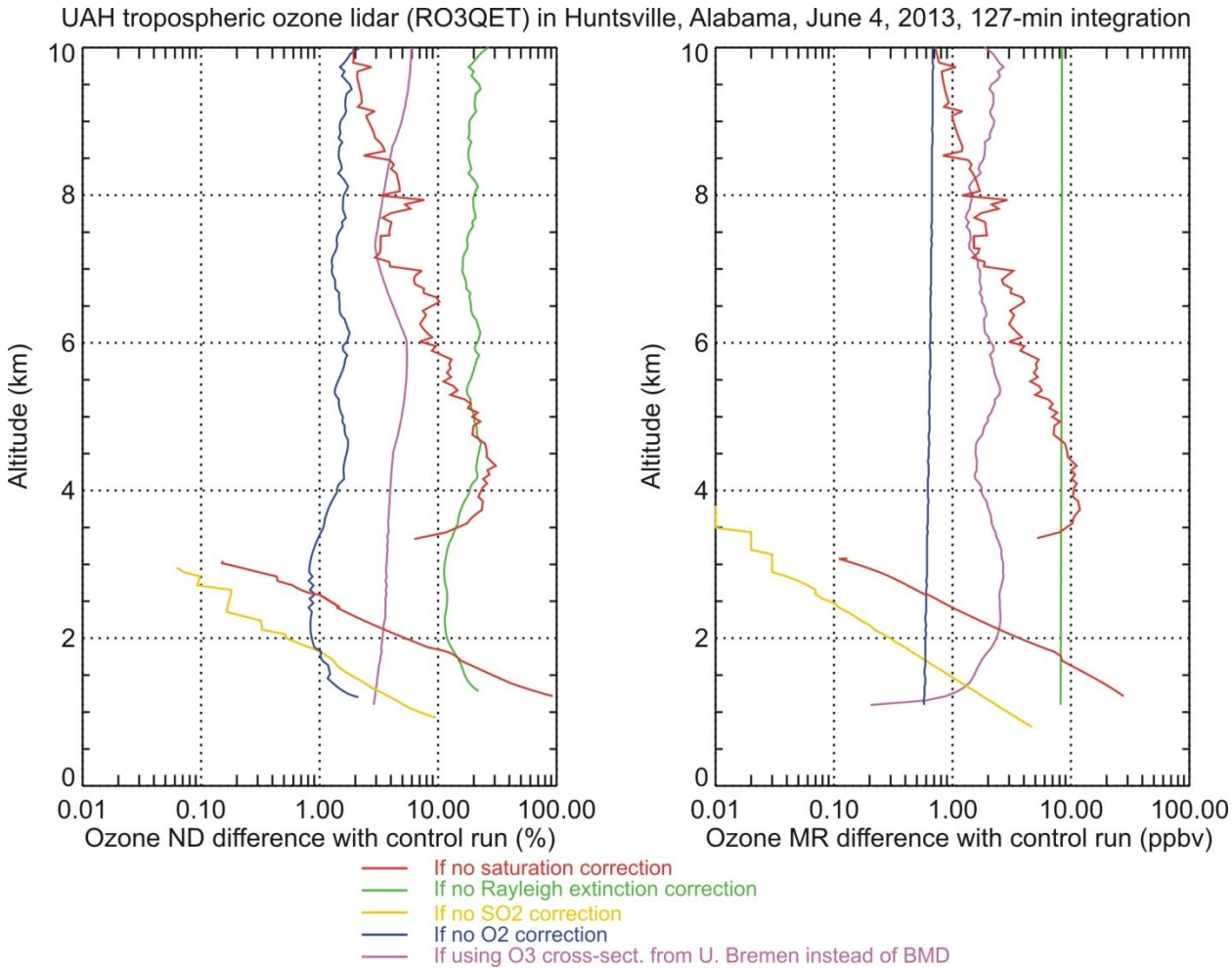


Figure 3.2 Same as Figure 3.1 but for tropospheric ozone. An actual measurement from the Univ. of Alabama tropospheric ozone lidar in Huntsville, Alabama (RO3QET lidar) was used to plot these differences

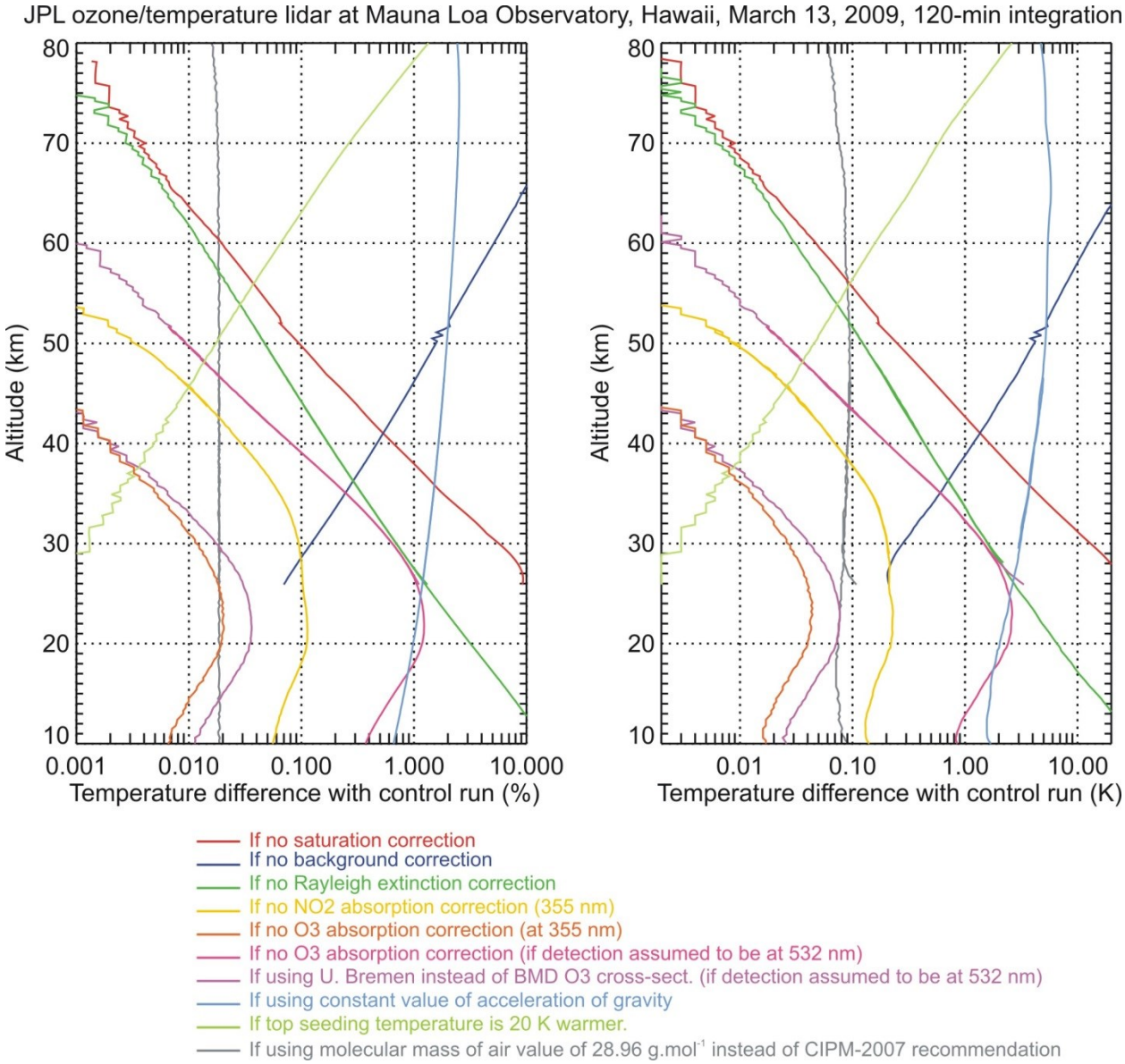


Figure 3.3 Same as Figure 3.1 but for the temperature. An actual measurement from the JPL stratospheric ozone and temperature lidar at Mauna Loa Observatory, Hawaii was used to plot these differences

The effect of particulate extinction and backscatter is not included here, and will not be treated in this report. Their contribution is negligible in a cloud-free, "clean" atmosphere, which is mostly true above 35 km (e.g., Godin-Beekmann, et al., 2003), and in most cases of tropospheric ozone DIAL clear-sky measurements for which the wavelength differential is small. When present and non-negligible, their contribution is highly variable from site to site, time to time, and highly dependent on the nature and quantity of the particulate matter at the time of measurement. Providing meaningful recommendations for a standardized treatment of these two sources of uncertainty is therefore beyond the scope of this report. However, the ISSI-team highly recommends the formation of a new Working Group (or ISSI Team) whose specific objective would be to provide recommendations for a standardized treatment of uncertainty owed to

660 particulate extinction and backscatter, similarly to what has been done in the present report for
661 the other uncertainty sources.

662 The following input quantities will be considered in this report, and should be taken into account
663 the NDACC-lidar standardized ozone uncertainty budget:

664 1) Saturation (pile-up) correction parameters (e.g., photon-counters' dead-time)
665 2) Background extraction parameters (typically, fitting parameters)
666 3) Ozone absorption cross-sections
667 4) Rayleigh extinction cross-sections
668 5) Ancillary air number density profile (or temperature and pressure profile)
669 6) NO_2 absorption cross-sections
670 7) Ancillary NO_2 number density profile (or mixing ratio profile)
671 8) SO_2 absorption cross-sections (UV only)
672 9) Ancillary SO_2 number density profile (or mixing ratio profile)
673 10) O_2 absorption cross-sections (at shorter UV wavelengths)

674 Not all of the above input quantities should necessarily be retained, depending on the instrument
675 and retrieval configuration considered. NO_2 and SO_2 absorption is typically negligible in the
676 stratospheric ozone retrieval (0.1%-1% ozone error or less if neglected), as well as most cases of
677 tropospheric ozone retrieval. However it is included in this report to account for the potentially
678 non-negligible effect of a heavily-polluted boundary layer, or potentially heavy volcanic aerosols
679 loading conditions (Godin-Beekmann, et al., 2003). The absorption by O_2 should be considered
680 only if any of the detection wavelengths is shorter than 294 nm as they relate to absorption by O_2
681 in the Herzberg bands. Note that the O_2 number density is directly proportional to air number
682 density (constant mixing ratio), and therefore no additional input quantity is needed for O_2
683 number density or mixing ratio.

684 Similarly, the following input quantities will be considered in this report and should be taken into
685 account in the NDACC-lidar standardized temperature uncertainty budget:

686 1) Saturation (pile-up) correction parameters (e.g., photon-counters' dead-time)
687 2) Background extraction parameters (typically, fitting parameters)
688 3) Rayleigh extinction cross-sections
689 4) Ancillary air number density (or temperature and pressure)
690 5) Ozone absorption cross-sections (for detection wavelengths in the Chappuis band)
691 6) Ancillary ozone number density or mixing ratio (detection in the Chappuis band)
692 7) NO_2 absorption cross-sections (for detection wavelengths in the Higgins band)
693 8) Ancillary NO_2 number density or mixing ratio (detection in the Higgins band)
694 9) The acceleration of gravity
695 10) The molecular mass of air

Just like for ozone, not all of the above input quantities should necessarily be retained, depending on the instrument and retrieval configuration considered. The impact of absorption by O_3 on the temperature retrieval is very small (<0.1 K) if working at wavelengths near the ozone minimum absorption region (e.g., 355 nm, 387 nm), but can account for up to 1 K error if neglected when working in the Chappuis band (e.g., 532 nm and 607 nm). Conversely, absorption by NO_2 is very small for temperature retrievals in the Chappuis band, but can account for up to a 0.2 K error if neglected at 355 nm and 387 nm. Absorption by SO_2 is negligible at all wavelengths relevant to temperature retrievals (typically longer than 340 nm).

The contributions of the acceleration of gravity to the temperature retrieval is very small (<0.1 K) providing the gravity model is altitude-dependent. In the upper mesosphere, the change in the air major species' mixing ratio induces a change with altitude of the air molecular mass and Rayleigh scattering cross-sections. However the induced changes remain below 0.1 K below 90 km, which is much less than the expected uncertainty owed to the other sources such as detection noise and tie-on temperature uncertainty (Argall, 2007). For temperature profiles seeded above 100 km (e.g., Sica et al., 2008), the change of the molecular mass of air with altitude must be taken into account.

3.1.2 On the choice of measurement models and sub-models

For both the ozone (respectively temperature) lidar measurement techniques (respectively DIAL and traditional density integration method), complex measurement models have been identified, and were briefly reviewed in **chapter 1**. The complete ozone and temperature lidar data processing chain can be divided in three processing stages: 1) signal processing common to both the ozone and temperature retrievals, 2) processing specific to ozone retrieval exclusively, and 3) processing specific to temperature retrieval exclusively. Furthermore, each stage must be split into multiple, simpler measurement sub-models where independent uncertainty components can be propagated step-by-step and in parallel until the final stage of processing.

The sub-models are described in **chapter 4** (signal processing), **chapter 5** (ozone retrieval) and **chapter 6** (temperature retrieval). The sub-models' input quantities X_i, are itemized in **these chapters**, and should be included and accounted for in the NDACC-standardized lidar measurement and retrieval uncertainty budget for ozone and temperature. These quantities are characterized by their input values x_i with standard uncertainty estimates u_i, For brevity, the values and their uncertainty estimates will often be referred to in this report as a pair (x_i,u_i). Special care was taken in this report to ensure that each input quantity introduced at any one sub-model level is independent from all other input quantities introduced elsewhere, so that the associated uncertainty component can be propagated in parallel with all the others before it is combined.

3.1.3 A standardized approach for the introduction of the input quantities

For each identified input quantity X_i known to have a non-zero uncertainty, the ISSI-Team provides one or more datasets that can be used a priori for the NDACC-standardized lidar uncertainty budgets. Each pair comprises the mean (or most likely) value x_R and its corresponding standard uncertainty (or standard deviation) u_R estimated either from a Type A or a Type B evaluation. These ancillary datasets, as of 2014, are reviewed in **section 3.5** below. The list is non-exhaustive and should be updated as needed. It is expected that the NDACC Lidar

738 Working Group regularly revisits the list of available ancillary datasets, and reviews and
739 proposes new datasets, as deemed necessary.

740 For each input quantity presented thereafter, NDACC lidar investigators can either choose to use
741 their "in-house" ancillary dataset having the value x_{PI} with the uncertainty u_{PI}, or one of the
742 ancillary datasets provided by the ISSI Team with input value x_R and uncertainty u_R. The
743 NDACC lidar investigators' choice of ancillary dataset should be consistent with the approach
744 described in one of the following four scenarios:

745 a. If an investigator chooses to use the dataset provided by the ISSI Team, then the
746 contribution of X_i to the measurement sub-model must be formed by the following
747 pair: $(x_i,u_i) = (x_R,u_R)$

748 b. If both x_{PI} and u_{PI} are available and the investigator chooses to use them, then the
749 contribution of X_i to the measurement sub-model must be formed by the following
750 pair: $(x_i,u_i) = (x_{PI},u_{PI})$

751 c. If x_{PI} is available but with no knowledge of u_{PI}, and the investigator chooses to
752 use it instead of the pair (x_R,u_R), recommended by the ISSI Team then the
753 contribution of X_i to the measurement sub-model must be formed by the following
754 pair: $(x_i,u_i) = (x_{PI},\max(u_R,|x_{PI}-x_R|))$

755 d. If no input quantity is used (i.e., if X_i is ignored in the measurement sub-model),
756 then the investigator must use 100% of the dataset value recommended by the
757 ISSI Team as the input quantity's uncertainty, i.e., the contribution of X_i to the
758 measurement sub-model must be formed by the following pair: $(x_i,u_i) = (0,|x_R|)$.

759 **Figure 3.4** illustrates in a more intuitive manner the four scenarios described above.

760

761

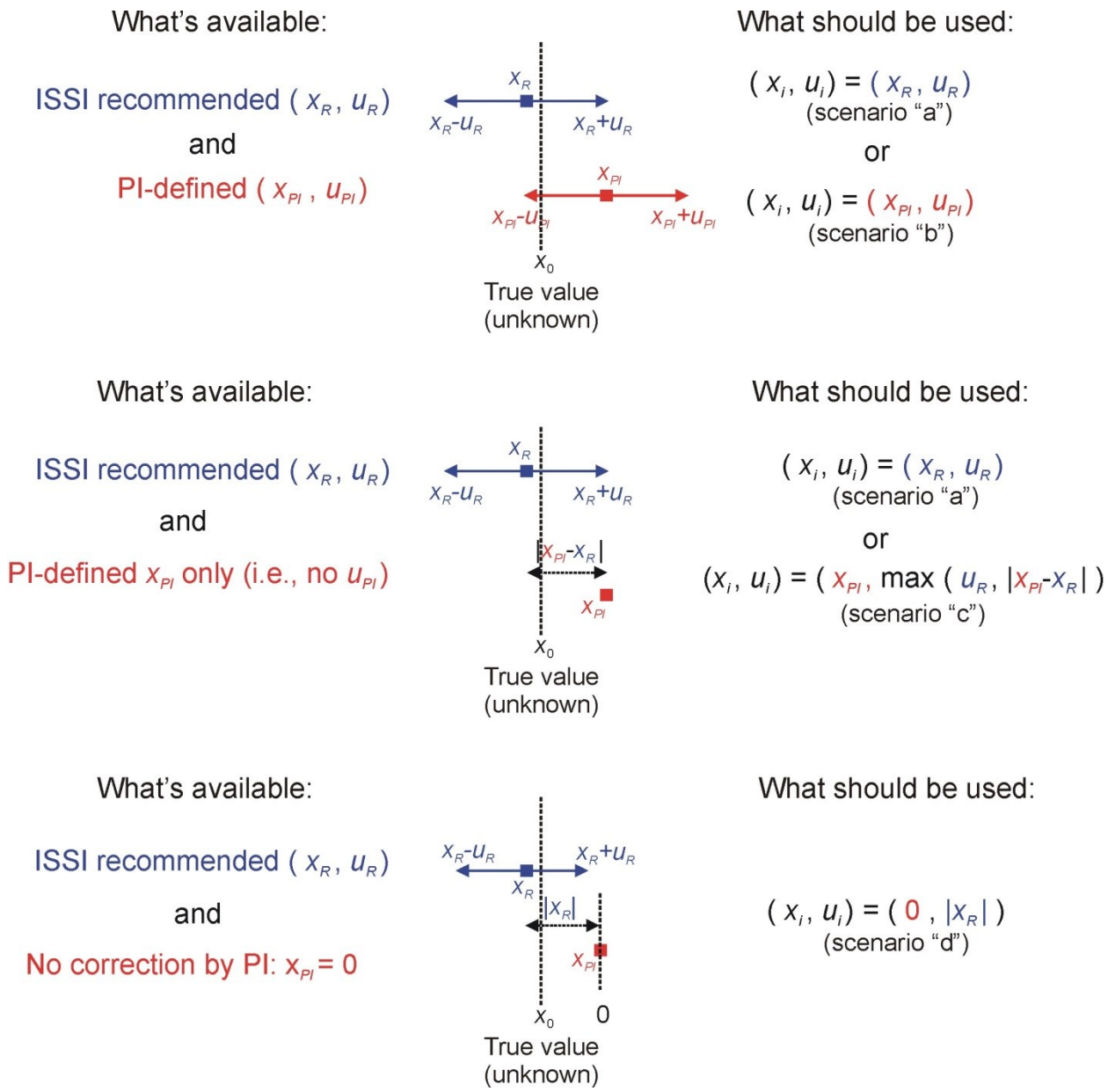


Figure 3.4: Schematics of the different scenarios available to the NDACC lidar investigators using the ISSI-Team proposed standardized approach for the introduction of the input quantities

3.1.4 Handling fundamental physical constants

In order to limit the complexity of the standardization process, the contribution of uncertainty associated with the fundamental physical constants is treated differently from that of the other sources. Similarly to the definition of uncertainty, we refer here to an internationally recognized and traceable standard for our recommendations on the use of physical constants, namely the International Council for Science (ICSU) Committee on Data for Science and Technology (CODATA, http://www.codata.org/), endorsed by the BIPM (Mohr et al., 2008). Within the

CODATA, the Task Group on Fundamental Constants (TGFC) "*periodically provides the scientific and technological communities with a self-consistent set of internationally recommended values of the basic constants and conversion factors of physics and chemistry based on all of the relevant data available at a given point in time*". A comprehensive set of physical constants and their uncertainty is available on the TGFC website: http://physics.nist.gov/cuu/Constants/index.html.

Our approach ensures that there is indeed no propagation of uncertainty for fundamental physical constants. To do so, we truncate the CODATA-reported values to the decimal level where uncertainty no longer affects rounding. For example, the molar gas constant value reported by the CODATA is 8.3144621 $Jmol^{-1}K^{-1}$ with an uncertainty of 0.0000075 $Jmol^{-1}K^{-1}$. If we truncate to the value of 8.3145 $Jmol^{-1}K^{-1}$, adding or subtracting its uncertainty does not modify the truncated value, and we therefore consider this value as "exact" (i.e., no uncertainty to be propagated). This approach is valid because the CODATA-reported uncertainty of the constants involved in the ozone and temperature lidar retrievals is always several orders of magnitude smaller than the value of the constant itself. This approach should not be followed when the uncertainty is of the same order of magnitude as the value of the constant. In such cases, the constant must be treated similarly to any other input quantities introduced in the measurement model, i.e., its uncertainty should be taken into account and propagated.

Examples of the CODATA-reported values of fundamental constants relevant to NDACC ozone and temperature lidar community is provided in **Table 3.1** together with their uncertainties. The table shows the current CODATA values, as well as the NDACC-standardized values following the methodology just described. It also shows values used within the NDACC community as of 2012 (i.e., before standardization efforts were undergone). Note that not all the constants listed this table are necessarily used in the ozone and/or temperature data processing chains.

Table 3.1 Examples of the range of values of the relevant physical constants used by a number of NDACC lidar investigators together with the CODATA recommended values

Constant [unit]	***Lowest value used in NDACC-lidar (2012)***	***Highest value used in NDACC-lidar (2012)***	***CODATA recommended value (2010)***	***CODATA reported uncertainty (2010)***	***ISSI Team recommended value (2014)***
Molar ideal gas constant [$Jmol^{-1}K^{-1}$]	8.31	8.314472	8.3144621	0.0000075	8.3145
Boltzmann constant [JK^{-1}]	1.3806E-23	1.380662E-23	1.3806488E-23	0.0000013E-23	1.38065E-23
Avogadro number [mol^{-1}]	6.02214179E23	6.0225E23	6.02214129 E23	0.00000027E23	6.02214E23
Atomic mass unit [kg]	1.66053E-27	1.66057E-27	1.660538 921E-27	0.000000073E-27	1.660539E-27
Speed of light in vacuum [ms^{-1}]	299790000	300000000	299792458	0	299792458

3.2 Proposed approach for the propagation of uncertainty

The practical implementation of **Eq. (1.4)** for the retrieval of ozone number density and **Eqs. (1.11)-(1.12)** for the retrieval of temperature requires that the raw lidar data be numerically transformed and corrected. The complete ozone and temperature lidar data processing chain can be divided in three processing stages: 1) signal processing common to both the ozone and temperature retrievals, 2) processing specific to ozone retrieval exclusively, and 3) processing specific to temperature retrieval exclusively. Because of its resulting complexity, the ozone measurement model must be split into multiple sub-models and independent uncertainty components must be propagated step-by-step and in parallel through these sub-models until the final stage of processing. The full data processing chain is summarized in the form of sample flowcharts that can be found in **chapters 3-5**..

In the remainder of this chapter we provide an overview of the sources of uncertainty that should be considered and introduced in the ozone and temperature lidar models. We also specify which sources of uncertainty will not be treated in this work, and the reasons for not treating them.

3.2.1 Signal processing

The data processing chain starts with the transformation of the lidar signals readout in the raw data files into number of backscattered photons actually reaching the lidar detector (which eventually relates to ozone and temperature through **Eqs. (1.4), (1.11) and (1.12)**). This transformation includes some, or all of the following: the conversion of the signals detected in photon-counting (PC) mode or analog-to-digital (AD) conversion mode, the PC signal correction for saturation (pile-up), the extraction and removal of background noise, the correction for incomplete overlap between the telescope field-of-view and the laser beam, the merging of multiple data streams into one stream covering all altitudes of interest, and the reduction of random noise by low-pass vertical filtering (smoothing). The corresponding independent uncertainty components to be considered and propagated in parallel are:

- Uncertainty due to detection noise for both PC and AD channels
- Uncertainty due to saturation (pile-up) correction for photon-counting (PC) channels
- Uncertainty due to background noise extraction (for both PC and AD channels)
- Uncertainty due to channel merging procedure (typically, PC and AD, or PC and PC)

All the above components except detection noise imply correlated terms in the vertical dimension, which means that covariance terms must be taken into account when vertical filtering is applied. In addition, if the same counting hardware is shared by two channels, covariance terms must be taken into account if dependent channels/ranges are combined (e.g., signal merging). These uncertainty components are introduced in **chapter 4**, and the expressions of their propagation are detailed in **chapters 4 and 5** for ozone, and in **chapters 4 and 6** for temperature.

Uncertainty due to analog-to-digital signal conversion (AD channels only) will not be estimated in details in this report. Its estimation is highly instrument-dependent, and there is therefore no practical standardized way to provide meaningful recommendations at the moment. For reference, the contribution of digitizing the signal from an analog channel is briefly described in **section 4.1.2**. Uncertainty due to incomplete overlap is also highly instrument-dependent, and

often time-dependent for the same instrument. One possible treatment is proposed in **appendix J** keeping in mind that the approach described is not unique.

3.2.2 Ozone DIAL retrieval

For the ozone DIAL retrieval, the data processing continues with the transformation of the corrected signals into ozone number density. The steps towards the final ozone profile include: the computation of the logarithm of the ratio of the "ON" and "OFF" channels, the differentiation with respect to altitude, optional smoothing, and correction for atmospheric extinction. From **Eq. (1.4)**, the independent uncertainty components that should be considered and propagated "in parallel" are:

- Uncertainty due to the a priori use of ozone cross-sections
- Uncertainty due to the a priori use of Rayleigh cross-sections
- Uncertainty due to the a priori use of ancillary air number density (or temperature and pressure)
- Uncertainty due to the a priori use of NO_2 absorption cross-sections
- Uncertainty due to the a priori use of ancillary NO_2 number density (or mixing ratio)
- Uncertainty due to the a priori use of SO_2 absorption cross-sections
- Uncertainty due to the a priori use of ancillary SO_2 number density (or mixing ratio)
- Uncertainty due to the a priori use of O_2 absorption cross-sections

The term "a priori" here does not mean that the ozone DIAL retrieval uses a variational/optimal estimation method (it does not), but simply means that the information comes from ancillary (i.e., non-lidar) measurements, and is input as "truth" in the data processing chain for use in the various lidar signal corrections needed. The above components are introduced in **chapter 5**, and the expressions for their propagation are detailed in that chapter.

When filtering or differentiating with respect to altitude, the covariance terms must be taken into account for all uncertainty components introduced earlier and that are known to be correlated in altitude (typically, saturation correction, background noise extraction, overlap correction). Furthermore, when taking the ratio of the logarithm of the signals of the "ON" and "OFF" channels, covariance should be taken into account when propagating the saturation correction uncertainty if these channels share the same counting hardware. This potentially applies to the background extraction uncertainty as well. Finally, when computing the ozone cross-section differentials and the interfering gases' cross-section differentials, the covariance terms should be taken into account if the same ancillary datasets are used for the "ON" and "OFF" wavelengths.

3.2.3 Temperature retrieval (density integration technique)

After the raw signal is processed as described in the above **section 3.2.1**, the data processing continues with range correction, low-pass vertical filtering (smoothing), correction for atmospheric extinction, the vertical integration of relative density, and the computation of temperature downward from the seeded value at the top of the profile. Based on **Eqs. (1.11)-(1.12)**, the independent uncertainty components that should be considered and propagated "in parallel" are:

- Uncertainty due to the a priori use of Rayleigh cross-section
- Uncertainty due to the a priori use of ancillary air number density
- Uncertainty due to the a priori use of O_3 absorption cross-sections
- Uncertainty due to the a priori use of ancillary O_3 number density (or mixing ratio)
- Uncertainty due to the a priori use of NO_2 absorption cross-sections
- Uncertainty due to the a priori use of ancillary NO_2 number density (or mixing ratio)
- Uncertainty due to the a priori use of ancillary temperature (or pressure) for tie-on at the top
- Uncertainty due to the a priori use of acceleration of gravity
- Uncertainty due to the a priori use of molecular mass of air

Once again, the term "a priori" here simply means that the information comes from ancillary (i.e., non-lidar) measurements, and is input as "truth" in the data processing chain for use in the signal corrections.

When vertically filtering, the covariance terms at multiple altitude bins should be taken into account when propagating the saturation and background correction uncertainties. When integrating the product of the signal by the acceleration of gravity, and when computing the temperature downward from the top, the covariance terms should be taken into account for all uncertainty components except uncertainty owed to detection noise. The uncertainty sources listed above are introduced in **chapter 6**, and the expressions of their propagation are detailed in that chapter.

3.3 Proposed approach for the reporting of uncertainty in the NDACC data files

Every source of uncertainty X_i identified in this report, whether or not accounted for in the NDACC investigator's uncertainty budget, must be reported in the NDACC-archived metadata file. Providing quantitative information on the ancillary datasets used (which typically could be the value of the pair (x_i,u_i)) is highly recommended. Whether or not using the NDACC-standardized uncertainty budget approach, the best estimate of the ozone (or temperature) combined standard uncertainty must be reported in the NDACC-archived data files. In addition, individual standard uncertainty components that contribute to the ozone (or temperature) combined uncertainty should be reported in the NDACC-archived data files whenever possible.

Typically, NDACC ozone and temperature lidar profiles are given as a function of altitude and for an averaging time period ranging between a few minutes and several hours. For each reported uncertainty component, the systematic or random nature of the effects it is associated with should be reported in both the altitude and time dimensions. When using multiple NDACC-archived ozone or temperature lidar profiles, for example to produce an ozone or temperature climatology, each reported uncertainty component must first be computed separately based on the expected systematic or random behavior of the process associated with it, and then be combined.

3.3.1 Uncertainty associated with processes of random nature

Detection noise is the only uncertainty source that should be considered having a random behavior both in the altitude and time dimensions. When averaging multiple NDACC-archived ozone (or temperature) profiles in time, the values of uncertainty due to detection noise for the time-averaged profile should be equal to the square-root of the quadratic sum of the detection noise uncertainty values reported for each profile used in the time average. When applying vertical smoothing to a given NDACC-archived ozone (or temperature) profile, the values of uncertainty due to detection noise for the vertically-smoothed profile should be equal to the square-root of the weighted quadratic sum of the uncertainty values reported at each altitude point used in the original profile.

The above time and vertical averaging rules should not be used for the combined ozone and combined temperature uncertainty unless it is proven that uncertainty due to detection noise is much larger than any other uncertainty components.

3.3.2 Uncertainty associated with processes of systematic nature

The following uncertainty components are associated with processes of systematic nature in both the altitude and time dimensions as long as no change in instrumentation, ancillary dataset, or signal correction algorithm occurs:

- Saturation (pile-up) correction
- Background noise extraction
- Overlap correction
- Channel (or range) vertical merging
- Rayleigh cross-sections
- Ancillary air number density
- Ozone absorption cross-sections
- Ancillary ozone number density or mixing ratio profile (temperature retrieval only)
- NO_2 absorption cross-sections
- Ancillary NO_2 number density profile
- SO_2 absorption cross-sections (ozone retrieval only)
- Ancillary SO_2 number density profile (ozone retrieval only)
- O_2 absorption cross-sections (wavelengths <294 nm only)
- Temperature tie-on (temperature retrieval only)
- Acceleration of gravity (temperature retrieval only)
- Molecular mass of air (temperature retrieval only)

When averaging multiple NDACC-archived ozone or temperature profiles in time, the time-averaged values of the corresponding uncertainty components should equal the time average of the uncertainty values reported for each profile. This averaging procedure is valid only over the

time period over which no change in instrumentation, ancillary dataset, or correction algorithm occurs. When applying vertical smoothing to a given NDACC-archived ozone or temperature profile, the values of the corresponding uncertainty components for the vertically-smoothed profile should be computed using the weighted average of the values reported at each altitude point in the original profile.

The time and vertical averaging rules just described should not be used for the combined ozone and combined temperature uncertainty unless it is proved that the uncertainty values owed to detection noise propagated to ozone or temperature are much smaller than the uncertainty values owed to the sources listed above.

3.4 Handling uncertainty for sources not identified or described in this report

Because each lidar instrument is unique, some sources of uncertainty are not necessarily identified or present in the ozone or temperature lidar measurement sub-models described in **chapters 4-6**. For these unidentified sources, as well as uncertainty owed to analog detection, overlap correction, and particulate backscatter and extinction corrections mentioned earlier, the NDACC lidar investigators should use the same generic approach as that used for the identified sources, and simply add those components to the uncertainty budget following the same definitions, methodologies, and propagation principles. As already mentioned, it is very desirable that dedicated working groups be formed in the near future to address the standardization of the treatment of uncertainty for these remaining components.

3.5 Ancillary datasets suitable for use in the NDACC lidar data processing algorithms

Here we provide a non-exhaustive list of ancillary datasets that can be used in the NDACC lidar data processing algorithms. These datasets were selected after compromising between their quality, availability, and ease-of-use. For the Rayleigh extinction and ozone absorption cross-sections, prioritization was made based on the results presented in **appendix D** and **appendix E** respectively. Often, more recent datasets have been given higher priority unless no uncertainty or standard deviation information is provided with the dataset, or unless part of the dataset was found unreliable.

3.5.1 Rayleigh cross-sections

Appendix D provides a brief review of the various equations used to express the Rayleigh scattering cross-section or Rayleigh scattering coefficient as a function of wavelength. Here, we suggest using the expression given by Eberhard (2010), who provides an excellent review of the approximations made by other authors leading to alternate expressions of the Rayleigh scattering cross-section or coefficient. The Rayleigh cross-section for the "mean" air (e.g., Eberhard, 2010) is given by:

$$\sigma_M(\lambda) = \frac{24\pi^3}{\lambda^4} L^2(\lambda) \sum_i \nu_i \frac{1}{N_{iS}^2} \left(\frac{n_{iS}^2(\lambda) - 1}{n_{iS}^2(\lambda) + 2} \right)^2 F_i(\lambda) \quad \textbf{(3.1)}$$

The process by which each term is introduced in **Eq. (3.1)** is detailed in **appendix D** which summarizes the work of Eberhard (2010). **Eq. (3.1)** differs from many other expressions found in literature due to the presence of the Lorentz factor L, and to the joint summation of the polarizability (expressed here as refractive index at standard temperature and pressure conditions n_{iS}) and the King factor F_i for each air constituent (i=N_2, O_2, Ar, CO_2, and H_2O) of mixing ratio ν_i and number density at standard temperature and pressure conditions N_{iS}.

The dominant source of uncertainty in the determination of Rayleigh cross-section is the King factor as revealed by the results presented in **appendix D**. **Figure 3.5** (exert of **appendix D**) shows the differences found between the Rayleigh scattering cross-sections formulation of Eberhard (2010) and other authors referred to in **appendix D**. The cross-section differences are plotted in the left panel (applies to lidar temperature retrieval). The impact on the cross-section differential (for the ozone DIAL retrieval) is shown in the right panel. If we exclude two "outliers" (Hoyt, 1976; Fröhlich and Shaw, 1980), all Rayleigh cross-section formulations reviewed in **appendix D** remain within +/-1% of that of Eberhard (2010) throughout the 200-670 nm spectral window of interest for temperature and ozone lidar retrievals.

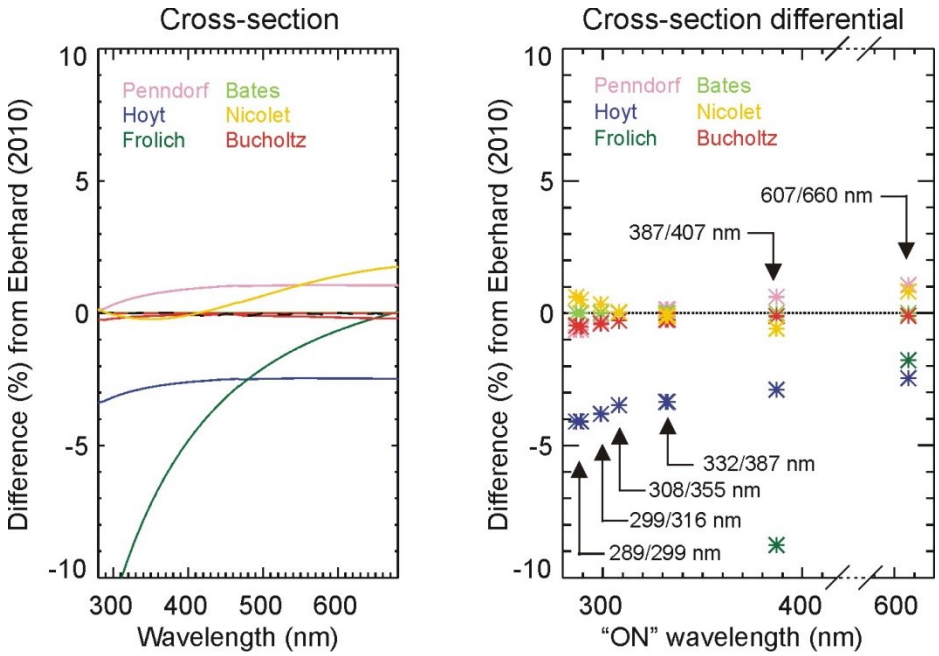


Figure 3.5 Relative differences (%) of the Rayleigh scattering cross-section (left) and cross-section differential (right) computed for the various equations found in literature and with respect to Eberhard (2010) (see appendix D for details)

After compromising between the strict application of **Eq. (3.1)** and the various approximations found in literature (Type-B evaluation), a good conservative estimate of uncertainty for the Rayleigh cross-section $u_{\sigma M}$ is 0.5-1% if using the formulations of Eberhard (2010) Bates (1984) and Bucholtz (1995), and 1% if using the formulations of Penndorf (1957) and Nicolet (1984). The formulations of Fröhlich and Shaw (1980) and Hoyt (1976) are not recommended. A relative uncertainty of 0.5 % holds in the visible region, but down to 200 nm the incorporation of theoretical values typically increases this value to 1%. As will be seen in the next paragraph, the

contribution of the ancillary air number density u_{Na} to the overall molecular extinction uncertainty is typically two to five times larger than that of the Rayleigh cross-section. Uncertainty estimate of five Rayleigh extinction cross-section sources reviewed by the ISSI-Team are compiled in **Table 3.2** below.

Table 3.2*: Rayleigh extinction cross-section datasets suitable for use in the NDACC lidar algorithms*

Dataset	Domain of validity	Uncertainty estimates	Knowledge base	ISSI Team recommendation for input uncertainty
Eberhard, 2010	200-700 nm	0-0.3 %	Comparisons	Use conservative 0.5% (UV-visible) and 1% (UV Hartley band)
Bates, 1984	200-700 nm	0-0.3 %	Comparisons	Use conservative 0.5% (UV-visible) and 1% (UV Hartley band)
Bucholtz, 1984	200-700 nm	0-0.5 %	Comparisons	Use conservative 0.5% (UV-visible) and 1% (UV Hartley band)
Nicolet, 1984	200-550 nm	0.5-1.0%	Comparisons	Use conservative 1%
Penndorf, 1957	200-700 nm	0-1.0%	Comparisons	Use conservative 1%

3.5.2 Air temperature and number density

An ancillary air temperature profile is needed in the ozone DIAL retrieval to compute the ozone absorption cross-sections as a function of height and wavelength, and is needed in the temperature retrieval to tie-on the measured profile at the top. The ancillary temperature profile can either be an actual measurement, or a profile obtained from an analysis, assimilation, forecast, climatological or empirical model. Temperature datasets can be found among a broad spectrum of measuring techniques and models. The associated uncertainty values vary widely and cannot be prescribed in a simple manner.

Below 30 km, most commonly used radiosondes have a temperature measurement uncertainty ranging from 0.2 K to 0.5 K (Hurst et al., 2012; Immler et al., 2010). Other techniques or models should be used in the ozone and temperature retrievals only if no nearby and near-simultaneous radiosonde profile is available. State-of-the-art operational, assimilation, or re-analysis models should be used as the next option as they yield the smallest temperature uncertainty after radiosonde measurements. Depending on the model and altitude range considered, this uncertainty is estimated to be between 2 K and 10 K. Whether using radiosondes or models, the air temperature and density uncertainties should take into account the uncertainty associated with pressure measurements, and when applicable, with the conversion between geometric altitude and geopotential height.

Between 30 km and 55 km, temperature uncertainty estimates from meteorological analysis such as NCEP range between 3 K and 9 K. If no radiosonde and no operational model is available, satellite measurements, either from infra-red sounders, such as SSU (Reale et al., 2008; Keckhut et al., 2011) and AMSU (Wang et al., 2014), or upper atmosphere research instruments such as Aura-MLS (Schwartz et al., 2008), OSIRIS (Sheese e tal., 2012), SOFIE (Stevens et al., 2012) and SABER (Remsberg et al., 2008) can be used. They usually yield uncertainties of the order of

1-10% (2-20 K). State-of-the-art assimilation or re-analysis models can also be used, with temperature uncertainties ranging from 10 to 20 K (Dee et al., 2011; Uppala et al., 2005; Kalnay et al., 1996).

Above 55 km (temperature retrieval only), upper atmosphere research instruments (e.g., Aura-MLS, SABER) are the best option. They typically yield uncertainties in the order of 1-10%, i.e., 2-15 K. Merged datasets such as GOZCARDS (Froidevaux et al., 2014, personal communication) can also be used in the stratosphere and mesosphere. In the absence of individual or merged datasets, empirical models such as CIRA-86, MSIS90 (Hedin, 1991), or MSISE00 (Picone et al., JGR, 2002) can be used with estimated uncertainties of about 20-30 K. Finally, the Whole Atmosphere Community Climate Model (WACCM) developed at NCAR (Garcia et al., 2007) has the advantage of covering four decades in time and the entire atmospheric profile sounded by the NDACC lidars (ground to 120 km).

If no standard uncertainty estimates are provided with a given dataset, standard deviations must be used following the examples provided in **appendix F**, or more sophisticated collocation uncertainty models must be used (Fassò et al., 2014; Sofieva et al., 2008). Coverage and uncertainty facts of a few ancillary temperature datasets are compiled in **Table 3.3**.

Ancillary air number density, in most cases, is not a measured quantity, but a quantity derived from measured (or modeled) temperature and pressure. The source datasets are therefore the same as those for ancillary temperature. Pressure p_a, temperature T_a and number density N_a are linked by the ideal gas law:

$$N_a(k) = \frac{p_a(k)}{k_B T_a(k)} \quad \textbf{(3.2)}$$

k_B is the Boltzmann constant. If the ancillary pressure and temperature profiles are fully correlated (i.e., correlation coefficient of 1), the ancillary air number density standard uncertainty should be written:

$$u_{Na} = N_a \left| \frac{u_{pa}}{p_a} - \frac{u_{Ta}}{T_a} \right| \quad \textbf{(3.3)}$$

If independent pressure and temperature profiles are used, the ancillary air number density standard uncertainty should be written:

$$u_{Na} = N_a \sqrt{\left(\frac{u_{pa}}{p_a}\right)^2 + \left(\frac{u_{Ta}}{T_a}\right)^2} \quad \textbf{(3.4)}$$

Among the datasets listed in **Table 3.3**, radiosonde is the only one that implies a non-ambiguous relationship between pressure and temperature. Radiosondes typically carry two different sensors, and though the pressure and temperature measurements are not completely independent, we can safely assume to use **Eq. (3.4)**. For all other datasets, a careful estimation of the pressure and temperature inter-dependence is needed before air number density uncertainty can be computed.

1090 **Table 3.3***: Ancillary air temperature and number density datasets suitable for use in the NDACC lidar algorithms*

Dataset	Domain of validity	Uncertainty estimates	Knowledge base	ISSI Team suggestion for uncertainty
Radiosonde	0-12 km	T_a: 0.2-0.5 K p_a: 0.3-0.5 hPa	Comparisons	Use conservative 0.5 K and 0.5 hPa
Radiosonde	12-30 km	T_a: 0.5-2.0 K p_a: 0.3-0.5 hPa	Comparisons	Use conservative 2 K and 0.5 hPa
NCEP	0-30 km	2-3 K	Data files	Use reported uncertainty
NCEP	30-50 km	3-9 K	Data files	Use reported uncertainty
WACCM	0-120 km	5-20 K	Comparisons	Use std.-dev.
GOZCARDS	300-0.001 hPa	5-10%	Data files	Use reported std.-dev.
SABER	100-0.001 hPa	5-20 K	Data files	Use reported uncertainty
CIRA, MSIS	0-30 km 30-100 km	10 K 20 K	Comparisons	Use conservative 10 K Use conservative 20 K

1091

1092 ### 3.5.3 Ozone absorption cross-sections

1093 The role of ozone absorption cross-sections is essential in the ozone DIAL retrieval (e.g., Godin-
1094 Beekmann and Nair, 1999). Any relative error in ozone absorption cross-section will directly
1095 translate into the same relative error in ozone number density. For the temperature retrieval,
1096 ozone absorption cross-sections have a smaller impact, but the impact is not negligible if using
1097 wavelengths in the Chappuis band (Sica et al., 2001). A detailed review of eight contemporary
1098 ozone cross-section datasets is provided in **appendix E**. The recommendations in the present
1099 chapter are based upon the recommendations given by the WMO Ad-hoc Working Group on
1100 Absorption Cross-sections of Ozone (ACSO) and the consensus decision made by the NDACC
1101 Lidar Working Group in November 2013, which was based on the results presented in **appendix**
1102 **E**. The differences between the cross-section differentials computed at selected "ON" and "OFF"
1103 wavelengths and for six of the eight datasets reviewed in **appendix E** are plotted in **Figure 3.6** as
1104 a function of temperature. Refer to **appendix E** for a complete description of how these plots
1105 were obtained. For DIAL pairs in the Hartley band, differences of up to 5% are found for the
1106 289/299 nm (tropospheric ozone) pair, while in the Huggins band, differences do not exceed 2-
1107 3%. **Figure 3.7** (exert from **appendix E**) shows the relative differences between the cross-
1108 sections of four datasets in the Chappuis band. Biases of 1-5% are apparent between the datasets.
1109 The differences observed at wavelengths longer than 675 nm are not discussed here because they
1110 do not impact the ozone DIAL and temperature retrievals.

1111

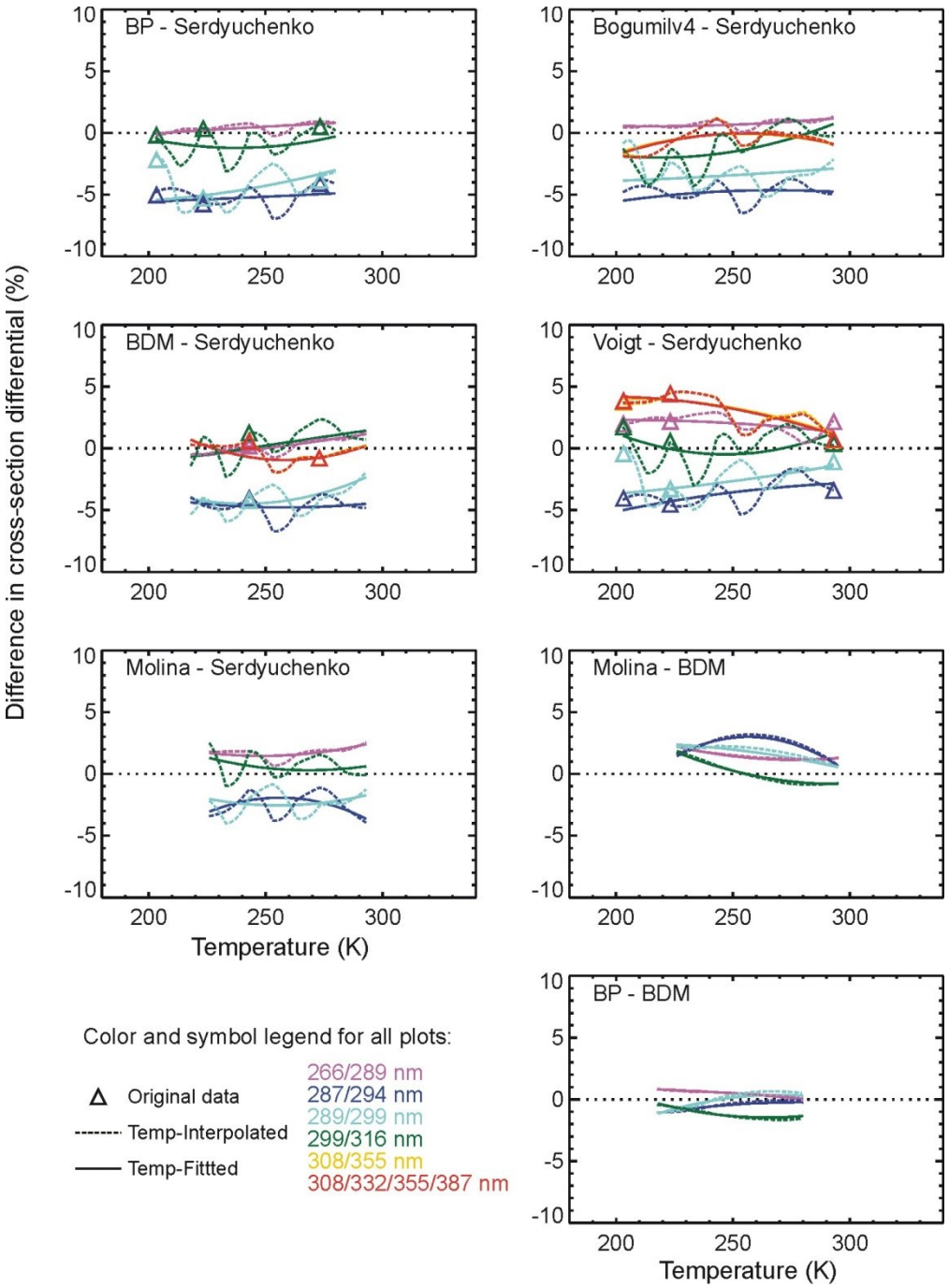


Figure 3.6 Relative differences (%) between the cross-section differentials at six DIAL wavelength pairs computed from six of the eight datasets reviewed in appendix E

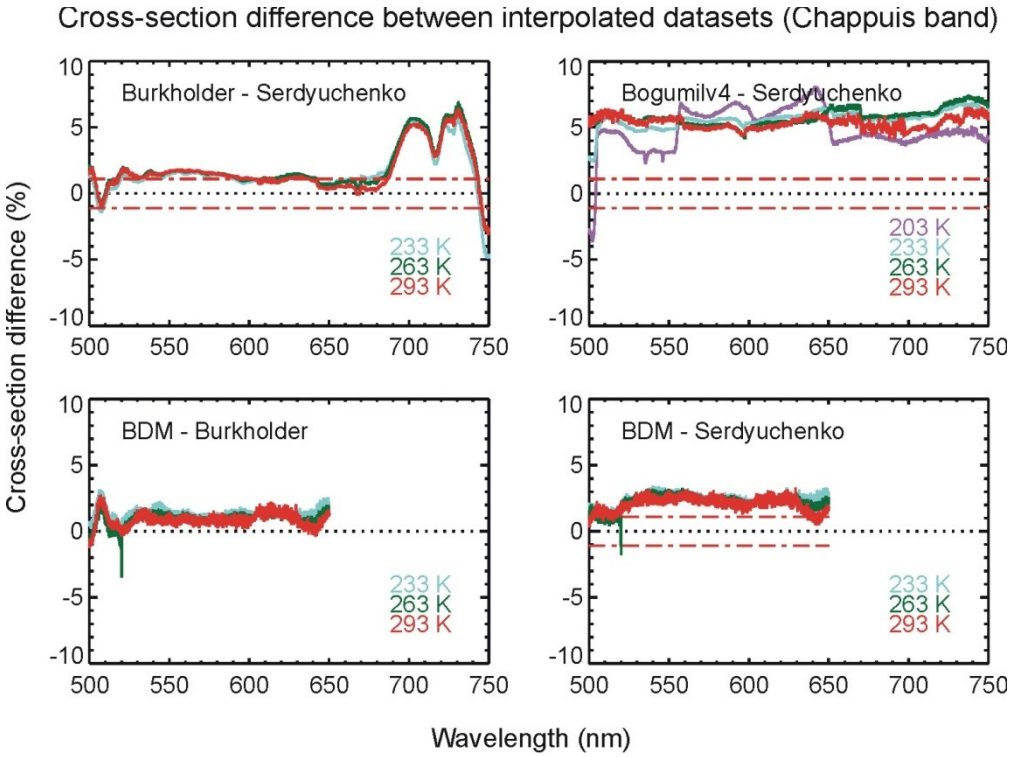


Figure 3.7 Cross-section relative differences (%) between selected datasets (Serdyuchenko, Bogumilv4, DMB and Burkholder) in the Chappuis band for selected temperatures, when available (see Appendix E for details)

Out of the eight datasets compared in **appendix E**, the datasets of DMB (Daumont et al., 1992; Malicet et al., 1995; Brion et al., 1998), and Serdyuchenko (Gorshelev et al., 2014) have the advantage of being tabulated at a 0.01-nm wavelength interval, of including cross-section measurements at multiple temperatures and over spectral regions covering all ozone and temperature lidar wavelengths (266-608 nm). They make therefore appropriate choices for use in the NDACC stratospheric ozone and temperature lidar retrievals. The Serdyuchenko cross-sections were measured at 11 temperatures between 193 and 293 K (Serdyuchenko et al., 2014), which represents the best available temperature coverage to date. However, the existing biases between this dataset and the DMB and BP (Bass and Paur, 1984) datasets for wavelengths relevant to tropospheric ozone DIAL must be kept in mind at all times. The Serdyuchenko data are provided as a function of wavelength in vacuum. A conversion to wavelength in air is necessary for comparison with DMB and BP. Finally, the historical BP dataset has not been used by the working groups of the most recent satellite instruments, and is not recommended by the ACSO Working Group. For the sake of consistency between tropospheric ozone DIAL, stratospheric ozone DIAL, and other measuring techniques (e.g., Dobson, satellite), the DMB dataset seems to provide the best compromise, and is ultimately the ISSI-Team preferred recommendation. For completeness, the characteristics of two other datasets are provided in **Table 3.4** below.

Table 3.4*: Ozone absorption cross-section datasets suitable for use in the NDACC lidar algorithms*

Dataset	Domain of validity	Uncertainty estimate	Knowledge base	ISSI Team recommendation for uncertainty
DMB, 1992-1998	200-310 nm 310-350 nm 350-410 nm 450-610 nm	1-1.5% 1.3-3.5% 5% 5%	Publications Comparisons	Use conservative 2% Use conservative 4% Use conservative 5% Use conservative 5%
Serdyuchenko, 2014	200-310 nm 310-350 nm 350-410 nm 450-610 nm	1.3-3% 1.7% 14-30% 5%	Publications Comparisons	Use conservative 3% Use conservative 2% Use conservative 20% Use conservative 5%
BP, 1984	245-330 nm 335-337.5 nm	1% 5%	Publications Comparisons	Use conservative 2% Use conservative 5%

3.5.4 Ozone number density

An ancillary ozone number density profile is needed in the temperature retrieval to account for ozone absorption at wavelengths in the Chappuis band. The ancillary ozone profile can either be an actual measurement, or a profile obtained from an analysis, assimilation, forecast, climatological or empirical model. Just like temperature, stratospheric ozone datasets can be found among a broad spectrum of measuring techniques and models. The associated uncertainty values vary widely and cannot be prescribed in a simple manner. Tropospheric ozone datasets (profiles) are rarer, and originate almost exclusively from ozone soundings.

Between the ground and 35 km, nearby, near-simultaneous electro-chemical-cell (ECC) ozonesondes should be used if available (Smit et al., 2007; Thompson et al., 2007). They yield a measurement uncertainty between 0.1 to 1 mPa (5-8%) (Smit et al., 2012; Stauffer et al., 2014). In the absence of nearby and simultaneous ozonesonde profile, publicly available climatologies can be used (e.g., Liu et al., 2013). Other datasets include the satellite measurements and/or climatologies of Aqua-Airs, Aura-TES in the troposphere, and Aura-MLS, SAGE, SAGE-II, UARS-MLS, and ENVISAT MIPAS, SCIAMACHY and GOMOS in the stratosphere (see Cracknell and Varotsos (2014) for a review of available datasets). State-of-the-art operational, assimilation, or re-analysis models (e.g., GEOS-5, ECMWF) can also be used. For ease of implementation, a climatology of the merged datasets from GOZCARDS (Froidevaux, 2014, personal communication), and a climatology of the model outputs from WACCM (Hannigan, 2014, personal communication) are suggested. In this case standard deviations are provided or easily computable with the climatology, and should be used for the estimation of uncertainty following the examples provided in **appendix F**. More sophisticated collocation uncertainty models Fassò et al. (2014) should be used if available. Coverage and uncertainty facts of a few ozone profile datasets are compiled in **Table 3.5** below.

1172 **Table 3.5*: Ancillary ozone profiles suitable for use in the NDACC lidar algorithms***

Dataset	Domain of validity	Uncertainty estimates	Knowledge base	ISSI Team recommendation for uncertainty
Ozonesonde	0-35 km	0.1-1 mPa (5-8%)	Publications Comparisons	Use conservative 8%
Env. Canada	0-30 km	5-25%	Publications Comparisons	Use std.-dev.
GOZCARDS	300-0.001 hPa	5-15%	Data files	Use std.-dev.
WACCM	0-100 km	5-25%	Comparisons	Use std.-dev.
Aqua-AIRS	1000-200 hPa	5-15%	Data files	Use reported uncertainty

1173

1174 ### 3.5.5 NO_2 and SO_2 absorption cross-sections

1175 In most cases, absorption by NO_2 and SO_2 has a minor impact on the ozone retrieval and a
1176 negligible impact on the temperature lidar retrieval. Only exceptional, highly-polluted conditions
1177 are susceptible to impact the ozone retrieval and its uncertainty budget, for example extreme-
1178 pollution peaks in urban areas or thick volcanic plumes. In these extreme cases, the air also
1179 contains a significant amount of particles (aerosols), which often proscribes meaningful ozone
1180 DIAL and temperature lidar measurements. The choice of NO_2 and SO_2 cross-sections is
1181 therefore not strongly constrained. Here we provide two well-known NO_2 and SO_2 cross-section
1182 datasets, one from the SCIAMACHY pre-flight model measurements of the University of
1183 Bremen (Bogumil et al., 2003), and the other from the Fourier Transform Spectrometer
1184 measurements of the Institut d'Aéronomie Spatiale de Belgique (Vandaele et al., 1998; 2009).
1185 Reported uncertainty estimates for these datasets are compiled in **Table 3.6** and **Table 3.7** for
1186 NO_2 and SO_2 respectively. These cross-sections are temperature dependent (Voigt et al., 2002;
1187 Orphal, 2003).

1188
1189 **Table 3.6 *NO_2 absorption cross-section datasets suitable for use in the NDACC lidar algorithms***

Dataset	Wavelength Range (UV)	Uncertainty estimates	Knowledge base	ISSI Team recommendation for uncertainty
Vandaele, 1998	238-250 nm 250-333 nm 333-900 nm	10% <3% 3%	Publications Comparisons	Use conservative 10% Use conservative 3% Use conservative 3%
Bogumil 2003	250-600 nm	3.5%	Publications Comparisons	Use conservative 4%

1190
1191 **Table 3.7 *SO_2 absorption cross-section datasets suitable for use in the NDACC lidar algorithms***

Dataset	Wavelength Range (UV)	Uncertainty estimates	Knowledge base	ISSI Team recommendation for uncertainty $u_{\sigma SO2}$
Vandaele 2009	227-345 nm	2%	Publications	Use conservative 5%
Bogumil 2003	295-320 nm	3%	Publications	Use conservative 5%

3.5.6 NO_2 and SO_2 number density profiles

Again the choice of NO_2 and SO_2 ancillary profiles does not need to be strongly constrained due to their minor impact on the ozone DIAL and temperature lidar retrievals. As of today, there is no single NO_2 or SO_2 measurement or climatology from the ground up to 100 km. A single profile can only be constructed using separate datasets.

Multi-year measurements of stratospheric NO_2 are available from several satellite instruments such as SCIAMACHY, HALOE, SAGE-II, ACE-FTS or OSIRIS (Bracher al., 2005; Butz et al., 2006; Dirksen et al., 2011; Dorf et al., 2006, Brohede et al., 2007, Jones et al., 2012, Rozanov et al., 2005; Sioris et al., 2007; Wetzel et al., 2007), or from ground-based DOAS or FTIR (Hendrick et al., 2003; Sussmann et al., 2005). In the troposphere (boundary layer), measurements or climatologies are much sparser (Boersma et al., 2009; Cao et al., 2006; Cede et al., 2006; Miyazaki et al., 2012a; 2012b; He et al., 2014; Meng et al., 2008; Volten et al., 2009). Reported uncertainty estimates for these datasets are compiled in **Table 3.8**.

For SO_2, measurements or climatologies are even sparser than for NO_2 (Brühl et al., 2013; He et al., 2014; Meng et al., 2008). Two datasets, namely MIPAS (Hopfner et al., 2013) and OMI (McLinden et al., 2014) can be used and have been listed for reference in **Table 3.9**.

Table 3.8 ***Ancillary NO_2 profiles suitable for use in the NDACC lidar algorithms***

Dataset	Domain of validity	Uncertainty estimates	Knowledge base	ISSI Team recommendation for uncertainty
WACCM	0-100 km	20%	Comparisons	Use std.-dev.
SCIAMACHY	15-30 km	15-20%	Publications	Use std.-dev. or conservative 20%
OSIRIS	15-25 km 25-35 km 35-40 km	22% 11-21% 11-31%	Publications	Use std.-dev. or conservative 25%
DIAL	Urban BL	6 ppb	Publications	Use conservative 10 ppb

Table 3.9 ***Ancillary SO_2 profiles suitable for use in the NDACC lidar algorithms***

Dataset	Domain of validity	Uncertainty estimates	Knowledge base	ISSI Team recommendation for uncertainty
MIPAS	15-45 km	5-20 ppt	Publications	Use std.-dev. or conservative 20 ppt
OMI	BL	30%	Publications	Use std.-dev. or conservative 30%

3.5.7 O_2 absorption cross-sections

Absorption by O_2 must be taken into account only for the tropospheric ozone retrievals using wavelengths shorter than 294 nm. The O_2 absorption features impacting the retrieval include the Herzberg continuum, the Herzberg bands and the Wulf bands. Because the Herzberg bands are made of very narrow lines, absorption by O_2 occurs only if the laser line width at the emission

wavelength overlaps with one or several of the Herzberg lines. Here we suggest using the O_2 cross-sections dataset obtained from the Fourier Transform Spectrometer measurements of the Institut d'Aéronomie Spatiale de Belgique (IASB) (Fally et al., 2000; Jenouvrier et al., 1999; Merienne et al., 2000), yielding rough uncertainty estimates of 10%.

In the ozone DIAL equation, the absorption terms by interfering species comprise not only the cross-section differential but also the number density of that species. The ancillary O_2 number density profile can be directly computed from the ancillary air number density profile using the following constant mixing ratio ((Picard et al., 2008), see **paragraph 3.5.9**):

$$N_{O2} = 0.209390 N_a \tag{3.5}$$

3.5.8 Acceleration of gravity

The acceleration of gravity is used in the temperature retrieval. Nowadays, the most accurate and complete gravity field models are 4-dimensional, i.e., a function of longitude, latitude, altitude, and time. Gravity and other geodetic data are now computed with high accuracy and very high horizontal resolution using spherical harmonics expansion techniques. High-accuracy models are needed for state-of-the-art geodetic applications such as GPS, tides and tectonics. A good review of models and methods to derive accurate gravity fields can be found in the NASA Technical Report on the Earth Gravitational Model EGM96 (Lemoine et al., 1998). The currently most recognized standard originates from the World Geodetic System 1984 (NIMA-WGS, 1984) with a 2008 upgraded version of the geoid (Pavlis et al., 2012) using the latest GRACE (Gravity Recovery and Climate Experiment) satellite measurements (Tapley et al., 2005). For high resolution models, the accuracy in the geoid determination is better than 1 mGal which is 10^{-4} ms^{-2} or approximately 0.00001%.

Time-invariant and zonal mean approximations are also commonly used. They typically introduce errors less than 10^{-3} ms^{-2} (100 mGal), mainly owed to the height difference between the local reference ellipsoid used in the approximation and the local geoid computed from the harmonics models of highest order (currently, maximum degree for EGM2008 is 2159). Several analytical formulations of altitude-latitude-dependent gravity field exist, typically producing r.m.s. values between them of the order of 10^{-6} ms^{-2} (0.1 mGal).

For the specific application of temperature lidar, any 2-dimensional approximation derived from a recognized, standard reference ellipsoid is sufficient (e.g., WGS 84, GRS 80, GRACE GGM03). The difference between the gravity field calculated from high resolution models (EGM2008) and that calculated using the WGS 84 reference ellipsoid is within +/-100 mGal which represents approximately +/-0.001%. The ISSI-Team-recommended two-dimensional expression of gravity as a function of latitude ϕ and height h is taken from WGS 84 (NIMA-WGS, 1984) and can be written:

$$g(\phi, h) = g(\phi, 0)\left(1 - \frac{2}{a_1}\left(1 + f + m - 2f\sin^2\phi\right)h + \frac{3}{a_1^2}h^2\right) \tag{3.6}$$

The acceleration of gravity at the surface of the reference ellipsoid is:

$$g(\phi,0) = g_E \frac{1 + k \sin^2 \phi}{\sqrt{1 - e^2 \sin^2 \phi}} \quad \textbf{(3.7)}$$

In **Eqs. (3.6) and (3.7)**, the following universal and derived constants are used:

$$m = \frac{\omega^2 a_1^2 a_2}{\mathrm{GM}} = 0.00344978650684$$

$$k = \frac{a_2 g_P}{\mathrm{a_1 g_E}} - 1 = 0.00193185265241$$

g_E = 9.7803253359 ms^{-2} (the normal gravity at the equator)

e = 8.1819190842622 10-2 (the ellipsoid's first eccentricity)

a_1 = 6378137.0 m (the ellipsoid's semi-major axis)

a_2 = 6356752.3142 m (the ellipsoid's semi-minor axis)

f = 1/298.257223563 (ellipsoidal flattening)

GM = $3986004.418.10^8$ (the product of Earth's gravitational constant and mass)

The height h is taken relative to the surface of the reference ellipsoid, not the local geoid (which is the closest approximation of mean sea level). In most cases, the height difference between the geoid and the reference ellipsoid is within 60 m, with exceptions in the northern Indian Ocean and Indonesia where it can exceed 100 m. The latitude ϕ used in **Eq. (3.6)** is not the geographic latitude, but the geodetic latitude, which is defined as the angular difference between the zenithal vector above the local ellipsoid surface and that at the equator.

The largest errors in the computation of gravity come from the assumptions of height-invariance or latitude-invariance of gravity. The variation of gravity with height and latitude cannot be ignored in the temperature lidar retrievals. **Figure 3.8** shows the magnitude of the error made when neglecting the height-dependence of gravity (top panel), the latitude dependence (middle panel), and when neglecting both, i.e., taking a constant value for g (bottom panel). The error introduced by assuming a latitude-invariant gravity reaches +/-0.3%. The error introduced by assuming a height-invariant gravity can reaches 1% at 30 km, 2% at 60 km, and 3% at 100 km. The impact of commonly-used approximations for the computation of the acceleration of gravity is summarized in **Table 3.10**.

If a high-resolution model traceable to a well-known international standard is used, the uncertainty in the gravity acceleration can be neglected in the temperature uncertainty budget. If a latitude-varying and altitude varying approximation model is used, and this model is traceable to a well-known standard (e.g., the WGS 84 ellipsoid), the uncertainty in the gravity acceleration can also be neglected. In any other case, i.e., if a latitude-invariant gravity model and/or a height-invariant gravity model is used, then the uncertainty in the gravity acceleration should be taken into account, and the values should be taken from **Table 3.10.**

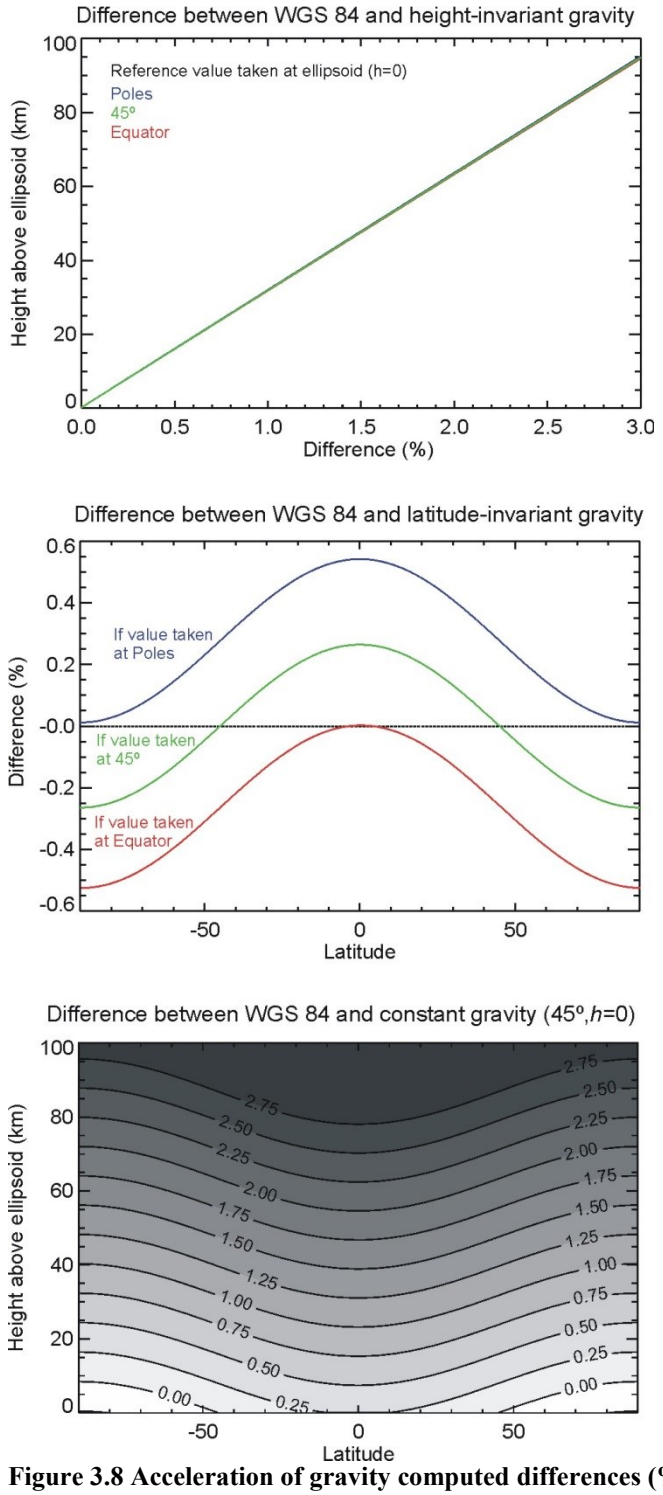


Figure 3.8 Acceleration of gravity computed differences (%) between the WGS 84 reference ellipsoid model and (top) a height-invariant model, (middle) a latitude-invariant model, and (bottom) a constant value (here taken at ϕ=45º and h=0, i.e., close to the standard acceleration of gravity g_0=9.806 ms^{-2}

1299 **Table 3.10 Commonly-used approximations on the acceleration of gravity and associated uncertainty**

Dataset	Estimated error	ISSI Team recommendation for uncertainty
WGS 84 (recommended)	<0.0002 ms^{-2} (from h)	Use conservative 0.0002 ms^{-2} (~0.002%)
Altitude-dependent but constant with latitude	0-0.6%	Use conservative 0.6%
Constant with altitude and latitude (not recommended)	0-3%	Use conservative 3%

1300

1301 ### 3.5.9 Molecular mass of air

1302 The molecular mass (or molecular weight) of air is needed in the temperature retrieval. Above
1303 about 10 km and below 80 km, the air can be considered well-mixed and the contribution of
1304 water vapor is negligible. An altitude-independent value for the mean molecular weight of dry
1305 air can therefore be computed. This value depends on the concentration in CO_2. In its
1306 computation, the Comité International des Poids et Mesures (CIPM-2007) assumes a CO_2 mixing
1307 ratio of 400 ppm and a value of Argon mixing ratio of 9334 ppm (Picard et al., 2008), leading to
1308 a standard value of the molecular mass of dry air of M_a = 0.02896546 kg. This value should be
1309 modified for different CO_2 mixing ratio values with the assumption that photosynthetic processes
1310 are dominant in the redistribution of the mole fractions of all constituents (which basically
1311 conserves the sum of the mixing ratios of O_2 and CO_2). This assumption leads to a modified
1312 expression of the mass of dry air as a function of CO_2 mixing ratio x_{CO2}:

1313 $$M_a(x_{CO2}) = 0.02896546 + 0.012011\left(x_{CO2} - 0.000400\right)$$

1314 **Table 3.11** shows the numerical values of the quantities involved in this calculation, and shows
1315 the values of the molecular mass of dry air corresponding to CO_2 mixing ratios measured in 1995
1316 (360 ppm) and 2014 (400 ppm). The relative difference between the 1995 and 2014 values is less
1317 than 0.002%, and the molecular mass of dry air can be used in this case without associated
1318 uncertainty. However, a rounded value of 0.02896 $kg.mol^{-1}$ is commonly used, and yields a
1319 difference with the CIPM-2007 of 0.02%. In this case, uncertainty must be propagated in the
1320 temperature retrieval.

1321

1322 **Table 3.11 Mass of dry air computed following CIPM-2007 and difference with other approximations**

Molecule	***Molecular mass*** (kg)	***Mixing ratio***	***Fractional contribution*** (kg)	**ISSI Team recommendation for uncertainty**
N_2	0.0280134	0.780848	0.0218742	/
O_2	0.0319988	0.209390	0.00670023	/
Ar	0.0399480	0.009332	0.000372795	/
CO_2 (400 ppm)	0.0440100	0.000400	0.0000176040	/
All others	/	/	$5.37156.10^{-7}$	/
Mass of dry air if 400 ppm of CO_2	/	/	**0.02896546**	Use as constant (no uncertainty)
Mass of dry air	/	/	**0.0289649**	Use as constant (no uncertainty)

if 360 ppm of CO_2				
Mass of dry air if rounded	/	/	**0.02896**	Use conservative 0.02%
At 90-110 km			**Variable with height**	Use conservative 0.5%

Above 80 km, the air is no longer well-mixed. The contribution of the variation of molecular mass of air with altitude remains very small (<1 K) compared to other contributions such as the temperature tie-on, signal detection noise, and background noise correction. However, the altitude-dependent, individual contributions of N_2, O_2 and O to the total air number density and to the Rayleigh backscatter cross-section should be taken into account if the lidar measurement extends beyond 100 km. Using for example the US Standard Atmosphere (1976), the resulting correction to the temperature can reach 1 K above 100 K (Argall, 2007).

4 Propagation of uncertainty common to both ozone and temperature retrievals

In the present chapter, we provide a detailed roadmap for the introduction and propagation of uncertainty through the initial data processing stage, i.e., the signal processing common to both ozone and temperature, starting from the lidar signals readout in the raw data files. The flowchart of **Figure 4.1** provides a convenient quick-look summary of this stage of data processing, and in particular shows the various locations at which new uncertainty terms are introduced and propagated. This flow-chart is not meant to prescribe or impose a specific data processing sequence or method, but rather to illustrate the generic approach of using successive sub-models through which independent uncertainty components are propagated.

For lidar channels operating in photon-counting (PC) mode, the data processing includes a correction for signal saturation as well as the removal of background noise. In most cases, the background noise is several orders of magnitude lower than the level of signal saturation, which allows saturation correction and background noise extraction to be permutated in the data processing sequence without altering the final results. In this chapter, we will address saturation correction first (**section 4.2**), and then background noise extraction (**section 4.3**).

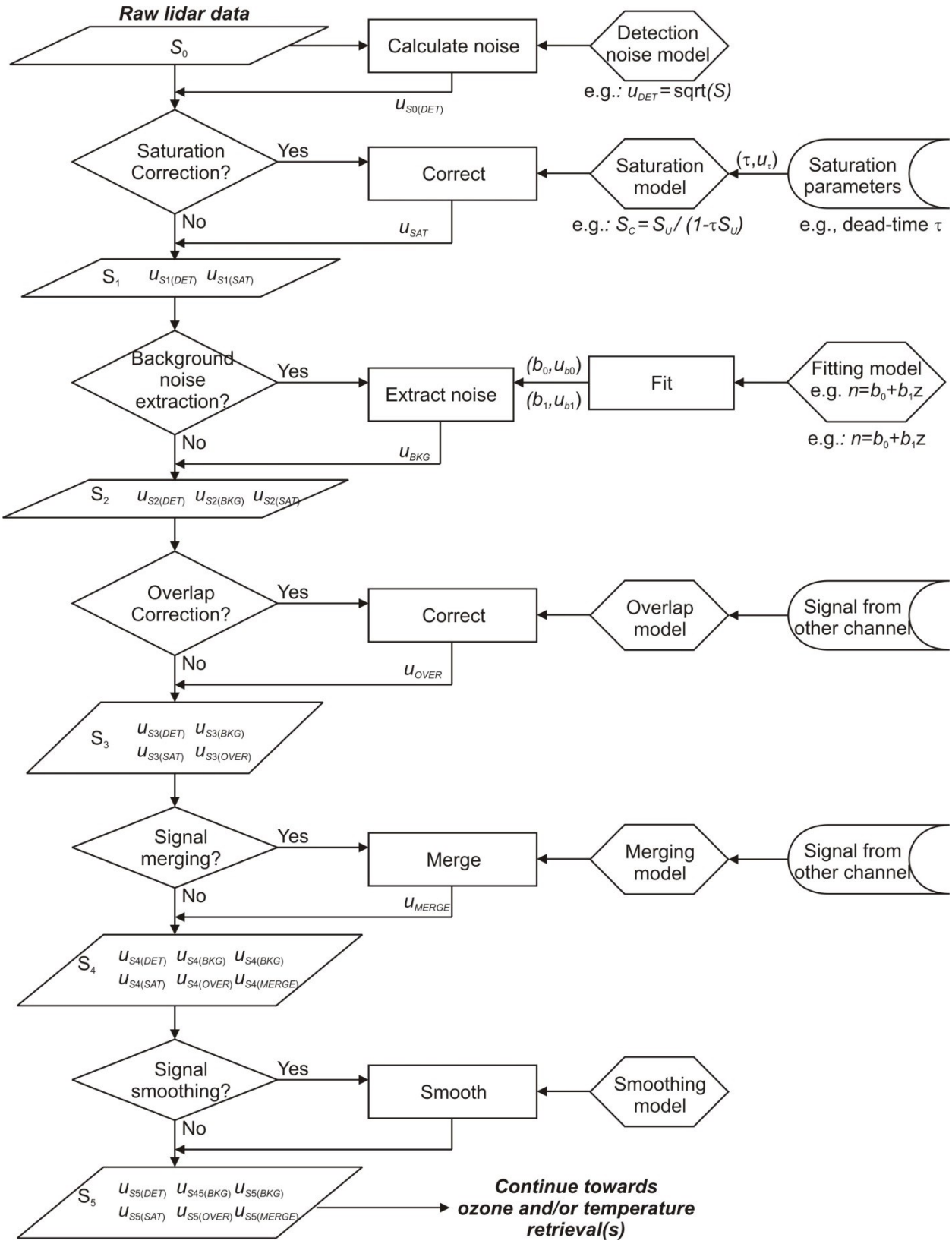


Figure 4.1 Flowchart representative of the lidar data processing common to both ozone and temperature retrievals

4.1 Raw signal detection noise

The variance (or noise) present in the signals recorded in the raw data files depends on the type of detection hardware, and on the type of signal recording hardware. Because this hardware can vary widely from an instrument to another, it is difficult to provide specific estimates of the uncertainty in the raw lidar signals recorded in the data files. However it is possible to provide guidance and recommendations for two well-known and distinct types of hardware: photon-counters (PC mode) and analog-to-digital converters (AD mode).

4.1.1 Photon Counting (PC mode)

When using a photon-counting data analyzer, the pulses output from the detector flow through a load resistor, a fast pulse amplifier and discriminator, and a high speed counter device. The variance of the noise in the signal S_0 recorded in the raw data files relates in first approximation to the theory of probability of detecting successive pulses given their amplitude, and is described by the common Poisson statistics. The signal standard deviation σ_0 therefore equals the square root of the mean expected value (Type A evaluation).

The single-shot standard uncertainty on signal s_0 recorded for channel i at altitude bin k is:

$$u_{s0(DET)}(i,k) = \sigma_0(i,k) = \sqrt{s_0(i,k)} \tag{4.1}$$

When summing the signals for L laser pulses, the summed signal S_0 is written:

$$S_0(i,k) = \sum_{l=1}^{L} s_0(i,k,l) \tag{4.2}$$

The standard uncertainty associated with the summed signal is:

$$u_{S0(DET)}(i,k) = \sqrt{S_0(i,k)} \tag{4.3}$$

The averaged (per pulse) signal $\overline{S_0}$ is written:

$$\overline{S_0}(i,k) = \frac{1}{L}\sum_{l=1}^{L} s_0(i,k,l) \tag{4.4}$$

The standard uncertainty associated with the averaged (per pulse) signal $\overline{S_0}$ is:

$$u_{\overline{S0}(DET)}(i,k) = \sqrt{\frac{\overline{S_0}(i,k)}{L}} \tag{4.5}$$

The three uncertainty terms defined here are often referred to as “precision” because they provide the best estimate of the repeatability of the lidar measurements. In the remainder of this

report, the summed signal S_0 for L laser shots will be considered and its standard uncertainty $u_{S0(DET)}$ will be propagated.

4.1.2 Analog-to-Digital Converter (AD)

In analog-to-digital conversion mode (AD), the signal output from the detector's backend is sent through an operational amplifier (OA) where it is converted into voltage, and then digitized. The bit-depth n of the digitizer typically ranges from 8-bit to 16-bit, depending on the manufacturer and hardware. The variance of the analog input signal v_0 is of similar form as that of the photon-counting case, but with additional dependencies on the detector gain m and the OA noise-equivalent bandwidth Δf and feedback resistance R_f:

$$u_{v0}(i,k) = \sigma_0(i,k) = \sqrt{v_0(i,k)} \tag{4.6}$$

With:

$$v_0(i,k) = 2m(i)eR_f^2(i)\Delta f(i)I_0(i,k) \tag{4.7}$$

I_0 is the intensity of the current (or voltage) output from the detector's backend, and e is the elementary charge of the electron. The analog signal is then digitized to form the readout signal S_0 for channel i at altitude bin k:

$$S_0(i,k) = v_0(i,k)\frac{2^{n(i)}}{V_{MAX}(i)} \tag{4.8}$$

V_{MAX} is the voltage range for channel i. This range is either pre-set (hardware specification), or set by the user at the time of data acquisition. In addition to the propagated statistical noise (Poisson) in the input analog signal, several sources of uncertainty specifically associated with the digitizing process must be taken into account:

- quantization, which is generally (but not always) taken as half the Least-Significant Bit (0.5 LSB) and which corresponds to the lowest increment in the analog voltage that can be discriminated by the digitizer (resolution), i.e., $V_{MAX}/2^n$ for a n-bit digitizer.

- differential non-linearity, a shift in the input voltage causing a bit change of the digital output at the incorrect analog value, and which is usually provided in units of LSB by the manufacturer

- integral non-linearity, the total shift in the input voltage value measured between the lowest and highest bit weights, also provided in units of LSB by the manufacturer

Other technical specifications provided by the manufacturer such as full-scale (or offset) error and gain error are alternate representations of the effects described above. All-in-all, typical, well-designed AD devices for lidar applications yield a total uncertainty in the digitizing process of 1 to 4 LSB. Because the analytical derivation of the total uncertainty in the digital readout varies from one hardware to another, the recommendation by the ISSI Team is to take the total uncertainty as the largest value between the detection noise (Poisson) and 4 LSB:

$$u_{S0(DET)}(i,k) = \max(\sqrt{S_0(i,k)}, 4)$$

(4.9)

4.2 Saturation correction (PC mode only)

During the counting process, the resolving time of the electronics (i.e., the dead-time τ) and the discriminator settings impact the linearity of the transfer between the signals at the detector's backend and the readout PC signals. The non-linearity effects are commonly referred to as "saturation" or "pile-up" effect. Well-known principles of probability of detection of two successive events associated with a Poisson process lead to analytical relationships of varying complexity between the true number of photons reaching the detector and the actual number of pulses counted. Practically, two types of detection models characterize photon-counting systems: models using non-extended dead-times ("non-paralyzable systems"), and models using extended dead-times ("paralyzable systems") (Muller, 1973; 1974).

4.2.1 Non-paralyzable systems

For non-paralyzable systems, the counting process "resets" at very large count rates at a time interval equal to the dead-time which causes the number of pulses counted to maximize (plateau) at the maximum counting rate of $P_{MAX}=1/\tau$. In this case, the relationship between the true count P_1 output from the detector and the actual counts P_0 transferred to the readout signals in single shot mode can be written:

$$P_1 = \frac{P_0}{1 - \tau P_0}$$

(4.10)

The uncorrected number of single-shot counts P_0 is linked to the uncorrected summed signal S_0 recorded in the raw data file for altitude bin k and channel i by the relation:

$$P_0 = \frac{2\delta z}{cL} S_0(i,k)$$

(4.11)

δz is the bin width, c is the speed of light and L is the number of laser shots considered for the summation of S_0. After saturation correction, the number of counts P_1 for altitude bin k and channel i is renormalized to obtain the corrected summed signal S_1 using:

$$S_1(i,k) = \frac{cL}{2\delta z} P_1$$

(4.12)

This leads to the actual saturation correction equation for summed signals:

$$S_1(i,k) = \frac{S_0(i,k)}{1 - \tau(i)\frac{2\delta z}{cL} S_0(i,k)}$$

(4.13)

There are two independent uncertainty components to propagate: uncertainty associated with the detection noise $u_{S0(DET)}$ and uncertainty associated with the dead-time u_τ. The value of the dead-

time and an estimate of its uncertainty are usually provided by the hardware's manufacturer as part of the system's technical specifications (Type A or Type B evaluation, depending on the manufacturer's evaluation procedure). It typically ranges from 2 ns to 10 ns, with a relative uncertainty estimate of a 2% to 10% percent (e.g., Keckhut et al., 1993; Leblanc et al., 1998; Welton and Campbell, 2002).

The uncertainty due to detection noise $u_{S0(DET)}$ propagated to the saturation-corrected signal S_1 is:

$$u_{S1(DET)}(i,k) = \left|\frac{\partial S_1(i,k)}{\partial S_0(i,k)}\right| u_{S0(DET)}(i,k) = \left(\frac{S_1(i,k)}{S_0(i,k)}\right)^2 u_{S0(DET)}(i,k) \quad \textbf{(4.14)}$$

The dead-time uncertainty u_τ propagated to the saturation-corrected signal S_1 is:

$$u_{S1(SAT)}(i,k) = \left|\frac{\partial S_1(i,k)}{\partial \tau(i)}\right| u_\tau(i) = \frac{2\delta z}{cL} S_1^2(i,k) u_\tau(i) \quad \textbf{(4.15)}$$

The combined uncertainty for the saturation-corrected signal S_1 is:

$$u_{S1}(i,k) = \sqrt{u_{S1(DET)}^2(i,k) + u_{S1(SAT)}^2(i,k)} \quad \textbf{(4.16)}$$

4.2.2 Paralyzable systems

For paralyzable systems, the counting process will not "reset" itself until two photon pulses are received within an interval equal to- or larger than the dead-time. This process causes the number of pulses counted to collapse to zero at very large count rates. In addition, the saturation process depends on the hardware's discriminator settings. Omitting for brevity the channel *i* and altitude range *k* dependencies, and using the formalism of Omote (1990) and Donovan et al. (1993), we can express the observed number of counts in single-shot mode P_0 as a function of the actual number of counts P_1 in single-shot mode:

$$P_0 = P_1 \exp\left(-\tau P_1\right)\left[\left(1-d\right) + \left(d - \frac{d^2}{2}\right)\tau P_1 + \sum_{j=2}^{\infty} \frac{1}{j!}\left(\frac{d^j}{j!} - \frac{d^{j+1}}{(j+1)!}\right)\left[\tau P_1\right]^j\right] \quad \textbf{(4.17)}$$

d is the counting system discriminator level ($0<d<1$, dimensionless) which is a parametric representation of the hardware's electronic pulse detection threshold usually expressed by the manufacturer in mV. A value of $d=0$ means that there is no discrimination, i.e., both strong and weak pulses output from the detector are recorded by the counting system, while a value close to 1 means that only the strongest pulses are recorded by the counting system. The zero-order approximation of **Eq. (4.17)** leads to a non-linear and non-invertible relationship:

$$P_0 = \left(1-d\right) P_1 \exp\left(-\tau P_1\right) \quad \textbf{(4.18)}$$

For practical reasons, it is recommended to minimize the hardware's discrimination (i.e., $d<<1$), which further simplifies to:

$$P_0 = P_1 \exp\left(-\tau P_1\right) \tag{4.19}$$

Using the same formalism as for non-paralyzable systems, the saturation-corrected summed signal S_1 for L laser shots can be expressed as a function of the uncorrected summed signal S_0:

$$S_0(i,k) = S_1(i,k) \exp\left(-\tau(i) \frac{2\delta z}{cL} S_1(i,k)\right) \tag{4.20}$$

This equation is still non-invertible, and for each value of the actual recorded signal $S_0(i,k)$, the corresponding value $S_1(i,k)$ must be found. This is typically done using a root-finding iterative method such as the Section or Newton-Raphson method (e.g., Press et al., 1986). The solution for the corrected signal S_1 is known with a precision set by the iteration algorithm itself. The uncertainty owed to the saturation correction in this case will include a component associated with the dead-time uncertainty, and a component associated with the accuracy of the root-finding method. No specific recommendation on the choice of a root-finding method can be provided, but for reference, the residual error from the Newton-Raphson method after *n*-iterations and the propagated dead-time and detection uncertainties are provided in **appendix H**.

The combined propagated uncertainty for the saturation-corrected signal S_1 for paralyzable systems can then be written:

$$u_{S1}(i,k) = \sqrt{u^2_{S1(SAT)}(i,k) + u^2_{S1(DET)}(i,k) + u^2_{S1(ITER)}(i,k)} \tag{4.21}$$

4.2.3 Two-channel correction method

In some cases the saturation correction does not use a priori values of the counting system dead-time. Instead, it is done by fitting the ratio of the signal recorded in a saturated channel to the signal recorded in a non-saturated ("reference") channel. The signal in the reference channel is collected at the same wavelength as that of the saturated signal, but has a lower intensity, or comes from an analog-to-digital converter device. It can also be of high intensity, but previously saturation-corrected. The fitting function can be either of the form of **Eq. (4.10)** or **Eq. (4.17)**, or any other form deemed appropriate to characterize the photon-counting hardware saturation behavior. In any case, the uncertainty associated with the saturation correction must take into account the covariance between the coefficients of the fit. If i_R is the reference channel, we have:

$$S_1(i_R,k) = S_0(i_R,k) \tag{4.22}$$

The saturation correction is done by fitting the ratio:

$$\frac{S_0(i_R,k)}{S_0(i,k)} \approx f_{SAT}(k, c_1, c_2, ..., c_m) \tag{4.23}$$

And by assuming that, after saturation correction, the ratio of the signals in the corrected and reference channels are proportional:

1526 $$\frac{S_1(i,k)}{S_1(i_R,k)} = cst = c_R$$

1527 **(4.24)**

1528 The saturation-corrected signal S_1 is then calculated from the best fit, and from the uncorrected
1529 signal S_0:

1530 $$S_1(i,k) = c_R f_{SAT}(k, c_1, c_2, ..., c_m) S_0(i,k)$$

1531 **(4.25)**

1532 A new uncertainty component $u_{S1(SAT)}$ must be introduced to account for the fitting procedure
1533 associated with the saturation correction:

1534 $$u_{S1(SAT)}(i,k) = c_R S_0(i,k) u_{fSAT}(k, c_1, c_2, ..., c_m)$$

1535 **(4.26)**

1536 The uncertainty u_{fSAT} associated with the fitting procedure can be written in generic form:

1537 $$u^2_{fSAT}(i,k) = \sqrt{\sum_{n=1}^{N}\left(\frac{\partial f_{SAT}(k, c_1, ..., c_m)}{\partial c_n}\right)^2 u^2_{cn} + 2\sum_{m=1}^{N-1}\sum_{n=m+1}^{N}\frac{\partial f_{SAT}(k, c_1, ..., c_m)}{\partial c_n}\frac{\partial f_{SAT}(k, c_1, ..., c_m)}{\partial c_m}\chi_{c_n, c_m}}$$

1538 **(4.27)**

1539 The fitting coefficients' uncertainty u_{cn} and covariance $\chi_{cn,cm}$ terms are calculated and returned
1540 by the fitting routine.

1541 The detection uncertainty propagated to the corrected signal S_1 can be written:

1542 $$u_{S1(DET)}(i,k) = c_R f_{SAT}(k, c_1, c_2, ..., c_m) u_{S0(DET)}(i,k)$$

1543 **(4.28)**

1544 The combined standard uncertainty for the saturation-corrected signal S_1 is:

1545 $$u_{S1}(i,k) = \sqrt{u^2_{S1(DET)}(i,k) + u^2_{S1(SAT)}(i,k)}$$

1546 **(4.29)**

1547 It is discouraged to use this saturation correction method if the correction function (i.e., the
1548 coefficients of the fitting function) is found to change significantly from one correction
1549 occurrence to the next one. The correction would be equivalent to a simple adjustment of the
1550 lidar measurement to an a priori state, and would not reflect the actual lidar measurement in this
1551 particular channel.

1552 For channels operating in analog detection mode, or for PC channels with very low intensity
1553 signals, no saturation correction is necessary. In this case the signals and their uncertainty
1554 components are unchanged:

1555 $$S_1(i,k) = S_0(i,k)$$

1556 $$u_{S1(DET)}(i,k) = u_{S0(DET)}(i,k)$$

1557 $$u_{S1(SAT)}(i,k) = 0$$

1558 $$u_{S1}(i,k) = u_{S0}(i,k)$$

4.3 Background noise extraction

For a given measurement time and channel, the background noise is either a constant baseline, or a function of altitude range. For most lidar system setups, the sky background and dark current components are constant. In some situations (undesired, but sometimes unavoidable), the detector is hit by high-intensity light causing so-called "signal-induced noise". This noise component is caused by additional photocathode emission with extended relaxation time, resulting in non-linear range-dependent background noise. Ultimately, whether the background noise is a constant or a complex function of range, it is extracted typically by fitting the total signal S_1 to a linear or non-linear function of range in an altitude range where only noise is believed to be present (i.e., no signal coming from the laser beam's backscattered light). The corrected signal S_2 therefore takes the following form:

$$S_2(i,k) = S_1(i,k) - B(i,k) \tag{4.30}$$

For constant and slowly-varying background noise, the fitting function is typically a polynomial of degree 0, 1 or 2:

$$B(i,k) = \sum_{m=0}^{M} b_m(i) z^m(k) \qquad (M=0,\ 1,\ \text{or}\ 2,\ \text{coefficients}\ b_m) \tag{4.31}$$

For background noise with stronger altitude dependence (not recommended), the fitting function is typically an exponential function of range:

$$B(i,k) = \exp\left(- \sum_{m=0}^{M} b_m(i) z^m(k) \right) \qquad (M\ \text{typically equal to 1 or 2}) \tag{4.32}$$

The fitting procedure can be performed in many ways. A Least-squares (LS) fitting method has the significant advantage that all the uncertainty and co-variance terms associated with the fitting coefficients can be analytically calculated. Furthermore, these analytical solutions take a simpler form for polynomials of degree 1 or 2 (Press et al., 1986). The analytical derivation of the fitting coefficients, their uncertainty, and the covariances terms is reviewed in **appendix I** for the Least squares (LS) and Singular Value Decomposition (SVD) methods. See textbooks (e.g., Press et al., 1986) for additional details.

In the case of background noise that can be assimilated to a linear function of altitude (which includes the case of constant background noise), the fitting function is a polynomial of degree 1, and the background extraction can expressed:

$$S_2(i,k) = S_1(i,k) - b_0 - b_1(i) z(k) \tag{4.33}$$

b_0 and b_1 are the coefficient of the linear fit.

A new uncertainty component, associated with the background noise extraction, must be introduced. This uncertainty component takes the form:

$$u_{S2(BKG)}(i,k) = \sqrt{\left(\frac{\partial S_2(i,k)}{\partial b_0}\right)^2 u_{b0}^2 + \left(\frac{\partial S_2(i,k)}{\partial b_1}\right)^2 u_{b1}^2 + 2\frac{\partial S_2(i,k)}{\partial b_0}\frac{\partial S_2(i,k)}{\partial b_1}\chi_{b0,b1}(i)}$$

(4.34)

which can be re-written:

$$u_{S2(BKG)}(i,k) = \sqrt{u_{b0}^2(i) + u_{b1}^2(i)z^2(k) + 2z(k)u_{b0}(i)u_{b1}(i)r_{b0,b1}(i)}$$

(4.35)

In general, there is no need for full access to the analytical derivation of the coefficients b_0 and b_1, their uncertainty u_{b0} and u_{b1}, and covariance $\chi_{b0,b1}$ (or correlation coefficient $r_{b0,b1}$) between them. Many scientific programming languages include bundled fitting routines that provide both the fitting coefficients and their uncertainty and covariances. If no background noise extraction is made, we have:

$$u_{S2(BKG)}(i,k) = 0$$

(4.36)

The standard uncertainty components associated with detection noise and saturation correction propagated to the background-corrected signal S_2 can be written:

$$u_{S2(DET)}(i,k) = \left|\frac{\partial S_2(i,k)}{\partial S_1(i,k)}\right| u_{S1(DET)} = u_{S1(DET)}(i,k)$$

(4.37)

$$u_{S2(SAT)}(i,k) = \left|\frac{\partial S_2(i,k)}{\partial S_1(i,k)}\right| u_{S1(SAT)} = u_{S1(SAT)}(i,k)$$

(4.38)

The combined standard uncertainty for the background-corrected signal is:

$$u_{S2}(i,k) = \sqrt{u_{S2(DET)}^2(i,k) + u_{S2(SAT)}^2(i,k) + u_{S2(BKG)}^2(i,k)}$$

(4.39)

4.4 Treatment of partial overlap and other caveats owed to instrumental setup

4.4.1 Partial overlap

Most lidar instruments have the inherent inability to provide measurements in the atmospheric layer located immediately above the instrument because the laser beam is not entirely encompassed within the lidar receiver field-of-view. The altitude range between the instrument and the lowest point at which the laser beam is fully seen is the region of partial overlap. In this region the lidar equation cannot be applied properly without correcting the backscattered signals for an altitude-dependent overlap factor, meant to compensate the missing fraction of the laser beam image collected on the surface of the detector. This factor, comprised between 0 and 1, is altitude-dependent and strongly dependent on the instrumental setup geometry and hardware

used for a particular channel. It is therefore impossible to provide any meaningful recommendation for this correction and its uncertainty in the present report. If any recommendation must be provided, it is to design and optimize the lidar instrumental setup so that the impact of partial overlap on the final ozone or temperature profile is minimized, ideally requiring no correction at all in the altitude range of interest. An example of how partial overlap can be corrected and the expression of the resulting uncertainty are described in **appendix J**.

4.4.2 Other caveats producing imperfect lidar signals

It is strongly discouraged to apply any large *a priori* or *a posteriori* signal correction unless the correction procedure relies on a physical process within the measurement system that produces a consistent and quantifiable systematic effect over time. The above example of overlap correction can be generalized to any empirical signal correction. A technical improvement of a system is superior to mathematical corrections.

4.5 Signal vertical merging

Most lidar instruments comprise multiple receiving channels for two main reasons: 1) the need to collect signals at different Rayleigh and Raman wavelengths and 2) the need to collect signals with different intensities. If one wants to minimize the negative impacts of detection noise and detector saturation, yet maximize the profiling range, the instrument setup then requires the inclusion of multiple channel "ranges", each range corresponding to a specific signal intensity (e.g.: high intensity range to cover higher altitudes, low intensity range to cover lower altitudes, etc.). It is then useful to vertically merge the signals of different intensities to form a single "channel" covering all altitudes of interest. This procedure is optional, and is usually ignored if the merging process is done after the species are retrieved. In this case, the ozone (or temperature) profiles retrieved from multiple intensity ranges are vertically merged. Also in some ozone DIAL cases, the signal slopes are merged instead of the signals themselves or the profiles. The propagation of uncertainties is treated similarly whether the merging process occurs at the signal, signal slope, or species level.

For signal merging, the procedure consists of scaling one channel to the other, then combine the scaled and unscaled (i.e., reference) channels together into one single profile. The scaling procedure can be as straightforward as a single-point normalization technique, or can consist of a Least-Squares Fitting (or linear regression) method applied to the ratio of the unscaled channel to the reference channel. In the following we will consider the channel index i_R for the "reference" channel, the channel index i_M for the channel to be merged with the reference channel, and the channel index i for the resulting merged profile.

4.5.1 Single-point merging methods

We start from a reference channel signal $S_3(i_R,k)$, and the signal $S_3(i_M,k)$ to be merged with it. In the altitude range where the reference channel should be preferably used, the merged signal is identical to the reference channel:

$$S_4(i,k) = S_3(i_R,k) \qquad \textbf{(4.40)}$$

The uncertainty components introduced earlier can be propagated to the merged signal S_4:

$$u_{S4(DET)}(i,k) = u_{S3(DET)}(i_R,k)$$

$$u_{S4(SAT)}(i,k) = u_{S3(SAT)}(i_R,k)$$

$$u_{S4(BKG)}(i,k) = u_{S3(BKG)}(i_R,k)$$

$$u_{S4(OVER)}(i,k) = u_{S3(OVER)}(i_R,k)$$

The combined standard uncertainty for the merged signal S_4 is:

$$u_{S4}(i,k) = \sqrt{u^2_{S4(DET)}(i,k) + u^2_{S4(SAT)}(i,k) + u^2_{S4(BKG)}(i,k) + u^2_{S4(OVER)}(i,k) + u^2_{S4(MERGE)}(i,k)} \quad \textbf{(4.41)}$$

In the altitude range where the channel i_M should be preferably used, the merged signal is:

$$S_4(i,k) = m_0(i) + m_1(i)S_3(i_M,k) \quad \textbf{(4.42)}$$

In theory, when corrections such as saturation, overlap and background have bene applied to all channels, the corrected signals should be strictly proportional, and the coefficient m_0 should therefore be non-zero only in the presence of an altitude or timing shift between channels i_M and channel i_R. m_0 and m_1 are determined by minimizing the difference between the two scaled channels over an altitude range where both channels are believed to be in a nominal regime (Whiteman et al., 2006; Newsom et al., 2009):

$$\sum_{k=k1}^{k2}\left(S_3(i_R,k) - \left(m_0(i) + m_1(i)S_3(i_M,k)\right)\right)^2 = \min \quad \textbf{(4.43)}$$

A new uncertainty component needs to be introduced to account for the merging procedure:

$$u_{S4(MERGE)}(i,k) = \sqrt{\left(\frac{\partial S_4(i,k)}{\partial m_0(i)}\right)^2 u^2_{m0}(i) + \left(\frac{\partial S_4(i,k)}{\partial m_1(i)}\right)^2 u^2_{m1}(i) + 2\frac{\partial S_4(i,k)}{\partial m_0(i)}\frac{\partial S_4(i,k)}{\partial m_1(i)}\chi_{m0,m1}(i)} \quad \textbf{(4.44)}$$

which can be re-written:

$$u_{S4(MERGE)}(i,k) = \sqrt{u^2_{m0}(i) + S^2_3(i_M,k)u^2_{m1}(i) + 2S_3(i_M,k)\chi_{m0,m1}(i)} \quad \textbf{(4.45)}$$

The uncertainty components introduced earlier can be propagated to the merged signal S_4 using:

$$u_{S4(DET)}(i,k) = \left|\frac{\partial S_4(i,k)}{\partial S_3(i)}\right| u_{S3(DET)}(i_M,k) = \left|m_1(i)\right| u_{S3(DET)}(i_M,k) \quad \textbf{(4.46)}$$

$$u_{S4(SAT)}(i,k) = \left|\frac{\partial S_4(i,k)}{\partial S_3(i)}\right| u_{S3(SAT)}(i_M,k) = \left|m_1(i)\right| u_{S3(SAT)}(i_M,k) \quad \textbf{(4.47)}$$

$$u_{S4(BKG)}(i,k) = \left|\frac{\partial S_4(i,k)}{\partial S_3(i)}\right| u_{S3(BKG)}(i_M,k) = |m_1(i)| u_{S3(BKG)}(i_M,k) \quad \textbf{(4.48)}$$

$$u_{S4(OVER)}(i,k) = \left|\frac{\partial S_4(i,k)}{\partial S_3(i)}\right| u_{S3(OVER)}(i_M,k) = |m_1(i)| u_{S3(OVER)}(i_M,k) \quad \textbf{(4.49)}$$

4.5.2 Merging methods with overlap

For a smoother transition, a linear combination between two channels i_R and i_M can be used in the region of nominal overlap. To form the merged signal, the signal of each channel is weighted by a coefficient adding up to unity:

$$S_4(i,k) = m_2(i,k)\big(m_0(i) + m_1(i) S_3(i_M,k)\big) + (1 - m_2(i,k)) S_3(i_R,k) \qquad 0 < m_2 < 1 \quad \textbf{(4.50)}$$

The uncertainty associated with the merging process becomes:

$$u_{S4(MERGE)}(i,k) = m_2(i,k)\sqrt{u_{m0}^2(i) + S_3^2(i_M,k) u_{m1}^2(i) + 2 S_3(i_M,k)\chi_{m0,m1}(i)} \quad \textbf{(4.51)}$$

This method has the advantage of avoiding discontinuities in the merged signal, but has the inconvenience of mixing the signals from two different channels, thus complicating the treatment of uncertainty.

The uncertainty associated with detection noise is propagated to the merged signal S_4 using:

$$u_{S4(DET)}(i,k) = \sqrt{m_2^2(i,k) m_1^2(i) u_{S3(DET)}^2(i_M,k) + \big(1 - m_2(i,k)\big)^2 u_{S3(DET)}^2(i_R,k)} \quad \textbf{(4.52)}$$

If the channels to merge are independent, the other uncertainty components previously introduced can be propagated to the merged signal S_4 using:

$$u_{S4(SAT)}(i,k) = \sqrt{m_2^2(i,k) m_1^2(i) u_{S3(SAT)}^2(i_M,k) + \big(1 - m_2(i,k)\big)^2 u_{S3(SAT)}^2(i_R,k)} \quad \textbf{(4.53)}$$

$$u_{S4(BKG)}(i,k) = \sqrt{m_2^2(i,k) m_1^2(i) u_{S3(BKG)}^2(i_M,k) + \big(1 - m_2(i,k)\big)^2 u_{S3(BKG)}^2(i_R,k)} \quad \textbf{(4.54)}$$

$$u_{S4(OVER)}(i,k) = \sqrt{m_2^2(i,k) m_1^2(i) u_{S3(OVER)}^2(i_M,k) + \big(1 - m_2(i,k)\big)^2 u_{S3(OVER)}^2(i_R,k)} \quad \textbf{(4.55)}$$

If the channels to merge share the same hardware from detector to the raw data acquisition system, these uncertainty components can be propagated to the merged signal S_4 using:

$$u_{S4(SAT)}(i,k) = \left|m_2(i,k) m_1(i) u_{S3(SAT)}(i_M,k) + (1 - m_2(i,k)) u_{S3(SAT)}(i_R,k)\right| \quad \textbf{(4.56)}$$

$$u_{S4(BKG)}(i,k) = \left| m_2(i,k) m_1(i) u_{S3(BKG)}(i_M,k) + (1 - m_2(i,k)) u_{S3(BKG)}(i_R,k) \right| \quad \textbf{(4.57)}$$

$$u_{S4(OVER)}(i,k) = \left| m_2(i,k) m_1(i) u_{S3(OVER)}(i_M,k) + (1 - m_2(i,k)) u_{S3(OVER)}(i_R,k) \right| \quad \textbf{(4.58)}$$

If the channels to merge share only part of the hardware between the detector and the raw data acquisition system, covariance terms must be taken into account. The uncertainty components associated with the saturation correction will be propagated to the merged signal S_4 using the covariance term $\chi_{S3M,S3R(SAT)}$ between this component in channel i_R and the same component in the channel i_M assuming all other components null:

$$u_{S4(SAT)}(i,k) = \sqrt{\begin{matrix} m_2^2(i,k) m_1^2(i) u_{S3(SAT)}^2(i_M,k) + \left(1 - m_2(i,k)\right)^2 u_{S3(SAT)}^2(i_R,k) \\ + 2 m_2(i,k) m_1(i) (1 - m_2(i,k)) \chi_{S3M,S3R(SAT)}(i,k) \end{matrix}} \quad \textbf{(4.59)}$$

The same approach applies to the other components previously introduced:

$$u_{S4(BKG)}(i,k) = \sqrt{\begin{matrix} m_2^2(i,k) m_1^2(i) u_{S3(BKG)}^2(i_M,k) + \left(1 - m_2(i,k)\right)^2 u_{S3(BKG)}^2(i_R,k) \\ + 2 m_2(i,k) m_1(i) (1 - m_2(i,k)) \chi_{S3M,S3R(BKG)}(i,k) \end{matrix}} \quad \textbf{(4.60)}$$

$$u_{S4(OVER)}(i,k) = \sqrt{\begin{matrix} m_2^2(i,k) m_1^2(i) u_{S3(OVER)}^2(i_M,k) + \left(1 - m_2(i,k)\right)^2 u_{S3(OVER)}^2(i_R,k) \\ + 2 m_2(i,k) m_1(i) (1 - m_2(i,k)) \chi_{S3M,S3R(OVER)}(i,k) \end{matrix}} \quad \textbf{(4.61)}$$

For both the one-point merging method and the merging method with overlap, the combined standard uncertainty for the merged signal S_4 is:

$$u_{S4}(i,k) = \sqrt{u_{S4(DET)}^2(i,k) + u_{S4(SAT)}^2(i,k) + u_{S4(BKG)}^2(i,k) + u_{S4(OVER)}^2(i,k) + u_{S4(MERGE)}^2(i,k)} \quad \textbf{(4.62)}$$

If no merging is done, the signals and their uncertainties remain unchanged:

$$u_{S4(DET)}(i,k) = u_{S3(DET)}(i,k)$$

$$u_{S4(SAT)}(i,k) = u_{S3(SAT)}(i,k)$$

$$u_{S4(BKG)}(i,k) = u_{S3(BKG)}(i,k)$$

$$u_{S4(OVER)}(i,k) = u_{S3(OVER)}(i,k)$$

$$u_{S4(MERGE)}(i,k) = 0$$

$$u_{S4}(i,k) = u_{S3}(i,k)$$

4.6 Signal vertical smoothing

The signals can be vertically smoothed to reduce detection noise. Smoothing the lidar signals is optional, especially if some smoothing is done during or after species retrieval (see **chapters 5 and 6**). Furthermore the smoothing process can be applied at almost any stage of the signal processing. For reference, we are introducing it after all signal corrections common to temperature and ozone retrievals have been made (i.e., applied to the corrected signal S_4), keeping in mind that the results presented here are valid at any other stage of the data processing chain. The smoothing procedure for a given channel consists of calculating a linear combination of signal values taken at neighboring altitude bins. The coefficients of smoothing filters are symmetrical with respect to the center bin at which the smoothed value is being calculated.

4.6.1 Filtering (smoothing) the signals

The smoothed value S_5 at altitude bin k calculated from the unsmoothed signal S_4 using a smoothing filter with $2n+1$ normalized coefficients is:

$$S_5(i,k) = \sum_{p=-n}^{n} f_p(i) S_4(i,k+p) \tag{4.63}$$

with $\sum_{p=-n}^{n} f_p(i) = 1$ and $f_p(i) = f_{-p}(i)$ for all p=1,2,…,n

The propagation of any uncertainty component X introduced earlier (X=*DET*, *SAT*, *BKG*, etc.) can be written:

$$u_{S5(X)}(i,k) = \sqrt{\sum_{p=-n}^{n}\sum_{q=-n}^{n} \frac{\partial S_5}{\partial S_4}(i,k+p)\frac{\partial S_5}{\partial S_4}(i,k+q)\chi_{S4,S4(X)}(i,k+p,k+q)} \tag{4.64}$$

which can be re-written:

$$u_{S5(X)}(i,k) = \sqrt{\sum_{p=-n}^{n}\sum_{q=-n}^{n} f_p(i) f_q(i) u_{S4}(i,k+p) u_{S4}(i,k+q) r_{S4,S4(X)}(i,k+p,k+q)} \tag{4.65}$$

The uncertainty associated with detection noise is uncorrelated from one altitude bin to another, and therefore can be propagated to the smoothed signal S_5 using:

$$u_{S5(DET)}(i,k) = \sqrt{\sum_{p=-n}^{n} f_p^2(i) u_{S4(DET)}^2(i,k+p)} \tag{4.66}$$

The other uncertainty components introduced earlier are all correlated in altitude, and covariance terms must be taken into account. The signals from the same channels taken at neighboring altitudes are highly correlated (correlation coefficients near 1) and each propagated component therefore simplifies to:

$$u_{S5(SAT)}(i,k) = \left| \sum_{p=-n}^{n} f_p(i) u_{S4(SAT)}(i,k+p) \right| \tag{4.67}$$

$$u_{S5(BKG)}(i,k) = \left| \sum_{p=-n}^{n} f_p(i) u_{S4(BKG)}(i,k+p) \right| \tag{4.68}$$

$$u_{S5(OVER)}(i,k) = \left| \sum_{p=-n}^{n} f_p(i) u_{S4(OVER)}(i,k+p) \right| \tag{4.69}$$

$$u_{S5(MERGE)}(i,k) = \left| \sum_{p=-n}^{n} f_p(i) u_{S4(MERGE)}(i,k+p) \right| \tag{4.70}$$

4.6.2 Filtering (smoothing) the logarithm of the signals

Because of the dynamic range of the lidar signals, it is often more appropriate to apply smoothing on the logarithm of the signals instead of the signals themselves. The implied transformations are:

$$S_{4L}(i,k) = \log\left(S_4(i,k)\right) \tag{4.71}$$

$$S_{5L}(i,k) = \sum_{p=-n}^{n} f_p(i) S_{4L}(i,k+p) \tag{4.72}$$

$$S_5(i,k) = \exp\left(S_{5L}(i,k)\right) \tag{4.73}$$

The uncertainty components are propagated the same way as they were for the smoothed signal S_5, but taking into account these transformations. For the general form of the propagation equation, we use again **Eq. (4.64)**. The uncertainty associated with detection noise can then be propagated to the smoothed log-signal S_5 using:

$$u_{S5(DET)}(i,k) = S_5(i,k) \sqrt{\sum_{p=-n}^{n} f_p^2(i) \frac{u_{S4(DET)}^2(i,k+p)}{S_{S4(DET)}^2(i,k+p)}} \tag{4.74}$$

The other uncertainty components can be propagated to the smoothed log-signal S_5 using:

$$u_{S5(SAT)}(i,k) = S_5(i,k) \left| \sum_{p=-n}^{n} f_p(i) \frac{u_{S4(SAT)}(i,k+p)}{S_4(i,k+p)} \right| \tag{4.75}$$

$$u_{S5(BKG)}(i,k) = S_5(i,k)\left|\sum_{p=-n}^{n} f_p(i)\frac{u_{S4(BKG)}(i,k+p)}{S_4(i,k+p)}\right| \quad \textbf{(4.76)}$$

$$u_{S5(OVER)}(i,k) = S_5(i,k)\left|\sum_{p=-n}^{n} f_p(i)\frac{u_{S4(OVER)}(i,k+p)}{S_4(i,k+p)}\right| \quad \textbf{(4.77)}$$

$$u_{S5(MERGE)}(i,k) = S_5(i,k)\left|\sum_{p=-n}^{n} f_p(i)\frac{u_{S4(MERGE)}(i,k+p)}{S_4(i,k+p)}\right| \quad \textbf{(4.78)}$$

The combined standard uncertainty on the smoothed signal and on the smooth log-signal S_5 is:

$$u_{S5}(i,k) = \sqrt{u^2_{S5(DET)}(i,k) + u^2_{S5(SAT)}(i,k) + u^2_{S5(BKG)}(i,k) + u^2_{S5(OVER)}(i,k) + u^2_{S5(MERGE)}(i,k)} \quad \textbf{(4.79)}$$

If no filtering is done, the signals and their uncertainties remain unchanged:

$$u_{S5(DET)}(i,k) = u_{S4(DET)}(i,k)$$

$$u_{S5(SAT)}(i,k) = u_{S4(SAT)}(i,k)$$

$$u_{S5(BKG)}(i,k) = u_{S4(BKG)}(i,k)$$

$$u_{S5(OVER)}(i,k) = u_{S4(OVER)}(i,k)$$

$$u_{S5(MERGE)}(i,k) = 0$$

$$u_{S5}(i,k) = u_{S4}(i,k)$$

5 Propagation of uncertainty specific to ozone retrieval

After applying the corrections reviewed in **chapter 4**, the saturation-background-overlap-corrected, merged and smoothed signals S_5 can be used to retrieve ozone number density using the DIAL equation (**Eq. (1.4)**). The first step is to calculate the logarithm of the ratio of the signals collected in the "ON" and "OFF" channels (channel index $i=i_{ON}$ and $i=i_{OFF}$ respectively). The second step is to vertically differentiate this ratio. The last step is to apply the DIAL equation. The flowchart of **Figure 5.1** provides a convenient quick-look summary of the data processing, and shows the various locations at which new uncertainty terms are introduced and propagated. It is worth mentioning that some of the processing steps can be performed in a slightly different order. For example, the vertical differentiation can be performed before the taking the logarithms. These slight variations in sequence do not impact the uncertainty budget.

In the following we define a new index i for each DIAL pair formed. Many lidar instruments use multiple DIAL pairs and this index is needed later when discussing the merging of multiple ozone ranges (see discussion on multiple intensity ranges in **chapter 4.5**). Typically, each DIAL pair has its own "ON" and "OFF" channels i_{ON} and i_{OFF}. For clarity and brevity we therefore modify the expressions of S_5 and u_{S5}, and write:

$$S_5(i_{ON},k) = S_{5\mathrm{ON}}(i,k),$$

$$S_5(i_{OFF},k) = S_{5\mathrm{OFF}}(i,k),$$

$$u_{S5}(i_{ON},k) = u_{\mathrm{S5ON}}(i,k)\ ,$$

$$u_{\mathrm{S5}}(i_{OFF},k) = u_{\mathrm{S5OFF}}(i,k).$$

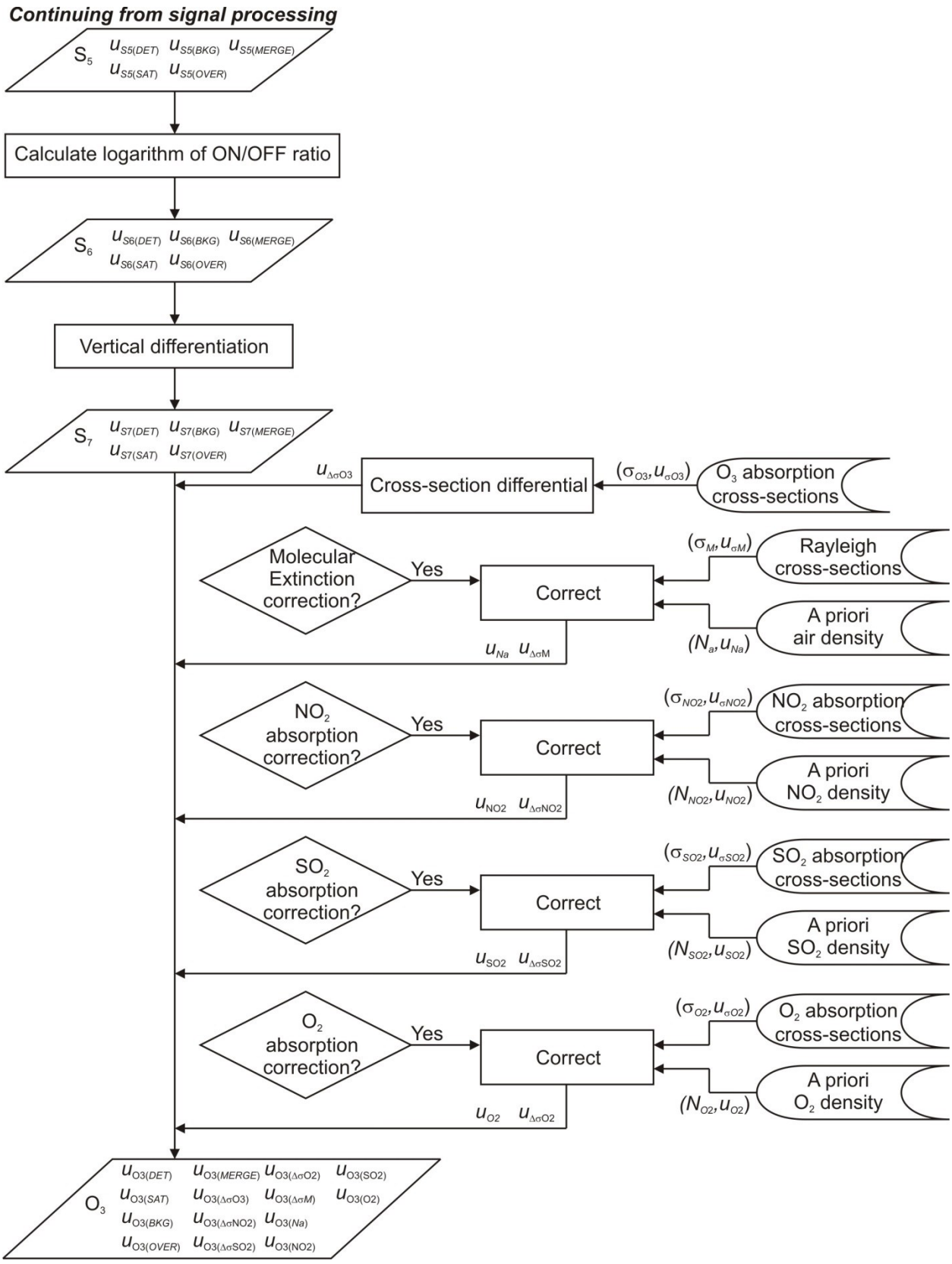


Figure 5.1 Flowchart representative of the lidar data processing specific to ozone retrieval

5.1 Logarithm of the ratio of the "ON" and "OFF" channels

The signal transformation is:

$$S_6(i,k) = \ln\left(\frac{S_5(i_{ON},k)}{S_5(i_{OFF},k)}\right) = \ln\left(\frac{S_{5ON}(i,k)}{S_{5OFF}(i,k)}\right) \quad \textbf{(5.1)}$$

All uncertainty components X introduced earlier (X=*DET*, *SAT*, *BKG*, etc.) should be propagated using the general expression:

$$u_{S6(X)}(i,k) = \sqrt{\left(\frac{\partial S_6(i,k)}{\partial S_{5ON}(i,k)}\right)^2 u^2_{S5ON}(i,k) + \left(\frac{\partial S_6(i,k)}{\partial S_{5OFF}(i,k)}\right)^2 u^2_{S5OFF}(i,k) + 2\frac{\partial S_6(i,k)}{\partial S_{5ON}(i,k)}\frac{\partial S_6(i,k)}{\partial S_{5OFF}(i,k)}\chi_{S5ON,S5OFF(X)}(i,k)} \quad \textbf{(5.2)}$$

which can be re-written:

$$u_{S6(X)}(i,k) = \sqrt{\left(\frac{u_{S5ON(X)}(i,k)}{S_{5ON}(i,k)}\right)^2 + \left(\frac{u_{S5OFF(X)}(i,k)}{S_{5OFF}(i,k)}\right)^2 - 2\frac{u_{S5ON(X)}(i,k)u_{S5OFF(X)}(i,k)}{S_{5ON}(i,k)S_{5OFF}(i,k)}r_{S5ON,S5OFF(X)}} \quad \textbf{(5.3)}$$

The "ON" and "OFF" channels may or may not share the same hardware. In any case, the uncertainty component associated with detection noise will be propagated to S_6 using:

$$u_{S6(DET)}(i,k) = \sqrt{\left(\frac{u_{S5ON(DET)}(i,k)}{S_{5ON}(i,k)}\right)^2 + \left(\frac{u_{S5OFF(DET)}(i,k)}{S_{5OFF}(i,k)}\right)^2} \quad \textbf{(5.4)}$$

For all other uncertainty components, there is an important distinction to make between the cases when the "ON" and "OFF" channels use different hardware, and the cases when they share the same hardware.

5.1.1 If the "ON" and "OFF" channels use different, independent hardware

If the "ON" and "OFF" channels do not share any hardware, the "ON" and "OFF" signals can be considered uncorrelated. All uncertainty components previously introduced can be propagated to S_6 using:

$$u_{S6(SAT)}(i,k) = \sqrt{\left(\frac{u_{S5ON(SAT)}(i,k)}{S_{5ON}(i,k)}\right)^2 + \left(\frac{u_{S5OFF(SAT)}(i,k)}{S_{5OFF}(i,k)}\right)^2} \quad \textbf{(5.5)}$$

$$u_{S6(BKG)}(i,k) = \sqrt{\left(\frac{u_{S5ON(BKG)}(i,k)}{S_{5ON}(i,k)}\right)^2 + \left(\frac{u_{S5OFF(BKG)}(i,k)}{S_{5OFF}(i,k)}\right)^2}$$

(5.6)

$$u_{S6(OVER)}(i,k) = \sqrt{\left(\frac{u_{S5ON(OVER)}(i,k)}{S_{5ON}(i,k)}\right)^2 + \left(\frac{u_{S5OFF(OVER)}(i,k)}{S_{5OFF}(i,k)}\right)^2}$$

(5.7)

$$u_{S6(MERGE)}(i,k) = \sqrt{\left(\frac{u_{S5ON(MERGE)}(i,k)}{S_{5ON}(i,k)}\right)^2 + \left(\frac{u_{S5OFF(MERGE)}(i,k)}{S_{5OFF}(i,k)}\right)^2}$$

(5.8)

5.1.2 If the "ON" and "OFF" channels share part of the hardware

If the "ON" and "OFF" channels share a significant fraction of the hardware, the "ON" and "OFF" signals cannot be considered uncorrelated. Covariance terms must be taken into account and uncertainty should be propagated to S_6 using an expression of the form of **Eq. (5.2)** or **(5.3)**. Depending on the degree of dependence of the "ON' and "OFF" channels, the covariance terms can be difficult to estimate and Monte-Carlo experiments similar to what is presented in **appendix A** may be necessary.

5.1.3 If the "ON" and "OFF" channels share the entire hardware

In this case the "ON" and "OFF" signals are fully correlated (correlation coefficient of 1). The uncertainty due to saturation correction propagated to S_6 can be written:

$$u_{S6(SAT)}(i,k) = \left|\frac{u_{S5ON(SAT)}(i,k)}{S_{5ON}(i,k)} - \frac{u_{S5OFF(SAT)}(i,k)}{S_{5OFF}(i,k)}\right|$$

(5.9)

A similar approach applies to the other three components:

$$u_{S6(BKG)}(i,k) = \left|\frac{u_{S5ON(BKG)}(i,k)}{S_{5ON}(i,k)} - \frac{u_{S5OFF(BKG)}(i,k)}{S_{5OFF}(i,k)}\right|$$

(5.10)

$$u_{S6(OVER)}(i,k) = \left|\frac{u_{S5ON(OVER)}(i,k)}{S_{5ON}(i,k)} - \frac{u_{S5OFF(OVER)}(i,k)}{S_{5OFF}(i,k)}\right|$$

(5.11)

$$u_{S6(MERGE)}(i,k) = \left|\frac{u_{S5ON(MERGE)}(i,k)}{S_{5ON}(i,k)} - \frac{u_{S5OFF(MERGE)}(i,k)}{S_{5OFF}(i,k)}\right|$$

(5.12)

The combined standard uncertainty on the log-signal S_6 is:

$$u_{S6}(i,k) = \sqrt{u^2_{S6(DET)}(i,k) + u^2_{S6(SAT)}(i,k) + u^2_{S6(BKG)}(i,k) + u^2_{S6(OVER)}(i,k) + u^2_{S6(MERGE)}(i,k)} \quad \textbf{(5.13)}$$

5.2 Vertical differentiation

The most basic expression of vertical differentiation of the signal S_6 in discretized form can be written:

$$S_7(i,k) = \frac{S_6(i,k+1) - S_6(i,k-1)}{z(k+1) - z(k-1)} \quad \textbf{(5.14)}$$

A derivative filter with $2n$+1 coefficients can also be used if a smooth version of the output signal S_7 is desired. In this case, the vertical differentiation takes the following discretized form:

$$S_7(i,k) = \frac{1}{\delta z} \sum_{p=-n}^{n} f_p(i) S_6(i,k+p) \quad \textbf{(5.15)}$$

with $f_0(i) = 0$, $f_p(i) = -f_{-p}(i)$ for all p=1,2,…,n and therefore $\sum_{p=-n}^{n} f_p(i) = 0$

Note that **Eq. (5.14)** is simply a particular case of the more general **Eq. (5.15)** with $2n$+1=3 and $f_1(i) = -f_{-1}(i) = 0.5$. Similarly to vertical smoothing (**section 4.3**), the vertical differentiation for a given channel is equivalent to computing a linear combination of signal values taken at neighboring altitude bins. The propagation of any uncertainty component X introduced earlier (X=*DET*, *SAT*, *BKG*, etc.) can therefore be written:

$$u_{S7(X)}(i,k) = \frac{1}{\delta z} \sqrt{\sum_{p=-n}^{n} \sum_{q=-n}^{n} \frac{\partial S_7}{\partial S_6}(i,k+p) \frac{\partial S_7}{\partial S_6}(i,k+q) \chi_{S6,S6(X)}(i,k+p,k+q)} \quad \textbf{(5.16)}$$

which can be re-written:

$$u_{S7(X)}(i,k) = \sqrt{\sum_{p=-n}^{n} \sum_{q=-n}^{n} f_p(i) f_q(i) u_{S6}(i,k+p) u_{S6}(i,k+q) r_{S6,S6(X)}(i,k+p,k+q)} \quad \textbf{(5.17)}$$

The uncertainty associated with detection noise is uncorrelated from one altitude bin to another, and therefore can be propagated to the differentiated signal S_7 using:

$$u_{S7(DET)}(i,k) = \frac{1}{\delta z} \sqrt{\sum_{p=-n}^{n} f_p^2(i) u^2_{S6(DET)}(i,k+p)} \quad \textbf{(5.18)}$$

The other uncertainty components introduced earlier are all correlated in altitude, and covariance terms must be taken into account. Since neighboring altitude bins are used, the correlation coefficients are very close to- or equal to 1, and the propagated uncertainty components become:

$$u_{S7(SAT)}(i,k) = \frac{1}{\delta z}\left|\sum_{p=-n}^{n} f_p(i)\, u_{S6(SAT)}(i,k+p)\right| \quad \textbf{(5.19)}$$

$$u_{S7(BKG)}(i,k) = \frac{1}{\delta z}\left|\sum_{p=-n}^{n} f_p(i)\, u_{S6(BKG)}(i,k+p)\right| \quad \textbf{(5.20)}$$

$$u_{S7(OVER)}(i,k) = \frac{1}{\delta z}\left|\sum_{p=-n}^{n} f_p(i)\, u_{S6(OVER)}(i,k+p)\right| \quad \textbf{(5.21)}$$

$$u_{S7(MERGE)}(i,k) = \frac{1}{\delta z}\left|\sum_{p=-n}^{n} f_p(i)\, u_{S6(MERGE)}(i,k+p)\right| \quad \textbf{(5.22)}$$

The total combined uncertainty on the differentiated signal S_7 is:

$$u_{S7}(i,k) = \sqrt{u^2_{S7(DET)}(i,k) + u^2_{S7(SAT)}(i,k) + u^2_{S7(BKG)}(i,k) + u^2_{S7(OVER)}(i,k) + u^2_{S7(MERGE)}(i,k)} \quad \textbf{(5.23)}$$

5.3 Ozone DIAL equation

The numerical implementation of the theoretical DIAL equation (**Eq. (1.4)**) for actual ozone DIAL instruments such as those from NDACC consists of an expression of ozone number density N_{O3} as a function of the corrected signals S_7 and the parameters included in the extinction terms of **Eq. (1.4)**:

$$N_{O3}(i,k) = \frac{S_7(i,k) - \Delta\sigma_M(i)N_a(k) - \Delta\sigma_{NO2}(i,k)N_{NO2}(k) - \Delta\sigma_{SO2}(i,k)N_{SO2}(k) - \Delta\sigma_{O2}(i,k)q_{O2}N_a(k)}{\Delta\sigma_{O3}(i,k)} \quad \textbf{(5.24)}$$

The term $\Lambda\eta$ present in **Eq. (1.4)** has been omitted since the only altitude-dependent contribution to η is from incomplete overlap and has already been treated in **chapter 4**. The terms $\Delta\alpha_P$ and $\Lambda\beta$ have also been removed since the impact of particulate backscatter and extinction is not considered in the present report. Following the recommendations of **section 3.2.2**, the interfering gases introduced here are NO_2, SO_2 and O_2. Here it is assumed that the NO_2 number density N_{NO2} and SO_2 number density N_{SO2} are independent from air number density N_a, while the O_2 number density is directly proportional to air number density through its constant mixing ratio q_{O2} = 0.209390 (CIPM-2007 value). An alternate formulation of **Eq. (5.24)** is provided in **paragraph 5.3.9** where ozone number density N_{O3} and mixing ratio q_{O3} are expressed as a function of the NO_2 mixing ratio q_{NO2} and SO_2 mixing ratio q_{SO2} instead of number density.

At this stage of processing we need to propagate the uncertainty components introduced earlier, as well as additional components associated with the absorption cross-sections and the extinction terms.

5.3.1 Ozone uncertainty due to prior processing (signal S_7)

The uncertainty components previously introduced and propagated to O_3 can be written:

$$u_{NO3(DET)}(i,k) = \left|\frac{\partial N_{O3}(i,k)}{\partial S_7(i,k)}\right| u_{S7(DET)}(i,k) = \frac{u_{S7(DET)}(i,k)}{\left|\Delta\sigma_{O3}(i,k)\right|} \quad \textbf{(5.25)}$$

$$u_{NO3(SAT)}(i,k) = \left|\frac{\partial N_{O3}(i,k)}{\partial S_7(i,k)}\right| u_{S7(SAT)}(i,k) = \frac{u_{S7(SAT)}(i,k)}{\left|\Delta\sigma_{O3}(i,k)\right|} \quad \textbf{(5.26)}$$

$$u_{NO3(BKG)}(i,k) = \left|\frac{\partial N_{O3}(i,k)}{\partial S_7(i,k)}\right| u_{S7(BKG)}(i,k) = \frac{u_{S7(BKG)}(i,k)}{\left|\Delta\sigma_{O3}(i,k)\right|} \quad \textbf{(5.27)}$$

$$u_{NO3(OVER)}(i,k) = \left|\frac{\partial N_{O3}(i,k)}{\partial S_7(i,k)}\right| u_{S7(OVER)}(i,k) = \frac{u_{S7(OVER)}(i,k)}{\left|\Delta\sigma_{O3}(i,k)\right|} \quad \textbf{(5.28)}$$

$$u_{NO3(MERGE)}(i,k) = \left|\frac{\partial N_{O3}(i,k)}{\partial S_7(i,k)}\right| u_{S7(MERGE)}(i,k) = \frac{u_{S7(MERGE)}(i,k)}{\left|\Delta\sigma_{O3}(i,k)\right|} \quad \textbf{(5.29)}$$

5.3.2 Ozone uncertainty due to ozone cross-section differential

The ozone absorption cross-section differential introduced in **chapter 1** (**Eq. (1.5)**) can be written in discretized form:

$$\Delta\sigma_{O3}(i,k) = \sigma_{O3}(i_1,k) + \sigma_{O3}(i_3,k) - \sigma_{O3}(i_2,k) + \sigma_{O3}(i_4,k) \quad \textbf{(5.30)}$$

The channel indices i_1 through i_4 have the same meaning as the indices used in **chapter 1**. The uncertainty in the ozone cross-section differential $u_{\Delta\sigma O3}$ is calculated using the values of cross-section uncertainty provided by laboratory measurements, and should be computed differently depending on the cross-section datasets used. Its general expression is:

$$u_{\Delta\sigma O3}(i,k) = \sqrt{\sum_{p=1}^{4}\sum_{q=1}^{4} \frac{\partial\Delta\sigma_{O3}(i,k)}{\partial\sigma_{O3}(i_p,k)} \frac{\partial\Delta\sigma_{O3}(i,k)}{\partial\sigma_{O3}(i_q,k)} \chi_{\sigma O3,\sigma O3}(i_p,i_q,k)} \quad \textbf{(5.31)}$$

For each cross-section value to consider, uncertainty owed to systematic effects must be distinguished from uncertainty owed to random effects. The type of backscatter (Rayleigh or Raman) must also be distinguished.

For Raman backscatter DIAL pairs, uncertainty owed to random effects can be propagated assuming that none of the cross-section values are correlated:

$$u_{\Delta\sigma O3}(i,k) = \sqrt{u^2_{\sigma O3}(i_1,k) + u^2_{\sigma O3}(i_2,k) + u^2_{\sigma O3}(i_3,k) + u^2_{\sigma O3}(i_4,k)} \quad \textbf{(5.32)}$$

For Rayleigh backscatter DIAL pairs, the emitted and received wavelengths are identical, and a modified expression must be used:

$$u_{\Delta\sigma O3}(i,k) = 2\sqrt{u^2_{\sigma O3}(i_1,k) + u^2_{\sigma O3}(i_2,k)} \quad \textbf{(5.33)}$$

The uncertainty components owed to systematic effects can be propagated assuming that all cross-section values are correlated within the same dataset.

If the same dataset is used for the cross-sections at the "ON" and "OFF" wavelengths, the uncertainty component owed to systematic effects should be propagated assuming that all cross-section values used are correlated, leading to just one expression for both Rayleigh and Raman backscatter DIAL pairs:

$$u_{\Delta\sigma O3}(i,k) = \left|u_{\sigma O3}(i_1,k) + u_{\sigma O3}(i_3,k) - u_{\sigma O3}(i_2,k) - u_{\sigma O3}(i_4,k)\right| \quad \textbf{(5.34)}$$

For Rayleigh backscatter DIAL pairs, **Eq. (5.34)** can be re-written in compact form:

$$u_{\Delta\sigma O3}(i,k) = 2\left|u_{\sigma O3}(i_1,k) - u_{\sigma O3}(i_2,k)\right| \quad \textbf{(5.35)}$$

If two datasets of different origin are used for the cross-section values at the "ON" and "OFF" wavelengths, the uncertainty component owed to systematic effects should be propagated assuming that the two cross-section datasets are independent, but that cross-section values are correlated within a given dataset. This leads again to leading to just one expression for both Rayleigh and Raman backscatter DIAL pairs:

$$u_{\Delta\sigma O3}(i,k) = \sqrt{\left(u_{\sigma O3}(i_1,k) + u_{\sigma O3}(i_3,k)\right)^2 + \left(u_{\sigma O3}(i_2,k) + u_{\sigma O3}(i_4,k)\right)^2} \quad \textbf{(5.36)}$$

Again for Rayleigh backscatter DIAL pairs, **Eq. (5.36)** can be re-written in compact form:

$$u_{\Delta\sigma O3}(i,k) = 2\sqrt{u^2_{\sigma O3}(i_1,k) + u^2_{\sigma O3}(i_2,k)} \quad \textbf{(5.37)}$$

If using a cross-section dataset that includes an uncertainty component owed to systematic effects and an uncertainty component owed to random effects, and if these components are known to be independent, then a separate computation for each component can be done using the appropriate combination of **Eqs. (5.32)-(5.37)**.

The uncertainty in ozone cross-section differential is propagated to ozone number density using the DIAL equation (**Eq. (5.24)**):

$$u_{NO3(\Delta\sigma O3)}(i,k) = \left|\frac{\partial N_{O3}(i,k)}{\partial \Delta\sigma_{O3}(i,k)}\right| u_{\Delta\sigma O3}(i,k) = \frac{N_{O3}(i,k)}{\left|\Delta\sigma_{O3}(i,k)\right|} u_{\Delta\sigma O3}(i,k) \quad \textbf{(5.38)}$$

If uncertainty components owed to systematic and random effects have been computed separately, each component should be kept and propagated in parallel, i.e., the above equation should be used for each component, leading to two separate ozone uncertainty components owed to the use of absorption cross-section.

5.3.3 Ozone uncertainty due to interfering gases' cross-section differential

The approach for the interfering gases' cross-section differential uncertainty $u_{\Delta\sigma X}$ (X=NO_2, SO_2 and O_2) is identical to that presented in the previous paragraph for the ozone cross-section differential. The computation of the cross-section differential uncertainty is therefore identical, i.e., **Eqs. ((5.30)-((5.37)** hold for all interfering species (replacing all subscripts "O3" by subscripts "NO2", "SO2", and "O2"). Just like for the ozone cross-sections, if using a dataset that includes an uncertainty component owed to systematic effects and an uncertainty component owed to random effects, a separate computation for each component should be done using the appropriate combination of **Eqs. (5.32)-(5.37)** (again with the modified subscripts). For all ozone retrievals involving wavelengths longer than 294 nm, the contribution of differential absorption by O_2 can be neglected.

The uncertainty in interfering gas cross-section differential is propagated to ozone number density using the DIAL equation:

$$u_{NO3(\Delta\sigma NO2)}(i,k) = \left|\frac{\partial N_{O3}(i,k)}{\partial \Delta\sigma_{NO2}(i,k)}\right| u_{\Delta\sigma NO2}(i,k) = \frac{N_{NO2}(k)}{\left|\Delta\sigma_{O3}(i,k)\right|} u_{\Delta\sigma NO2}(i,k) \quad \textbf{(5.39)}$$

$$u_{NO3(\Delta\sigma SO2)}(i,k) = \left|\frac{\partial N_{O3}(i,k)}{\partial \Delta\sigma_{SO2}(i,k)}\right| u_{\Delta\sigma SO2}(i,k) = \frac{N_{SO2}(k)}{\left|\Delta\sigma_{O3}(i,k)\right|} u_{\Delta\sigma SO2}(i,k) \quad \textbf{(5.40)}$$

$$u_{NO3(\Delta\sigma O2)}(i,k) = \left|\frac{\partial N_{O3}(i,k)}{\partial \Delta\sigma_{O2}(i,k)}\right| u_{\Delta\sigma O2}(i,k) = \frac{q_{O2} N_a(k)}{\left|\Delta\sigma_{O3}(i,k)\right|} u_{\Delta\sigma O2}(i,k) \quad \textbf{(5.41)}$$

Just like for the ozone cross-sections, if uncertainty components owed to systematic and random effects have been computed separately, each component should be kept and propagated in parallel, i.e., the above equations should be used for each component, leading to two separate ozone uncertainty components owed to the use of absorption cross-section for a given interfering gas.

5.3.4 Ozone uncertainty due to Rayleigh cross-section differential

The Rayleigh extinction cross-section differential introduced in **chapter 1** (**Eq. (1.6)**) can be written in discretized form:

$$\Delta\sigma_M(i) = \sigma_M(i_1) + \sigma_M(i_3) - \sigma_M(i_2) + \sigma_M(i_4) \quad \textbf{(5.42)}$$

The approach to compute the Rayleigh cross-section differential uncertainty is identical to that presented in the previous paragraphs for the ozone and the interfering gases, but with the

exception that the Rayleigh cross-section values are typically computed from an analytical function instead of being measured (see **section 3.3** and **appendix A**). As a result, it is assumed by default that the cross-section values at the "ON" and "OFF" wavelengths are fully correlated, and depend only on wavelength (no temperature or altitude dependence). For both the Rayleigh and Raman backscatter DIAL pairs, this assumption leads to the expression:

$$u_{\Delta\sigma M}(i) = \left| u_{\sigma M}(i_1) + u_{\sigma M}(i_3) - u_{\sigma M}(i_2) - u_{\sigma M}(i_4) \right| \tag{5.43}$$

For Rayleigh backscatter DIAL pairs, **Eq. (5.43)** can be re-written in compact form:

$$u_{\Delta\sigma M}(i) = 2\left| u_{\sigma M}(i_1) - u_{\sigma M}(i_2) \right| \tag{5.44}$$

The ozone uncertainty associated with the use of molecular extinction cross-sections can be derived easily from the DIAL equation:

$$u_{NO3(\Delta\sigma M)}(i,k) = \left| \frac{\partial N_{O3}(i,k)}{\partial \Delta\sigma_M(i)} \right| u_{\Delta\sigma M}(i) = \frac{N_a(k)}{\left|\Delta\sigma_{O3}(i,k)\right|} u_{\Delta\sigma M}(i) \tag{5.45}$$

5.3.5 Ozone uncertainty due to the use of ancillary air number density

The expression and values of ancillary air number density uncertainty depends on the dataset used. For air number density derived from radiosonde measurements, meteorological analysis, or an assimilation model, the air number density is usually derived from air temperature and pressure:

$$N_a(k) = \frac{p(k)}{k_B T(k)} \tag{5.46}$$

The general expression of the ancillary air number density uncertainty is:

$$u_{Na}(k) = \sqrt{\left(\frac{\partial N_a(k)}{\partial p(k)}\right)^2 u_p^2(k) + \left(\frac{\partial N_a(k)}{\partial T(k)}\right)^2 u_T^2(k) + 2\frac{\partial N_a(k)}{\partial p(k)}\frac{\partial N_a(k)}{\partial T(k)}\chi_{p,T}(k)} \tag{5.47}$$

which can be re-written:

$$u_{Na}(k) = N_a(k)\sqrt{\frac{u_p^2(k)}{p^2(k)} + \frac{u_T^2(k)}{T^2(k)} - 2\frac{u_p(k)u_T(k)}{p(k)T(k)}r_{p,T}(k)} \tag{5.48}$$

If temperature and pressure are measured or computed independently, the air number density uncertainty will be written:

2102 $$u_{Na}(k) = N_a(k)\sqrt{\frac{u_p^2(k)}{p^2(k)} + \frac{u_T^2(k)}{T^2(k)}}$$

2103 **(5.49)**

2104 If the measured or modeled temperature and pressure are fully correlated, the air number density
2105 uncertainty will be written:

2106 $$u_{Na}(k) = N_a(k)\left|\frac{u_p(k)}{p(k)} - \frac{u_T(k)}{T(k)}\right|$$

2107 **(5.50)**

2108 The ozone uncertainty associated with the use of ancillary air number density can be derived
2109 from the DIAL equation (**Eq. ((5.24))**):

2110 $$u_{NO3(Na)}(i,k) = \left|\frac{\partial N_{O3}(i,k)}{\partial N_a(k)}\right| u_{Na}(k) = \left|\frac{\Delta\sigma_M(i) + q_{O2}\Delta\sigma_{O2}(i,k)}{\Delta\sigma_{O3}(i,k)}\right| u_{Na}(k)$$

2111 **(5.51)**

2112

2113 ### 5.3.6 Ozone uncertainty due to the interfering gases' number density

2114 The ozone uncertainty associated with the use of ancillary number density for the interfering
2115 gases NO_2 and SO_2 can be derived easily from the DIAL equation:

2116 $$u_{NO3(NNO2)}(i,k) = \left|\frac{\partial N_{O3}(i,k)}{\partial N_{NO2}(k)}\right| u_{NNO2}(k) = \left|\frac{\Delta\sigma_{NO2}(i,k)}{\Delta\sigma_{O3}(i,k)}\right| u_{NNO2}(k)$$

2117 **(5.52)**

2118 $$u_{NO3(NSO2)}(i,k) = \left|\frac{\partial N_{O3}(i,k)}{\partial N_{SO2}(k)}\right| u_{NSO2}(k) = \left|\frac{\Delta\sigma_{SO2}(i,k)}{\Delta\sigma_{O3}(i,k)}\right| u_{NSO2}(k)$$

2119 **(5.53)**

2120

2121 ### 5.3.7 Ozone number density combined standard uncertainty

2122 Assuming that the air number density profile and the profiles of NO_2, SO_2 and O_2 are all
2123 uncorrelated, the combined standard uncertainty on ozone number density can be written:

2124 $$u_{NO3}(i,k) = \sqrt{\begin{aligned} &u_{NO3(DET)}^2(i,k) + u_{NO3(SAT)}^2(i,k) + u_{NO3(BKG)}^2(i,k) + u_{NO3(OVER)}^2(i,k) + u_{NO3(MERGE)}^2(i,k) \\ &+ u_{NO3(\Delta\sigma O3)}^2(i,k) + u_{NO3(\Delta\sigma M)}^2(i,k) + u_{NO3(\Delta\sigma NO2)}^2(i,k) + u_{NO3(\Delta\sigma SO2)}^2(i,k) + u_{NO3(\Delta\sigma O2)}^2(i,k) \\ &+ u_{NO3(Na)}^2(i,k) + u_{NO3(NNO2)}^2(i,k) + u_{NO3(NSO2)}^2(i,k) \end{aligned}}$$

2125 **(5.54)**

2126 Examples of a complete uncertainty budget for the JPL stratospheric ozone lidar at Mauna Loa,
2127 Hawaii, and the JPL tropospheric ozone lidar at Table Mountain, California are provided for
2128 reference in **Figure 5.2** and **Figure 5.3** respectively. On these figures, all the uncertainty
2129 components are included, and they are computed for the multiple altitude ranges available for

these lidars. For the stratospheric ozone lidar instrument, the red curves correspond to the DIAL pair of vibrational Raman channels (3332/387 nm), the blue curves correspond to the pair of Rayleigh low-intensity channels (08/355 nm), and the green curves correspond to the pair of Rayleigh high-intensity channels 308/355 nm). For the tropospheric ozone lidar, the red curves correspond to the pair of low-intensity channels (289/299 nm), the blue curves correspond to the pair of high-intensity channels (289/299 nm), and the green curves correspond to a tropospheric-stratospheric hybrid pair of channels 299/355 nm. Generally speaking, the dominant sources of uncertainty at the bottom of the profiles are saturation and Rayleigh extinction differential, the dominant source in the middle of the profiles is the ozone absorption cross-section differential, and the dominant source at the top of the profile is detection noise.

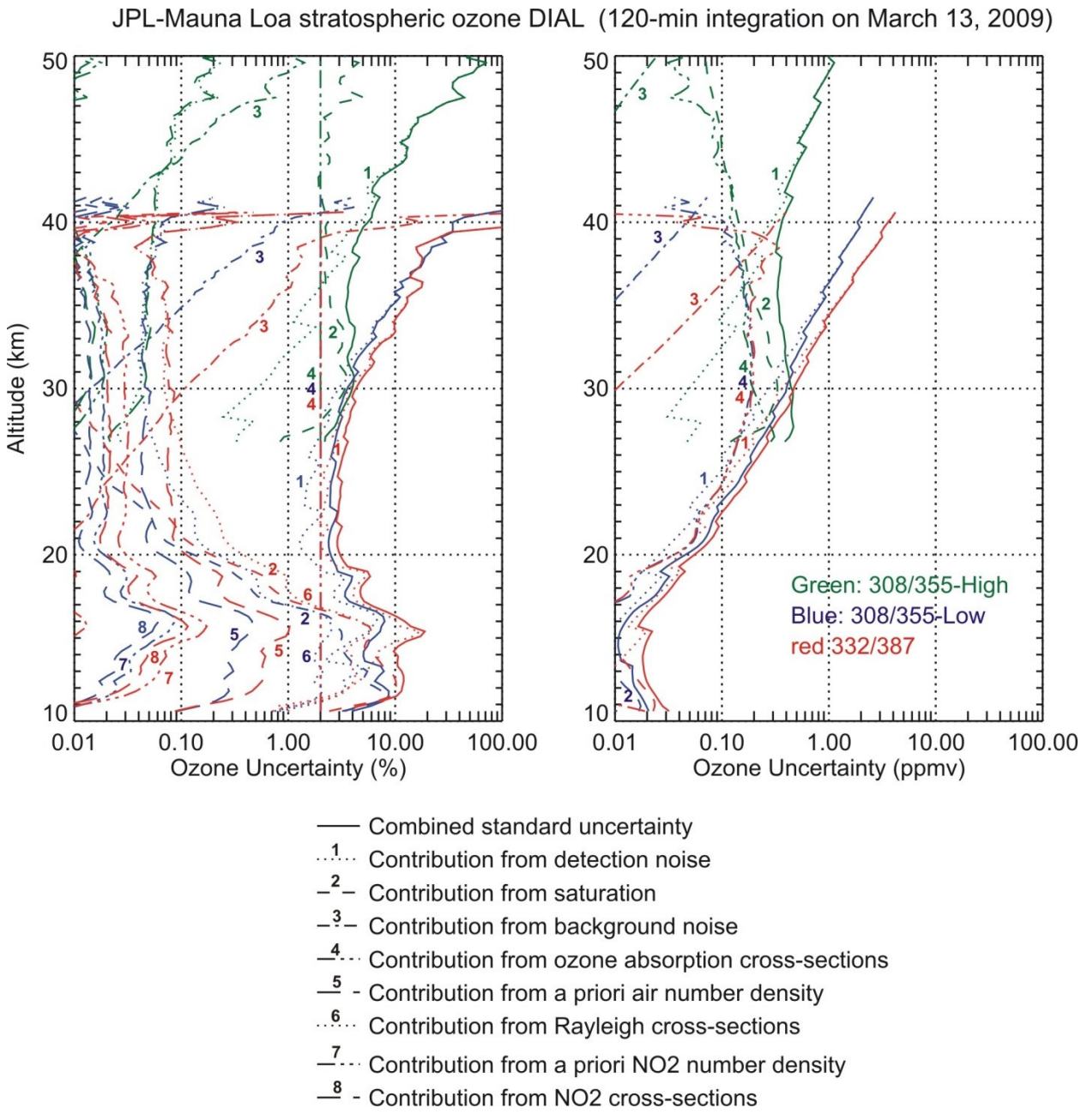


Figure 5.2 Example of full uncertainty budget for the JPL ozone differential absorption lidar at Mauna Loa Observatory (data taken on March 13, 2009), as computed using the present recommendations

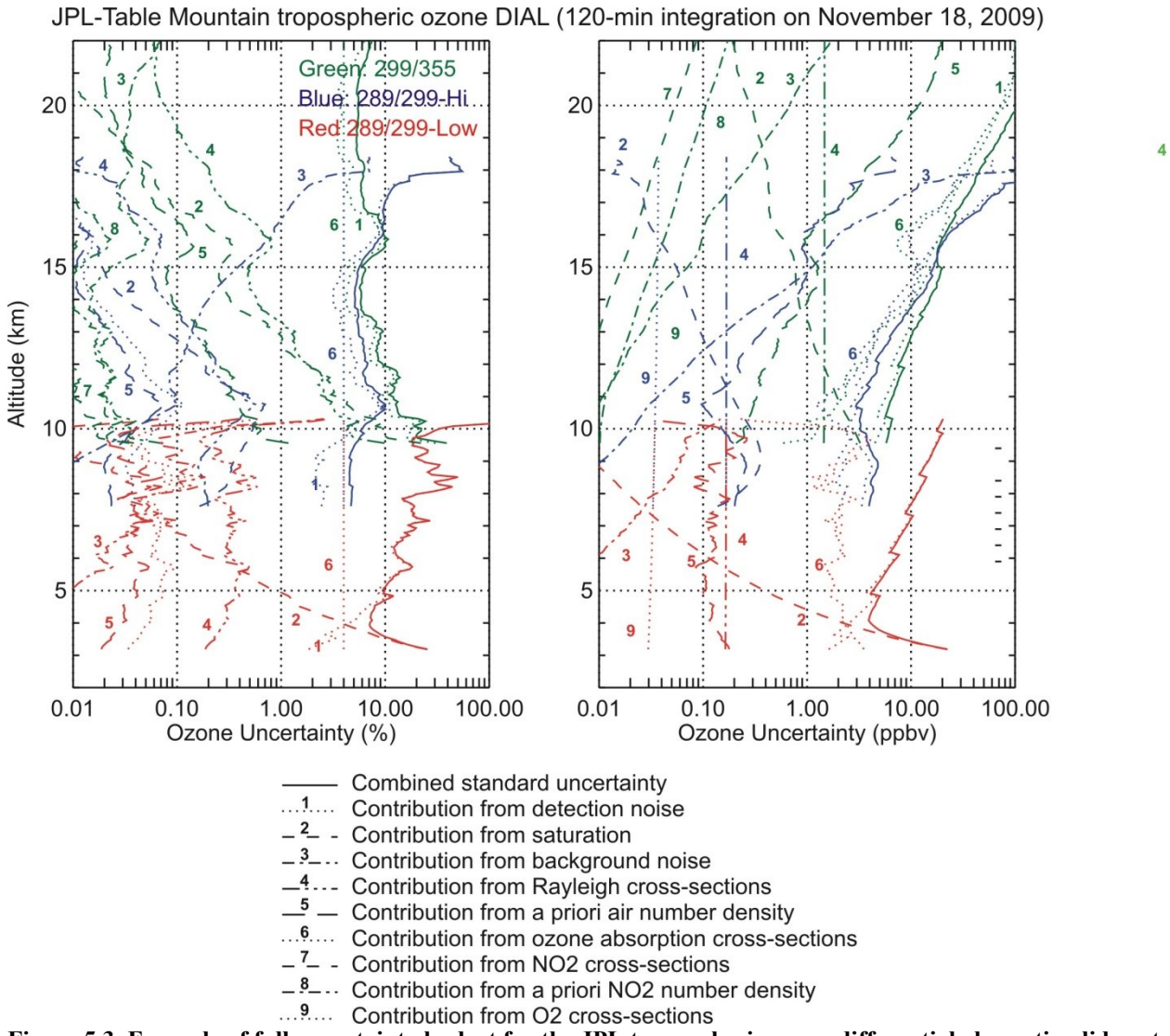


Figure 5.3 Example of full uncertainty budget for the JPL tropospheric ozone differential absorption lidar at Table Mountain Facility, CA (data taken on November 18, 2009), as computed using the present recommendations

5.3.8 Derived ozone mixing ratio uncertainty

In the DIAL technique, the ozone number density profile is the measured quantity. For various scientific reasons it is useful to derive ozone mixing ratio from the measured number density. Ozone volume mixing ratio q_{O3} is derived by computing the ratio of the ozone number density N_{O3} to the air number density N_a. This can be written from the DIAL equation (**Eq. (5.24)**):

$$q_{O3}(i,k)=\frac{1}{\Delta\sigma_{O3}(i,k)}\left(\frac{S_7(i,k)-\Delta\sigma_{NO2}(i,k)N_{NO2}(k)-\Delta\sigma_{SO2}(i,k)N_{SO2}(k)}{N_a(k)}-\Delta\sigma_M(i)-q_{O2}\Delta\sigma_{O2}(i,k)\right) \quad \textbf{(5.55)}$$

The uncertainty components introduced earlier can be propagated to ozone mixing ratio using **Eq. (5.55)**:

$$u_{qO3(DET)}(i,k)=\left|\frac{\partial q_{O3}(i,k)}{\partial S_7(i,k)}\right|u_{S7(DET)}(i,k)=\frac{u_{S7(DET)}(i,k)}{N_a(k)\left|\Delta\sigma_{O3}(i,k)\right|} \quad \textbf{(5.56)}$$

$$u_{qO3(SAT)}(i,k)=\left|\frac{\partial q_{O3}(i,k)}{\partial S_7(i,k)}\right|u_{S7(SAT)}(i,k)=\frac{u_{S7(SAT)}(i,k)}{N_a(k)\left|\Delta\sigma_{O3}(i,k)\right|} \quad \textbf{(5.57)}$$

$$u_{qO3(BKG)}(i,k)=\left|\frac{\partial q_{O3}(i,k)}{\partial S_7(i,k)}\right|u_{S7(BKG)}(i,k)=\frac{u_{S7(BKG)}(i,k)}{N_a(k)\left|\Delta\sigma_{O3}(i,k)\right|} \quad \textbf{(5.58)}$$

$$u_{qO3(OVER)}(i,k)=\left|\frac{\partial q_{O3}(i,k)}{\partial S_7(i,k)}\right|u_{S7(OVER)}(i,k)=\frac{u_{S7(OVER)}(i,k)}{N_a(k)\left|\Delta\sigma_{O3}(i,k)\right|} \quad \textbf{(5.59)}$$

$$u_{qO3(MERGE)}(i,k)=\left|\frac{\partial q_{O3}(i,k)}{\partial S_7(i,k)}\right|u_{S7(MERGE)}(i,k)=\frac{u_{S7(MERGE)}(i,k)}{N_a(k)\left|\Delta\sigma_{O3}(i,k)\right|} \quad \textbf{(5.60)}$$

$$u_{qO3(\Delta\sigma O3)}(i,k)=\left|\frac{\partial q_{O3}(i,k)}{\partial\Delta\sigma_{O3}(i,k)}\right|u_{\Delta\sigma O3}(i,k)=\frac{q_{O3}(i,k)}{\left|\Delta\sigma_{O3}(i,k)\right|}u_{\Delta\sigma O3}(i,k) \quad \textbf{(5.61)}$$

$$u_{qO3(\Delta\sigma M)}(i,k)=\left|\frac{\partial q_{O3}(i,k)}{\partial\Delta\sigma_{M}(i,k)}\right|u_{\Delta\sigma M}(i,k)=\frac{u_{\Delta\sigma M}(i,k)}{\left|\Delta\sigma_{O3}(i,k)\right|} \quad \textbf{(5.62)}$$

$$u_{qO3(\Delta\sigma NO2)}(i,k)=\left|\frac{\partial q_{O3}(i,k)}{\partial\Delta\sigma_{NO2}(i,k)}\right|u_{\Delta\sigma NO2}(i,k)=\frac{1}{N_a(k)}\frac{N_{NO2}(k)}{\left|\Delta\sigma_{O3}(i,k)\right|}u_{\Delta\sigma NO2}(i,k) \quad \textbf{(5.63)}$$

$$u_{qO3(\Delta\sigma SO2)}(i,k)=\left|\frac{\partial q_{O3}(i,k)}{\partial\Delta\sigma_{SO2}(i,k)}\right|u_{\Delta\sigma SO2}(i,k)=\frac{1}{N_a(k)}\frac{N_{SO2}(k)}{\left|\Delta\sigma_{O3}(i,k)\right|}u_{\Delta\sigma SO2}(i,k) \quad \textbf{(5.64)}$$

$$u_{qO3(\Delta\sigma O2)}(i,k)=\left|\frac{\partial q_{O3}(i,k)}{\partial\Delta\sigma_{O2}(i,k)}\right|u_{\Delta\sigma O2}(i,k)=\frac{q_{O2}}{\left|\Delta\sigma_{O3}(i,k)\right|}u_{\Delta\sigma O2}(i,k) \quad \textbf{(5.65)}$$

$$u_{qO3(Na)}(i,k) = \left|\frac{\partial q_{O3}(i,k)}{\partial N_a(k)}\right| u_{Na}(k) = \frac{1}{N_a(k)}\left| q_{O3}(i,k) + \frac{\Delta\sigma_M(i)}{\Delta\sigma_{O3}(i,k)} + \frac{q_{O2}\Delta\sigma_{O2}(i)}{\Delta\sigma_{O3}(i,k)}\right| u_{Na}(k) \quad \textbf{(5.66)}$$

$$u_{qO3(NNO2)}(i,k) = \left|\frac{\partial q_{O3}(i,k)}{\partial N_{NO2}(k)}\right| u_{NNO2}(k) = \frac{1}{N_a(k)}\left|\frac{\Delta\sigma_{NO2}(i)}{\Delta\sigma_{O3}(i,k)}\right| u_{NNO2}(k) \quad \textbf{(5.67)}$$

$$u_{qO3(NSO2)}(i,k) = \left|\frac{\partial q_{O3}(i,k)}{\partial N_{SO2}(k)}\right| u_{NSO2}(k) = \frac{1}{N_a(k)}\left|\frac{\Delta\sigma_{SO2}(i)}{\Delta\sigma_{O3}(i,k)}\right| u_{NSO2}(k) \quad \textbf{(5.68)}$$

The combined ozone mixing ratio uncertainty can then be written:

$$u_{qO3}(i,k) = \sqrt{\begin{array}{l} u^2_{qO3(DET)}(i,k) + u^2_{qO3(SAT)}(i,k) + u^2_{qO3(BKG)}(i,k) + u^2_{qO3(OVER)}(i,k) + u^2_{qO3(MERGE)}(i,k) \\ + u^2_{qO3(\Delta\sigma O3)}(i,k) + u^2_{qO3(\Delta\sigma M)}(i,k) + u^2_{qO3(\Delta\sigma NO2)}(i,k) + u^2_{qO3(\Delta\sigma SO2)}(i,k) + u^2_{qO3(\Delta\sigma O2)}(i,k) \\ + u^2_{qO3(Na)}(i,k) + u^2_{qO3(NNO2)}(i,k) + u^2_{qO3(NSO2)}(i,k) \end{array}} \quad \textbf{(5.69)}$$

5.3.9 Using ancillary mixing ratio instead of number density

Until now, it was assumed that the independent input quantities for the absorption by NO_2 and SO_2 were the number densities N_{NO2} and N_{SO2} with uncertainties u_{NNO2}, and u_{NSO2}, respectively. These uncertainties were introduced and propagated assuming that the number densities were uncorrelated with the air number density N_a. If it is assumed instead that the independent input quantities are the mixing ratios q_{NO2}, and q_{SO2} (with uncertainties u_{qNO2}, and u_{qSO2} respectively), the DIAL equation must be reformulated to take into account the interdependence between the gases' number densities and the air number density. The ozone number density equation **Eq. (5.24)** can be re-formulated as:

$$N_{O3}(i,k) = \frac{S_7(i,k) - N_a(k)\left[\Delta\sigma_M(i) + \Delta\sigma_{NO2}(i,k)q_{NO2}(k) + \Delta\sigma_{SO2}(i,k)q_{SO2}(k) + q_{O2}\Delta\sigma_{O2}(i,k)\right]}{\Delta\sigma_{O3}(i,k)} \quad \textbf{(5.70)}$$

The ozone number density uncertainty component associated with the ancillary air number density becomes:

$$u_{NO3(Na)}(i,k) = \left|\frac{\Delta\sigma_M(i) + \Delta\sigma_{NO2}(i,k)q_{NO2}(k) + \Delta\sigma_{SO2}(i,k)q_{SO2}(k) + q_{O2}\Delta\sigma_{O2}(i,k)}{\Delta\sigma_{O3}(i,k)}\right| u_{Na}(k) \quad \textbf{(5.71)}$$

The ozone number density uncertainty component associated with the ancillary NO_2 and SO_2 mixing ratio profiles become:

$$u_{NO3(qNO2)}(i,k) = \left|\frac{N_a(k)\Delta\sigma_{NO2}(i,k)}{\Delta\sigma_{O3}(i,k)}\right| u_{qNO2}(k) \quad \textbf{(5.72)}$$

2208 $$u_{NO3(qSO2)}(i,k) = \left| \frac{N_a(k)\Delta\sigma_{SO2}(i,k)}{\Delta\sigma_{O3}(i,k)} \right| u_{qSO2}(k)$$

2209 **(5.73)**

2210 The combined ozone number density uncertainty becomes:

2211 $$u_{NO3}(i,k) = \sqrt{\begin{aligned} & u^2_{NO3(DET)}(i,k) + u^2_{NO3(SAT)}(i,k) + u^2_{NO3(BKG)}(i,k) + u^2_{NO3(OVER)}(i,k) + u^2_{NO3(MERGE)}(i,k) \\ & + u^2_{NO3(\Delta\sigma O3)}(i,k) + u^2_{NO3(\Delta\sigma M)}(i,k) + u^2_{NO3(\Delta\sigma NO2)}(i,k) + u^2_{NO3(\Delta\sigma SO2)}(i,k) + u^2_{NO3(\Delta\sigma O2)}(i,k) \\ & + u^2_{NO3(Na)}(i,k) + u^2_{NO3(qNO2)}(i,k) + u^2_{NO3(qSO2)}(i,k) \end{aligned}}$$

2212 **(5.74)**

2213 The ozone mixing ratio equation (**Eq. (5.55))** can be re-written:

2214 $$q_{O3}(i,k) = \frac{1}{\Delta\sigma_{O3}(i,k)} \left(\frac{S_7(i,k)}{N_a(k)} - \Delta\sigma_M(i) - \Delta\sigma_{NO2}(i,k) q_{NO2}(k) - \Delta\sigma_{SO2}(i,k) q_{SO2}(k) - q_{O2}\Delta\sigma_{O2}(i,k) \right)$$

2215 **(5.75)**

2216 The ozone mixing ratio uncertainty component associated with the ancillary NO_2 and SO_2
2217 mixing ratio profiles become:

2218 $$u_{qO3(qNO2)}(i,k) = \left| \frac{\Delta\sigma_{NO2}(i,k)}{\Delta\sigma_{O3}(i,k)} \right| u_{qNO2}(k)$$

2219 **(5.76)**

2220 $$u_{qO3(qSO2)}(i,k) = \left| \frac{\Delta\sigma_{SO2}(i,k)}{\Delta\sigma_{O3}(i,k)} \right| u_{qSO2}(k)$$

2221 **(5.77)**

2222 The ozone mixing ratio uncertainty component associated with the ancillary air number density
2223 becomes:

2224 $$u_{qO3(Na)}(i,k) = \left| q_{O3}(i,k) + \frac{\Delta\sigma_M(i)}{\Delta\sigma_{O3}(i,k)} + \frac{\Delta\sigma_{NO2}(i)}{\Delta\sigma_{O3}(i,k)} q_{NO2}(k) + \frac{\Delta\sigma_{SO2}(i)}{\Delta\sigma_{O3}(i,k)} q_{SO2}(k) + \frac{\Delta\sigma_{O2}(i)}{\Delta\sigma_{O3}(i,k)} q_{O2} \right| \frac{u_{Na}(k)}{N_a(k)}$$

2225 **(5.78)**

2226 This equation differs from **Eq. (5.66)** by two additional terms that reflect the inter-dependence
2227 between the number densities N_{NO2} and N_{SO2} and the air number density N_a.

2228 All other uncertainty components are propagated using expressions similar to those presented in
2229 **paragraphs 5.3.6 and 5.3.7**, but using the pairs (q_{NO2},u_{qNO2}) and (q_{SO2},u_{qSO2}) instead of
2230 (N_{NO2},u_{NO2}) and (N_{SO2},u_{SO2}), respectively.

2231 The combined ozone mixing ratio uncertainty becomes:

2232 $$u_{qO3}(i,k) = \sqrt{\begin{aligned} & u^2_{qO3(DET)}(i,k) + u^2_{qO3(SAT)}(i,k) + u^2_{qO3(BKG)}(i,k) + u^2_{qO3(OVER)}(i,k) + u^2_{qO3(MERGE)}(i,k) \\ & + u^2_{qO3(\Delta\sigma O3)}(i,k) + u^2_{qO3(\Delta\sigma M)}(i,k) + u^2_{qO3(\Delta\sigma NO2)}(i,k) + u^2_{qO3(\Delta\sigma SO2)}(i,k) + u^2_{qO3(\Delta\sigma O2)}(i,k) \\ & + u^2_{qO3(Na)}(i,k) + u^2_{qO3(qNO2)}(i,k) + u^2_{qO3(qSO2)}(i,k) \end{aligned}}$$

2233 **(5.79)**

2234

5.3.10 Merging the ozone profiles from multiple channels into one profile

The approach is similar to that presented in **section 4.3.2** for merging the signals. It is indeed more straightforward if a single point merging method is used. In this case the merging consists of a simple collation of two individual channels (i.e., no overlap), and there is no need to introduce and propagate a merging uncertainty. When the merging procedure uses overlap, the approach is identical to that described for signal merging. Because no further range or channel combination is expected after the present merging procedure, the merged profile is now a one-dimensional array, function of altitude only. Merging a high intensity channel i_H with a low-intensity channel i_L together can be written:

$$N_{O3}(k) = m(k)N_{O3}(i_H,k) + (1-m(k))N_{O3}(i_L,k) \qquad 0 < m < 1 \tag{5.80}$$

The uncertainty due to detection noise can be combined assuming that the signals from the high-intensity and low-intensity channels are independent:

$$u_{NO3(DET)}(k) = \sqrt{m^2(k)u^2_{NO3(DET)}(i_H,k) + (1-m(k))^2 u^2_{NO3(DET)}(i_L,k)} \tag{5.81}$$

Assuming that the extinction correction of the high- and low-intensity channels' signals are made consistently (i.e., same datasets used for the corrections of the low-intensity and high-intensity channels), the uncertainty components for the extinction correction terms can be propagated as follows:

$$u_{NO3(\Delta\sigma O3)}(k) = m(k)u_{NO3(\Delta\sigma O3)}(i_H,k) + (1-m(k))u_{NO3(\Delta\sigma O3)}(i_L,k) \tag{5.82}$$

$$u_{NO3(\Delta\sigma M)}(k) = m(k)u_{NO3(\Delta\sigma M)}(i_H,k) + (1-m(k))u_{NO3(\Delta\sigma M)}(i_L,k) \tag{5.83}$$

$$u_{NO3(\Delta\sigma NO2)}(k) = m(k)u_{NO3(\Delta\sigma NO2)}(i_H,k) + (1-m(k))u_{NO3(\Delta\sigma NO2)}(i_L,k) \tag{5.84}$$

$$u_{NO3(\Delta\sigma SO2)}(k) = m(k)u_{NO3(\Delta\sigma SO2)}(i_H,k) + (1-m(k))u_{NO3(\Delta\sigma SO2)}(i_L,k) \tag{5.85}$$

$$u_{NO3(\Delta\sigma O2)}(k) = m(k)u_{NO3(\Delta\sigma O2)}(i_H,k) + (1-m(k))u_{NO3(\Delta\sigma O2)}(i_L,k) \tag{5.86}$$

$$u_{NO3(Na)}(k) = m(k)u_{NO3(Na)}(i_H,k) + (1-m(k))u_{NO3(Na)}(i_L,k) \tag{5.87}$$

$$u_{NO3(NNO2)}(k) = m(k)u_{NO3(NNO2)}(i_H,k) + (1-m(k))u_{NO3(NNO2)}(i_L,k) \tag{5.88}$$

$$u_{NO3(qNO2)}(k) = m(k)u_{NO3(qNO2)}(i_H,k) + (1-m(k))u_{NO3(qNO2)}(i_L,k) \tag{5.89}$$

$$u_{NO3(NSO2)}(k) = m(k)u_{NO3(NSO2)}(i_H,k) + (1-m(k))u_{NO3(NSO2)}(i_L,k) \tag{5.90}$$

$$u_{NO3(qSO2)}(k) = m(k)u_{NO3(qSO2)}(i_H,k) + (1-m(k))u_{NO3(qSO2)}(i_L,k) \quad \textbf{(5.91)}$$

For all other uncertainty components previously introduced, the signals in the low- and high intensity channels may or may not be considered correlated, depending on the design of the instrument and on the consistency of the data processing. If no hardware is shared, the remaining uncertainty components are propagated as follows:

$$u_{NO3(SAT)}(k) = \sqrt{m^2(k)u^2_{NO3(SAT)}(i_H,k) + (1-m(k))^2 u^2_{NO3(SAT)}(i_L,k)} \quad \textbf{(5.92)}$$

$$u_{NO3(BKG)}(k) = \sqrt{m^2(k)u^2_{NO3(BKG)}(i_H,k) + (1-m(k))^2 u^2_{NO3(BKG)}(i_L,k)} \quad \textbf{(5.93)}$$

$$u_{NO3(OVER)}(k) = \sqrt{m^2(k)u^2_{NO3(OVER)}(i_H,k) + (1-m(k))^2 u^2_{NO3(OVER)}(i_L,k)} \quad \textbf{(5.94)}$$

$$u_{NO3(MERGE)}(k) = \sqrt{m^2(k)u^2_{NO3(MERGE)}(i_H,k) + (1-m(k))^2 u^2_{NO3(MERGE)}(i_L,k)} \quad \textbf{(5.95)}$$

If the same hardware is shared, and the data processing is assumed to be consistent for both channels, the remaining uncertainty components are propagated as follows:

$$u_{NO3(SAT)}(k) = m(k)u_{NO3(SAT)}(i_H,k) + (1-m(k))u_{NO3(SAT)}(i_L,k) \quad \textbf{(5.96)}$$

$$u_{NO3(BKG)}(k) = m(k)u_{NO3(BKG)}(i_H,k) + (1-m(k))u_{NO3(BKG)}(i_L,k) \quad \textbf{(5.97)}$$

$$u_{NO3(OVER)}(k) = m(k)u_{NO3(OVER)}(i_H,k) + (1-m(k))u_{NO3(OVER)}(i_L,k) \quad \textbf{(5.98)}$$

$$u_{NO3(MERGE)}(k) = m(k)u_{NO3(MERGE)}(i_H,k) + (1-m(k))u_{NO3(MERGE)}(i_L,k) \quad \textbf{(5.99)}$$

If the input quantities for the interfering gases terms of the DIAL equation are number densities, the merged ozone number density profile combined standard uncertainty can be written:

$$u_{NO3}(k) = \sqrt{\begin{aligned} &u^2_{NO3(\Delta\sigma O3)}(k) + u^2_{NO3(DET)}(k) + u^2_{NO3(SAT)}(k) + u^2_{NO3(BKG)}(k) + u^2_{NO3(OVER)}(k) + u^2_{NO3(MERGE)}(k) \\ &+ u^2_{NO3(\Delta\sigma M)}(k) + u^2_{NO3(\Delta\sigma NO2)}(k) + u^2_{NO3(\Delta\sigma SO2)}(k) + u^2_{NO3(\Delta\sigma O2)}(k) \\ &+ u^2_{NO3(Na)}(k) + u^2_{NO3(NNO2)}(k) + u^2_{NO3(NSO2)}(k) \end{aligned}} \quad \textbf{(5.100)}$$

If the input quantities for the interfering gases terms of the DIAL equation are mixing ratios, the merged ozone number density profile combined standard uncertainty can be written:

2302 $$u_{NO3}(k)=\sqrt{\begin{array}{l}u^2_{NO3(\Delta\sigma O3)}(k)+u^2_{NO3(DET)}(k)+u^2_{NO3(SAT)}(k)+u^2_{NO3(BKG)}(k)+u^2_{NO3(OVER)}(k)+u^2_{NO3(MERGE)}(k)\\+u^2_{NO3(\Delta\sigma M)}(k)+u^2_{NO3(\Delta\sigma NO2)}(k)+u^2_{NO3(\Delta\sigma SO2)}(k)+u^2_{NO3(\Delta\sigma O2)}(k)\\+u^2_{NO3(Na)}(k)+u^2_{NO3(qNO2)}(k)+u^2_{NO3(qSO2)}(k)\end{array}}$$

2303 **(5.101)**

2304 **Eqs. (5.80)-(5.101)** presented for ozone number density apply similarly to ozone mixing ratio.

2305
2306

6 Propagation of uncertainty specific to temperature retrieval

After applying the signal corrections reviewed in **chapter 4**, the saturation-background-overlap-corrected, merged and smoothed signals S_5 can be used to retrieve temperature using the air number density integration assuming hydrostatic balance and assuming that air is an ideal gas. (**Eq. (1.11)**). The first step is to correct the signals for the solid angle factor $(z\text{-}z_L)^2$. The second step is to correct for atmospheric extinction. The last step is to vertically sum the product of the corrected signal by the gravity acceleration, and integrate this product downward from the top of the profile to obtain temperature. The flowchart of **Figure 6.1** provides a convenient quick-look summary of the data processing, and in particular shows the various locations at which new uncertainty terms are introduced and propagated.

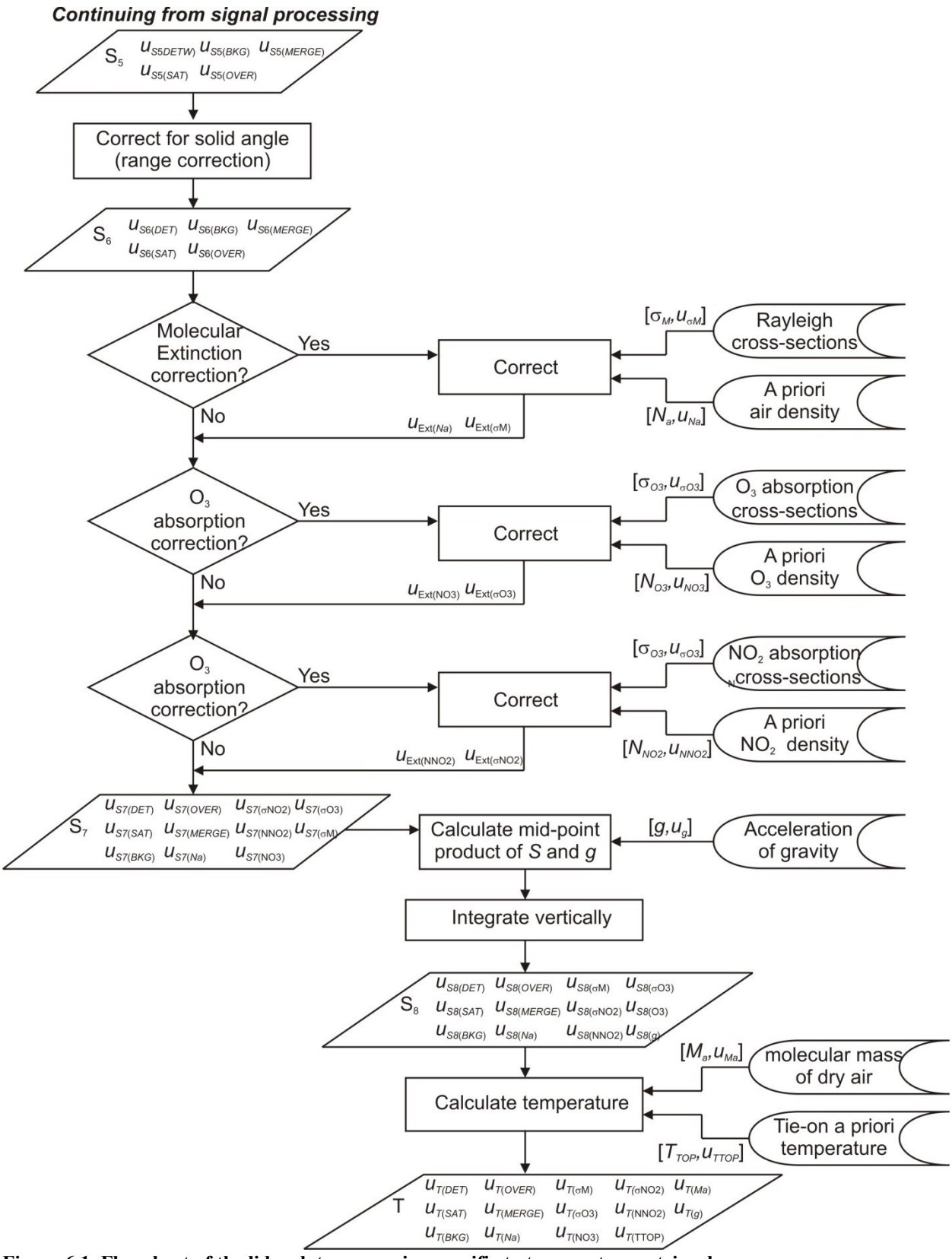


Figure 6.1 Flowchart of the lidar data processing specific to temperature retrieval

6.1 Range (or solid angle) correction

The signal transformation is:

$$S_6(i,k) = \left(z(k) - z_L\right)^2 S_5(i,k) \tag{6.1}$$

z_L is the altitude of the lidar instrument. It is assumed that the emitter and receiver are located at the same altitude. The exact altitude of each data bin k can be determined experimentally, for example by tracking the exact position in the data stream of the laser beam backscattering off the laser room hatch (assuming that the receiver and the transmission of the laser beam in the atmosphere are located in the same room). The time (i.e., altitude) resolution of today's lidar data acquisition hardware is very high (of the order of nanoseconds, i.e., a few meters). The exact altitude of the lidar instrument can also be determined to a precision better than a meter using today's standard geo-positioning methods. For well-designed and well-validated lidar instruments, there is therefore no uncertainty due to the determination of altitude, and the standard uncertainty components introduced earlier can be propagated to the corrected signal S_6 as follows:

$$u_{S6(DET)}(i,k) = \left|\frac{\partial S_6(i,k)}{\partial S_5(i,k)}\right| u_{S6(DET)}(i,k) = \left(z(k) - z_L\right)^2 u_{S5(DET)}(i,k) \tag{6.2}$$

$$u_{S6(SAT)}(i,k) = \left|\frac{\partial S_6(i,k)}{\partial S_5(i,k)}\right| u_{S6(SAT)}(i,k) = \left(z(k) - z_L\right)^2 u_{S5(SAT)}(i,k) \tag{6.3}$$

$$u_{S6(BKG)}(i,k) = \left|\frac{\partial S_6(i,k)}{\partial S_5(i,k)}\right| u_{S6(BKG)}(i,k) = \left(z(k) - z_L\right)^2 u_{S5(BKG)}(i,k) \tag{6.4}$$

$$u_{S6(OVER)}(i,k) = \left|\frac{\partial S_6(i,k)}{\partial S_5(i,k)}\right| u_{S6(OVER)}(i,k) = \left(z(k) - z_L\right)^2 u_{S5(OVER)}(i,k) \tag{6.5}$$

$$u_{S6(MERGE)}(i,k) = \left|\frac{\partial S_6(i,k)}{\partial S_5(i,k)}\right| u_{S6(MERGE)}(i,k) = \left(z(k) - z_L\right)^2 u_{S5(MERGE)}(i,k) \tag{6.6}$$

The combined standard uncertainty on the corrected signal S_6 becomes:

$$u_{S6}(i,k) = \sqrt{u^2_{S6(DET)}(i,k) + u^2_{S6(SAT)}(i,k) + u^2_{S6(BKG)}(i,k) + u^2_{S6(OVER)}(i,k) + u^2_{S6(MERGE)}(i,k)} \tag{6.7}$$

6.2 Extinction correction

For Rayleigh backscatter channels, the emitted and received wavelengths are identical. The discretized version of equation (**Eq. (1.12)**) presented in the introduction can be written:

$$S_7(i,k) = S_6(i,k)\exp\left(\sum_{k'=0}^{k}\left(\begin{array}{l}\left(\sigma_M(\lambda_E)+\sigma_M(\lambda_R)\right)N_a(k')+\left(\sigma_{O3}(\lambda_E,k')+\sigma_{O3}(\lambda_R,k')\right)N_{O3}(k')\\+\left(\sigma_{NO2}(\lambda_E,k')+\sigma_{NO2}(\lambda_R,k')\right)N_{NO2}(k')\end{array}\right)\right) \quad \textbf{(6.8)}$$

For Rayleigh backscatter channels, the emitted wavelength λ_E and the received wavelength λ_R are identical, and **Eq. (6.8)** becomes:

$$S_7(i,k) = S_6(i,k)\exp\left(2\sum_{k'=0}^{k}\left(\sigma_M(\lambda_E)N_a(k')+\sigma_{O3}(\lambda_E,k')N_{O3}(k')+\sigma_{NO2}(\lambda_E,k')N_{NO2}(k')\right)\right) \quad \textbf{(6.9)}$$

The term η present in **Eq. (1.12)** has been removed here. The only altitude-dependent contribution to η is from incomplete overlap and has already been treated in **chapter 4**. Every other contribution to η is constant with altitude and therefore is not needed here as it cancels out when calculating the ratio of the signals at two successive altitude bins as part of the temperature computation. The only atmospheric extinction terms included in the above equation is absorption by ozone and by NO_2.

6.2.1 Uncertainty components propagated from prior signal processing (S_6)

The uncertainty components previously introduced and propagated to S_7 can be written:

$$u_{S7(DET)}(i,k) = \left|\frac{\partial S_7(i,k)}{\partial S_6(i,k)}\right| u_{S6(DET)}(i,k) = \frac{S_7(i,k)}{S_6(i,k)} u_{S6(DET)}(i,k) \quad \textbf{(6.10)}$$

$$u_{S7(SAT)}(i,k) = \left|\frac{\partial S_7(i,k)}{\partial S_6(i,k)}\right| u_{S6(SAT)}(i,k) = \frac{S_7(i,k)}{S_6(i,k)} u_{S6(SAT)}(i,k) \quad \textbf{(6.11)}$$

$$u_{S7(BKG)}(i,k) = \left|\frac{\partial S_7(i,k)}{\partial S_6(i,k)}\right| u_{S6(BKG)}(i,k) = \frac{S_7(i,k)}{S_6(i,k)} u_{S6(BKG)}(i,k) \quad \textbf{(6.12)}$$

$$u_{S7(OVER)}(i,k) = \left|\frac{\partial S_7(i,k)}{\partial S_6(i,k)}\right| u_{S6(OVER)}(i,k) = \frac{S_7(i,k)}{S_6(i,k)} u_{S6(OVER)}(i,k) \quad \textbf{(6.13)}$$

$$u_{S7(MERGE)}(i,k) = \left|\frac{\partial S_7(i,k)}{\partial S_6(i,k)}\right| u_{S6(MERGE)}(i,k) = \frac{S_7(i,k)}{S_6(i,k)} u_{S6(MERGE)}(i,k) \quad \textbf{(6.14)}$$

6.2.2 Uncertainty components due to Rayleigh cross-section

The molecular extinction term comprises the Rayleigh cross-section values $\sigma_M(\lambda_E)$ and $\sigma_M(\lambda_R)$ at the emitted and received wavelengths respectively. The general expression of uncertainty associated with these cross-sections and propagated to the extinction-corrected signal S_7 is:

$$u_{S7(\sigma M)}(i,k) = \sqrt{\sum_{p=E}^{R}\sum_{q=E}^{R} \frac{\partial S_7(i,k)}{\partial \sigma_M(\lambda_p)} \frac{\partial S_7(i,k)}{\partial \sigma_M(\lambda_q)} \chi_{\sigma M,\sigma M}(\lambda_p,\lambda_q)}$$

(6.15)

As reviewed in **appendix D**, the Rayleigh scattering cross-section values are typically computed using an analytical function of wavelength. It is therefore assumed that the cross-section values at the emitted and received wavelengths are fully correlated. The extinction-corrected signal uncertainty associated with the use of Rayleigh extinction cross-sections can therefore be written:

$$u_{S7(\sigma M)}(i,k) = S_7(i,k)\sum_{k'=0}^{k} N_a(k')\left(u_{\sigma M}(\lambda_E) + u_{\sigma M}(\lambda_R)\right)$$

(6.16)

For Rayleigh backscatter channels, the received and emitted wavelengths are identical, and **Eq. (6.16)** can be re-written:

$$u_{S7(\sigma M)}(i,k) = 2S_7(i,k)\sum_{k'=0}^{k} N_a(k')u_{\sigma M}(\lambda_E)$$

(6.17)

6.2.3 Uncertainty components due to ancillary air number density

Assuming that the ancillary air number density values are fully correlated in altitude, the extinction-corrected signal uncertainty associated with the use of ancillary air number density can be written:

$$u_{S7(Na)}(i,k) = \left|\frac{\partial S_7(i,k)}{\partial N_a(k)}\right| u_{Na}(k) = S_7(i,k)\sum_{k'=0}^{k}\left(\sigma_M(\lambda_E) + \sigma_M(\lambda_R)\right)u_{Na}(k')$$

(6.18)

For Rayleigh backscatter channels, this expression becomes:

$$u_{S7(Na)}(i,k) = 2S_7(i,k)\sum_{k'=0}^{k}\sigma_M(\lambda_E)u_{Na}(k')$$

(6.19)

6.2.4 Uncertainty components due to ozone and NO_2 absorption cross-sections

Similarly to molecular extinction, the ozone absorption term comprises the cross-section values $\sigma_{O3}(\lambda_E,k)$ and $\sigma_{O3}(\lambda_R,k)$ at the emitted and received wavelengths respectively. The general expression of uncertainty associated with these ozone absorption cross-sections and propagated to the extinction-corrected signal S_7 is similar to that of **Eq. (6.15)**, but taking into account the altitude dependence of the cross-sections.

For vibrational Raman channels, the emitted and received wavelengths are different. The uncertainty owed to random effects must be propagated to the absorption-corrected signal S_7 using:

$$u_{S7(\sigma O3)}(i,k) = S_7(i,k)\sqrt{\sum_{k'=0}^{k} N_{O3}^2(k')\left(u_{\sigma O3}^2(\lambda_E,k') + u_{\sigma O3}^2(\lambda_R,k')\right)}$$

(6.20)

For Rayleigh scattering channels, the emitted and received wavelengths are identical. The uncertainty owed to random effects must be propagated to the absorption-corrected signal S_7 using:

$$u_{S7(\sigma O3)}(i,k) = 2S_7(i,k)\sqrt{\sum_{k'=0}^{k} N_{O3}^2(k')u_{\sigma O3}^2(\lambda_E,k')}$$

(6.21)

For uncertainty owed to systematic effects, the values of the cross-sections can be assumed fully correlated from one wavelength to another and from one altitude to another. For vibrational Raman backscatter channels, their uncertainty must therefore be propagated to the absorption-corrected signal S_7 using:

$$u_{S7(\sigma O3)}(i,k) = S_7(i,k)\sqrt{\sum_{k'=0}^{k} N_{O3}(k')\left(u_{\sigma O3}(\lambda_E,k') + u_{\sigma O3}(\lambda_R,k')\right)}$$

(6.22)

For Rayleigh backscatter channels, their uncertainty **Eq. (6.22)** simplifies to

$$u_{S7(\sigma O3)}(i,k) = 2S_7(i,k)\sum_{k'=0}^{k} N_{O3}(k')u_{\sigma O3}(\lambda_E,k')$$

(6.23)

For the NO_2 cross-sections, **Eqs. (6.20)-(6.23)** can just be written equivalently. For vibrational Raman channels, the uncertainty owed to random effects are propagated to the absorption-corrected signal S_7 using:

$$u_{S7(\sigma NO2)}(i,k) = S_7(i,k)\sqrt{\sum_{k'=0}^{k} N_{NO2}^2(k')\left(u_{\sigma NO2}^2(\lambda_E,k') + u_{\sigma NO2}^2(\lambda_R,k')\right)}$$

(6.24)

For Rayleigh scattering channels, the uncertainty owed to random effects are propagated to the absorption-corrected signal S_7 using:

$$u_{S7(\sigma NO2)}(i,k) = 2S_7(i,k)\sqrt{\sum_{k'=0}^{k} N_{NO2}^2(k')u_{\sigma NO2}^2(\lambda_E,k')}$$

(6.25)

For uncertainty owed to systematic effects, and vibrational Raman backscatter channels, uncertainty is propagated to the absorption-corrected signal S_7 using:

$$u_{S7(\sigma NO2)}(i,k) = S_7(i,k)\sqrt{\sum_{k'=0}^{k} N_{NO2}(k')\left(u_{\sigma NO2}(\lambda_E,k') + u_{\sigma NO2}(\lambda_R,k')\right)}$$

(6.26)

For Rayleigh backscatter channels, the above equation simplifies to

$$u_{S7(\sigma NO2)}(i,k) = 2S_7(i,k)\sum_{k'=0}^{k} N_{NO2}(k')u_{\sigma NO2}(\lambda_E,k') \quad \textbf{(6.27)}$$

6.2.5 Uncertainty components due to the ozone and NO_2 number densities

The ozone (respectively NO_2) absorption term comprises the sum of ancillary ozone (respectively NO_2) number density values taken at all altitudes from the ground to the altitude considered $z(k)$. The general expression of uncertainty associated with this term and propagated to the absorption-corrected signal S_7 is therefore:

$$u_{S7(X)}(i,k) = \sqrt{\sum_{k'=0}^{k}\sum_{k''=0}^{k} \frac{\partial S_7(i,k)}{\partial N_X(k')}\frac{\partial S_7(i,k)}{\partial N_X(k'')}\chi_{X,X}(k',k'')} \quad \text{with } X = O3,\ NO2 \quad \textbf{(6.28)}$$

Assuming that all values within the same ancillary profile are fully correlated, the above expression can be written similarly for ozone and NO_2:

$$u_{S7(NO3)}(i,k) = S_7(i,k)\sum_{k'=0}^{k}\left(\sigma_{O3}(\lambda_E,k') + \sigma_{O3}(\lambda_R,k')\right)u_{NO3}(k') \quad \textbf{(6.29)}$$

$$u_{S7(NNO2)}(i,k) = S_7(i,k)\sum_{k'=0}^{k}\left(\sigma_{NO2}(\lambda_E,k') + \sigma_{NO2}(\lambda_R,k')\right)u_{NNO2}(k') \quad \textbf{(6.30)}$$

For Rayleigh backscatter channels, they are further simplified to:

$$u_{S7(NO3)}(i,k) = 2S_7(i,k)\sum_{k'=0}^{k}\sigma_{O3}(i,k')u_{NO3}(k') \quad \textbf{(6.31)}$$

$$u_{S7(NNO2)}(i,k) = 2S_7(i,k)\sum_{k'=0}^{k}\sigma_{NO2}(i,k')u_{NNO2}(k') \quad \textbf{(6.32)}$$

The combined standard uncertainty on the extinction-corrected signal S_7 is:

$$u_{S7}(i,k) = \sqrt{\begin{array}{l} u^2_{S7(DET)}(i,k) + u^2_{S7(SAT)}(i,k) + u^2_{S7(BKG)}(i,k) + u^2_{S7(OVER)}(i,k) + u^2_{S7(MERGE)}(i,k) \\ + u^2_{S7(\sigma M)}(i,k) + u^2_{S7(\sigma O3)}(i,k) + u^2_{S7(\sigma NO2)}(i,k) + u^2_{S7(Na)}(i,k) + u^2_{S7(NO3)}(i,k) + u^2_{S7(NNO2)}(i,k) \end{array}} \quad \textbf{(6.33)}$$

6.2.6 Using ancillary ozone and NO_2 mixing ratio instead of number density

Until now, it was assumed that the ancillary dataset used as input quantity for the absorption by ozone (respectively NO_2) was number density N_{O3} (respectively N_{NO2}). The corresponding uncertainty components were introduced and propagated assuming that the ozone and NO_2 number density were uncorrelated with the air number density N_a. When the ancillary datasets used as input quantity is mixing ratio q_{O3} and q_{NO2} instead of number density, **Eq. (6.8)** must be reformulated to take into account the interdependence between the ozone number density and the air number density:

$$S_7(i,k) = S_6(i,k)\exp\left(\sum_{k'=0}^{k}\left(\begin{pmatrix}\sigma_M(\lambda_E)+\sigma_M(\lambda_R)\\ +\left(\sigma_{O3}(\lambda_E,k')+\sigma_{O3}(\lambda_R,k')\right)q_{O3}(k')\\ +\left(\sigma_{NO2}(\lambda_E,k')+\sigma_{NO2}(\lambda_R,k')\right)q_{NO2}(k')\end{pmatrix}N_a(k')\right)\right) \quad \textbf{(6.34)}$$

With this new expression, several uncertainty components propagated to S_7 must be re-written. Uncertainty owed to air number density and propagated to S_7 for vibrational Raman backscatter channels becomes:

$$u_{S7(Na)}(i,k) = S_7(i,k)\sum_{k'=0}^{k}\left(\left(\sigma_M(\lambda_E)+\sigma_M(\lambda_R)\right)+\left(\sigma_{O3}(\lambda_E,k')+\sigma_{O3}(\lambda_R,k')\right)q_{O3}(k')\right)u_{Na}(k') \quad \textbf{(6.35)}$$

For Rayleigh backscatter channels this expression simplifies to:

$$u_{S7(Na)}(i,k) = 2S_7(i,k)\sum_{k'=0}^{k}\left(\sigma_M(\lambda_E)+\sigma_{O3}(\lambda_E)q_{O3}(k')\right)u_{Na}(k') \quad \textbf{(6.36)}$$

Uncertainty owed to the use of ancillary ozone and NO_2 mixing ratio and propagated to S_7 for vibrational Raman backscatter channels becomes

$$u_{S7(qO3)}(i,k) = S_7(i,k)\sum_{k'=0}^{k}\left(\sigma_{O3}(\lambda_E,k')+\sigma_{O3}(\lambda_R,k')\right)N_a(k')u_{qO3}(k') \quad \textbf{(6.37)}$$

$$u_{S7(qNO2)}(i,k) = S_7(i,k)\sum_{k'=0}^{k}\left(\sigma_{NO2}(\lambda_E,k')+\sigma_{NO2}(\lambda_R,k')\right)N_a(k')u_{qNO2}(k') \quad \textbf{(6.38)}$$

For Rayleigh backscatter channels these expressions simplify to:

$$u_{S7(qO3)}(i,k) = 2S_7(i,k)\sum_{k'=0}^{k}\sigma_{O3}(i,k')N_a(k')u_{qO3}(k') \quad \textbf{(6.39)}$$

$$u_{S7(qNO2)}(i,k) = 2S_7(i,k)\sum_{k'=0}^{k}\sigma_{NO2}(i,k')N_a(k')u_{qNO2}(k') \quad \textbf{(6.40)}$$

The combined standard uncertainty on the extinction-corrected signal S_7 becomes:

$$u_{S7}(i,k)=\sqrt{\begin{array}{l}u^2_{S7(DET)}(i,k)+u^2_{S7(SAT)}(i,k)+u^2_{S7(BKG)}(i,k)+u^2_{S7(OVER)}(i,k)+u^2_{S7(MERGE)}(i,k)\\+u^2_{S7(\sigma M)}(i,k)+u^2_{S7(\sigma O3)}(i,k)+u^2_{S7(\sigma NO2)}(i,k)+u^2_{S7(Na)}(i,k)+u^2_{S7(qO3)}(i,k)+u^2_{S7(qNO2)}(i,k)\end{array}} \quad \textbf{(6.41)}$$

6.3 Temperature integration

The discretized version of equation **Eq. (1.11)** presented in the introduction can be written:

$$T(i,k)=\frac{S_7(i,k_{TOP})}{S_7(i,k)}T(i,k_{TOP})+\frac{M_a\delta z}{R_a S_7(i,k)}\sum_{k'=k}^{kTOP-1}\overline{S_7}(i,k')\overline{g}(k') \qquad k<k_{TOP} \quad \textbf{(6.42)}$$

As in **Eq. (1.11)**, the term $\overline{S_7}(k')$ (respectively $\overline{g}(k')$) denotes the mean value of S_7 (respectively g) in the altitude layer comprised between $z(k')$ and $z(k'+1)$.

6.3.1 Uncertainty propagated to the layer-averaged product of S_7 and g

The vertical decrease of the signal S_7 is of exponential form. We can therefore rewrite the layer-averaged value of S_7 as:

$$\overline{S_7}(i,k)=\sqrt{S_7(i,k)S_7(i,k+1)} \quad \textbf{(6.43)}$$

For all uncertainty components X introduced earlier (X=*DET*, *SAT*, *BKG*, *OVER*, *MERGE*, σM, *Na*, $\sigma O3$, *NO*3, $\sigma NO2$, *NNO*2), the general expression of propagation associated with **Eq. (6.43)** is:

$$u_{\overline{S7}(X)}(i,k)=\sqrt{\sum_{k_1=k}^{k+1}\sum_{k_2=k}^{k+1}\frac{\partial\overline{S_7}(i,k)}{\partial S_7(i,k_1)}\frac{\partial\overline{S_7}(i,k)}{\partial S_7(i,k_2)}\chi_{S7,S7}(i,k_1,k_2)} \quad \textbf{(6.44)}$$

The uncertainty due to detection noise can be propagated assuming that the signals are uncorrelated between altitudes:

$$u_{\overline{S7}(DET)}(i,k)=\frac{1}{2}\sqrt{\frac{S_7(i,k+1)}{S_7(i,k)}u^2_{S7(DET)}(i,k')+\frac{S_7(i,k)}{S_7(i,k+1)}u^2_{S7(DET)}(i,k+1)} \quad \textbf{(6.45)}$$

For all other uncertainty components introduced earlier, the signals at neighboring points are highly correlated, and the covariance terms must be taken into account. The uncertainty due to saturation correction can be propagated using:

$$u_{\overline{S7}(SAT)}(i,k)=\frac{\overline{S_7}(i,k)}{2}\left(\frac{u_{S7(SAT)}(i,k)}{S_7(i,k)}+\frac{u_{S7(SAT)}(i,k+1)}{S_7(i,k+1)}\right) \quad \textbf{(6.46)}$$

The same approach can be used for all remaining components:

$$u_{\overline{S7}(BKG)}(i,k) = \frac{\overline{S_7}(i,k)}{2}\left(\frac{u_{S7(BKG)}(i,k)}{S_7(i,k)} + \frac{u_{S7(BKG)}(i,k+1)}{S_7(i,k+1)}\right) \quad \textbf{(6.47)}$$

$$u_{\overline{S7}(OVER)}(i,k) = \frac{\overline{S_7}(i,k)}{2}\left(\frac{u_{S7(OVER)}(i,k)}{S_7(i,k)} + \frac{u_{S7(OVER)}(i,k+1)}{S_7(i,k+1)}\right) \quad \textbf{(6.48)}$$

$$u_{\overline{S7}(MERGE)}(i,k) = \frac{\overline{S_7}(i,k)}{2}\left(\frac{u_{S7(MERGE)}(i,k)}{S_7(i,k)} + \frac{u_{S7(MERGE)}(i,k+1)}{S_7(i,k+1)}\right) \quad \textbf{(6.49)}$$

$$u_{\overline{S7}(\sigma M)}(i,k) = \frac{\overline{S_7}(i,k)}{2}\left(\frac{u_{S7(\sigma M)}(i,k)}{S_7(i,k)} + \frac{u_{S7(\sigma M)}(i,k+1)}{S_7(i,k+1)}\right) \quad \textbf{(6.50)}$$

$$u_{\overline{S7}(Na)}(i,k) = \frac{\overline{S_7}(i,k)}{2}\left(\frac{u_{S7(Na)}(i,k)}{S_7(i,k)} + \frac{u_{S7(Na)}(i,k+1)}{S_7(i,k+1)}\right) \quad \textbf{(6.51)}$$

$$u_{\overline{S7}(\sigma O3)}(i,k) = \frac{\overline{S_7}(i,k)}{2}\left(\frac{u_{S7(\sigma O3)}(i,k)}{S_7(i,k)} + \frac{u_{S7(\sigma O3)}(i,k+1)}{S_7(i,k+1)}\right) \quad \textbf{(6.52)}$$

$$u_{\overline{S7}(NO3)}(i,k) = \frac{\overline{S_7}(i,k)}{2}\left(\frac{u_{S7(NO3)}(i,k)}{S_7(i,k)} + \frac{u_{S7(NO3)}(i,k+1)}{S_7(i,k+1)}\right) \quad \textbf{(6.53)}$$

$$u_{\overline{S7}(qO3)}(i,k) = \frac{\overline{S_7}(i,k)}{2}\left(\frac{u_{S7(qO3)}(i,k)}{S_7(i,k)} + \frac{u_{S7(qO3)}(i,k+1)}{S_7(i,k+1)}\right) \quad \textbf{(6.54)}$$

$$u_{\overline{S7}(\sigma NO2)}(i,k) = \frac{\overline{S_7}(i,k)}{2}\left(\frac{u_{S7(\sigma NO2)}(i,k)}{S_7(i,k)} + \frac{u_{S7(\sigma NO2)}(i,k+1)}{S_7(i,k+1)}\right) \quad \textbf{(6.55)}$$

$$u_{\overline{S7}(NNO2)}(i,k) = \frac{\overline{S_7}(i,k)}{2}\left(\frac{u_{S7(NNO2)}(i,k)}{S_7(i,k)} + \frac{u_{S7(NNO2)}(i,k+1)}{S_7(i,k+1)}\right) \quad \textbf{(6.56)}$$

$$u_{\overline{S7}(qNO2)}(i,k) = \frac{\overline{S_7}(i,k)}{2}\left(\frac{u_{S7(qNO2)}(i,k)}{S_7(i,k)} + \frac{u_{S7(qNO2)}(i,k+1)}{S_7(i,k+1)}\right) \quad \textbf{(6.57)}$$

At small vertical scales, the vertical decrease of the acceleration of gravity g is nearly linear. For the altitude layer comprised between $z(k)$ and $z(k+1)$, we can therefore approximate the height of the local ellipsoid h defined in **Eq.(4.6)** using:

$$\bar{h}(k) = \frac{1}{2}\left(h(k) + h(k+1)\right) \tag{6.58}$$

The discretized, layer-averaged value of the acceleration of gravity g defined by **Eq. (4.6)** can be written:

$$\bar{g}(k) = g_0\left(1 + g_1\bar{h}(k) + g_2\bar{h}^2(k)\right) \tag{6.59}$$

The constants g_0, g_1 and g_2 relate to the Earth's geometry and to the geodetic latitude of the lidar site. They can be calculated using **Eqs. (4.6) and (4.7)** provided in **chapter 4**. If a value of the local ellipsoid height at the lidar site $h(0)$ is not known, we can approximate it to the site's altitude above mean sea level $z(0)$. For all altitude-dependent and latitude-dependent formulations of the acceleration of gravity, the difference between $h(0)$ and $z(0)$ is by far the largest source of error in the computation of the acceleration of gravity. We therefore can define a new uncertainty component u_h associated with the approximation of h. The values of h at neighboring altitudes are fully correlated, and their standard uncertainty can be deduced directly from **Eq. (6.58)**:

$$u_{\bar{h}}(k) = \frac{1}{2}\left(u_h(k) + u_h(k+1)\right) \tag{6.60}$$

The height uncertainty propagated to the layer-averaged acceleration of gravity is:

$$u_{\bar{g}}(k) = \left|\frac{\partial\bar{g}(k)}{\partial\bar{h}(k)}\right| u_{\bar{h}}(k) = g_0\left(g_1 + 2g_2\bar{h}(k)\right)u_{\bar{h}}(k) \tag{6.61}$$

The layer-averaged value of the product of S_7 by g is written:

$$S_8(i,k) = \overline{S_7}(i,k)\bar{g}(k) \tag{6.62}$$

The standard uncertainty component owed to the acceleration of gravity and propagated to S_8 can be written:

$$u_{S8(g)}(i,k) = \overline{S_7}(i,k)u_{\bar{g}}(k) \tag{6.63}$$

The standard uncertainty components previously introduced can be propagated to S_8 using:

$$u_{S8(DET)}(i,k) = \bar{g}(k)u_{\overline{S7}(DET)}(i,k) \tag{6.64}$$

$$u_{S8(SAT)}(i,k) = \bar{g}(k)u_{\overline{S7}(SAT)}(i,k) \tag{6.65}$$

$$u_{S8(BKG)}(i,k) = \bar{g}(k)u_{\overline{S7}(BKG)}(i,k) \tag{6.66}$$

$$u_{S8(OVER)}(i,k) = \bar{g}(k)u_{\overline{S7}(OVER)}(i,k) \quad \textbf{(6.67)}$$

$$u_{S8(MERGE)}(i,k) = \bar{g}(k)u_{\overline{S7}(MERGE)}(i,k) \quad \textbf{(6.68)}$$

$$u_{S8(\sigma M)}(i,k) = \bar{g}(k)u_{\overline{S7}(\sigma M)}(i,k) \quad \textbf{(6.69)}$$

$$u_{S8(Na)}(i,k) = \bar{g}(k)u_{\overline{S7}(Na)}(i,k) \quad \textbf{(6.70)}$$

$$u_{S8(\sigma O3)}(i,k) = \bar{g}(k)u_{\overline{S7}(\sigma O3)}(i,k) \quad \textbf{(6.71)}$$

$$u_{S8(NO3)}(i,k) = \bar{g}(k)u_{\overline{S7}(NO3)}(i,k) \quad \textbf{(6.72)}$$

$$u_{S8(qO3)}(i,k) = \bar{g}(k)u_{\overline{S7}(qO3)}(i,k) \quad \textbf{(6.73)}$$

$$u_{S8(\sigma NO2)}(i,k) = \bar{g}(k)u_{\overline{S7}(\sigma NO2)}(i,k) \quad \textbf{(6.74)}$$

$$u_{S8(NNO2)}(i,k) = \bar{g}(k)u_{\overline{S7}(NNO2)}(i,k) \quad \textbf{(6.75)}$$

$$u_{S8(qNO2)}(i,k) = \bar{g}(k)u_{\overline{S7}(qNO2)}(i,k) \quad \textbf{(6.76)}$$

6.3.2 Uncertainty propagated to the summation term

We can re-write the summation term:

$$S_9(i,k) = \sum_{k'=k}^{kTOP-1} S_8(i,k') \quad \textbf{(6.77)}$$

The standard uncertainty due to detection noise can be propagated to S_9 assuming that the signal values are uncorrelated between neighboring altitude bins:

$$u_{S9(DET)}(i,k) = \sum_{k'=k}^{kTOP-1} u^2_{S8(DET)}(i,k') \quad \textbf{(6.78)}$$

All other uncertainty components can be propagated assuming full correlation between the neighboring points. The standard uncertainty associated with saturation correction can be written:

$$u_{S9(SAT)}(i,k) = \sum_{k'=k}^{kTOP-1} u_{S8(SAT)}(i,k') \tag{6.79}$$

The same approach can be used for the other components:

$$u_{S9(BKG)}(i,k) = \sum_{k'=k}^{kTOP-1} u_{S8(BKG))}(i,k') \tag{6.80}$$

$$u_{S9(OVER)}(i,k) = \sum_{k'=k}^{kTOP-1} u_{S8(OVER)}(i,k') \tag{6.81}$$

$$u_{S9(MERGE)}(i,k) = \sum_{k'=k}^{kTOP-1} u_{S8(MERGE)}(i,k') \tag{6.82}$$

$$u_{S9(\sigma M)}(i,k) = \sum_{k'=k}^{kTOP-1} u_{S8(\sigma M)}(i,k') \tag{6.83}$$

$$u_{S9(Na)}(i,k) = \sum_{k'=k}^{kTOP-1} u_{S8(Na)}(i,k') \tag{6.84}$$

$$u_{S9(\sigma O3)}(i,k) = \sum_{k'=k}^{kTOP-1} u_{S8(\sigma O3)}(i,k') \tag{6.85}$$

$$u_{S9(NO3)}(i,k) = \sum_{k'=k}^{kTOP-1} u_{S8(NO3)}(i,k') \tag{6.86}$$

$$u_{S9(qO3)}(i,k) = \sum_{k'=k}^{kTOP-1} u_{S8(qO3)}(i,k') \tag{6.87}$$

$$u_{S9(\sigma NO2)}(i,k) = \sum_{k'=k}^{kTOP-1} u_{S8(\sigma NO2)}(i,k') \tag{6.88}$$

$$u_{S9(NNO2)}(i,k) = \sum_{k'=k}^{kTOP-1} u_{S8(NNO2)}(i,k') \tag{6.89}$$

$$u_{S9(qNO2)}(i,k) = \sum_{k'=k}^{kTOP-1} u_{S8(qNO2)}(i,k') \tag{6.90}$$

$$u_{S9(g)}(i,k) = \sum_{k'=k}^{kTOP-1} u_{S8(g)}(i,k')$$

2645 **(6.91)**

2646

2647 ### 6.3.3 Temperature combined standard uncertainty

2648 After the transformations of the signals S_7 leading to the computation of the summed term S_9, we
2649 can re-write the temperature integration equation:

2650 $$T(i,k) = \frac{1}{S_7(i,k)}\left(S_7(i,k_{TOP})T(i,k_{TOP}) + \frac{M_a \delta z}{R_a} S_9(i,k) \right) \qquad k < k_{TOP}$$

2651 **(6.92)**

2652 This equation shows that an ancillary temperature $T(i,k_{TOP}) = T_a(k_{TOP})$ is needed to initialize the
2653 profile at the top. We therefore introduce the uncertainty associated with the ancillary
2654 temperature u_{TTOP}, and we propagate it to the retrieved temperature using:

2655 $$u_{T(TTOP)}(i,k) = \frac{S_7(i,k_{TOP})}{S_7(i,k)} u_{TTOP}$$

2656 **(6.93)**

2657 We also introduce the uncertainty associated with the molecular mass of dry air u_{Ma}, and we
2658 propagate it to the retrieved temperature using:

2659 $$u_{T(Ma)}(i,k) = \frac{\delta z}{R_a} \frac{S_9(i,k)}{S_7(i,k)} u_{Ma}$$

2660 **(6.94)**

2661 The temperature standard uncertainty due to detection noise can be written assuming that none of
2662 the input quantities $S_7(i,k)$, $S_7(i,k_{TOP})$, $T_a(k_{TOP})$ and $S_9(i,k)$ are correlated:

2663 $$u_{T(DET)}(i,k) = \frac{1}{S_7(i,k)T(i,k)} \sqrt{T^2(i,k)u^2_{S7(DET)}(i,k) + T_a^2(k_{TOP})u^2_{S7(DET)}(i,k_{TOP}) + \left(\frac{M_a \delta z}{R_a}\right)^2 u^2_{S9(DET)}(i,k)}$$

2664 **(6.95)**

2665 For all other uncertainty components, the input quantities $S_7(i,k)$, $S_7(i,k_{TOP})$, T_{TOP}, and $S_9(i,k)$ are
2666 all correlated. The temperature uncertainty due to saturation correction can therefore be written:

2667 $$u_{T(SAT)}(i,k) = \frac{1}{S_7(i,k)} \left| T(i,k)u_{S7(SAT)}(i,k) - T_a(k_{TOP})u_{S7(SAT)}(i,k_{TOP}) - \frac{M_a \delta z}{R_a} u_{S9(SAT)}(i,k) \right|$$

2668 **(6.96)**

2669 The same approach can be used for all other uncertainty components:

2670 $$u_{T(BKG)}(i,k) = \frac{1}{S_7(i,k)} \left| T(i,k)u_{S7(BKG)}(i,k) - T_a(k_{TOP})u_{S7(BKG)}(i,k_{TOP}) - \frac{M_a \delta z}{R_a} u_{S9(BKG)}(i,k) \right|$$

2671 **(6.97)**

2672 $$u_{T(OVER)}(i,k) = \frac{1}{S_7(i,k)} \left| T(i,k)u_{S7(OVER)}(i,k) - T_a(k_{TOP})u_{S7(OVER)}(i,k_{TOP}) - \frac{M_a \delta z}{R_a} u_{S9(OVER)}(i,k) \right|$$

2673 **(6.98)**

$$u_{T(MERGE)}(i,k)=\frac{1}{S_7(i,k)}\left|T(i,k)u_{S7(MERGE)}(i,k)-T_a(k_{TOP})u_{S7(MERGE)}(i,k_{TOP})-\frac{M_a\delta z}{R_a}u_{S9(MERGE)}(i,k)\right| \quad \textbf{(6.99)}$$

$$u_{T(\sigma M)}(i,k)=\frac{1}{S_7(i,k)}\left|T(i,k)u_{S7(\sigma M)}(i,k)-T_a(k_{TOP})u_{S7(\sigma M)}(i,k_{TOP})-\frac{M_a\delta z}{R_a}u_{S9(\sigma M)}(i,k)\right| \quad \textbf{(6.100)}$$

$$u_{T(Na)}(i,k)=\frac{1}{S_7(i,k)}\left|T(i,k)u_{S7(Na)}(i,k)-T_a(k_{TOP})u_{S7(Na)}(i,k_{TOP})-\frac{M_a\delta z}{R_a}u_{S9(Na)}(i,k)\right| \quad \textbf{(6.101)}$$

$$u_{T(\sigma O3)}(i,k)=\frac{1}{S_7(i,k)}\left|T(i,k)u_{S7(\sigma O3)}(i,k)-T_a(k_{TOP})u_{S7(\sigma O3)}(i,k_{TOP})-\frac{M_a\delta z}{R_a}u_{S9(\sigma O3)}(i,k)\right| \quad \textbf{(6.102)}$$

$$u_{T(NO3)}(i,k)=\frac{1}{S_7(i,k)}\left|T(i,k)u_{S7(NO3)}(i,k)-T_a(k_{TOP})u_{S7(NO3)}(i,k_{TOP})-\frac{M_a\delta z}{R_a}u_{S9(NO3)}(i,k)\right| \quad \textbf{(6.103)}$$

$$u_{T(qO3)}(i,k)=\frac{1}{S_7(i,k)}\left|T(i,k)u_{S7(qO3)}(i,k)-T_a(k_{TOP})u_{S7(qO3)}(i,k_{TOP})-\frac{M_a\delta z}{R_a}u_{S9(qO3)}(i,k)\right| \quad \textbf{(6.104)}$$

$$u_{T(\sigma NO2)}(i,k)=\frac{1}{S_7(i,k)}\left|T(i,k)u_{S7(\sigma NO2)}(i,k)-T_a(k_{TOP})u_{S7(\sigma NO2)}(i,k_{TOP})-\frac{M_a\delta z}{R_a}u_{S9(\sigma NO2)}(i,k)\right| \quad \textbf{(6.105)}$$

$$u_{T(NNO2)}(i,k)=\frac{1}{S_7(i,k)}\left|T(i,k)u_{S7(NNO2)}(i,k)-T_a(k_{TOP})u_{S7(NNO2)}(i,k_{TOP})-\frac{M_a\delta z}{R_a}u_{S9(NNO2)}(i,k)\right| \quad \textbf{(6.106)}$$

$$u_{T(qNO2)}(i,k)=\frac{1}{S_7(i,k)}\left|T(i,k)u_{S7(qNO2)}(i,k)-T_a(k_{TOP})u_{S7(qNO2)}(i,k_{TOP})-\frac{M_a\delta z}{R_a}u_{S9(qNO2)}(i,k)\right| \quad \textbf{(6.107)}$$

$$u_{T(g)}(i,k)=\frac{1}{S_7(i,k)}\frac{M_a\delta z}{R_a}u_{S9(g)}(i,k) \quad \textbf{(6.108)}$$

If using ozone and NO_2 number density as input quantities for absorption, the temperature combined standard uncertainty can be written:

$$u_T(i,k)=\sqrt{\begin{aligned}&u^2_{T(RAW)}(i,k)+u^2_{T(SAT)}(i,k)+u^2_{T(OVER)}(i,k)+u^2_{T(MERGE)}(i,k)\\&+u^2_{T(\sigma M)}(i,k)+u^2_{T(\sigma O3)}(i,k)+u^2_{T(\sigma NO2)}(i,k)\\&+u^2_{T(Na)}(i,k)+u^2_{T(NO3)}(i,k)+u^2_{T(NNO2)}(i,k)+u^2_{T(g)}(i,k)+u^2_{T(TTOP)}(i,k)+u^2_{T(Ma)}(i,k)\end{aligned}} \quad \textbf{(6.109)}$$

If using ozone and NO_2 mixing ratio as input quantities for absorption, the temperature combined standard uncertainty can be written:

$$u_T(i,k) = \sqrt{\begin{aligned} &u^2_{T(RAW)}(i,k) + u^2_{T(SAT)}(i,k) + u^2_{T(OVER)}(i,k) + u^2_{T(MERGE)}(i,k) \\ &+ u^2_{T(\sigma M)}(i,k) + u^2_{T(\sigma O3)}(i,k) + u^2_{T(\sigma NO2)}(i,k) \\ &+ u^2_{T(Na)}(i,k) + u^2_{T(qO3)}(i,k) + u^2_{T(qNO2)}(i,k) + u^2_{T(g)}(i,k) + u^2_{T(TTOP)}(i,k) + u^2_{T(Ma)}(i,k) \end{aligned}} \quad \textbf{(6.110)}$$

Note that at the tie-on altitude $z(k_{TOP})$, all uncertainty components should be set to zero except uncertainty owed to the ancillary temperature $u_{T(TTOP)}$.

An example of a complete uncertainty budget for the JPL temperature lidar at Mauna Loa, Hawaii, is provided for reference in figure **Figure 6.2**. On this figure we show all the uncertainty components previously introduced except those owed to overlap correction and signal merging, which are not included in the JPL retrieval at this stage of processing.

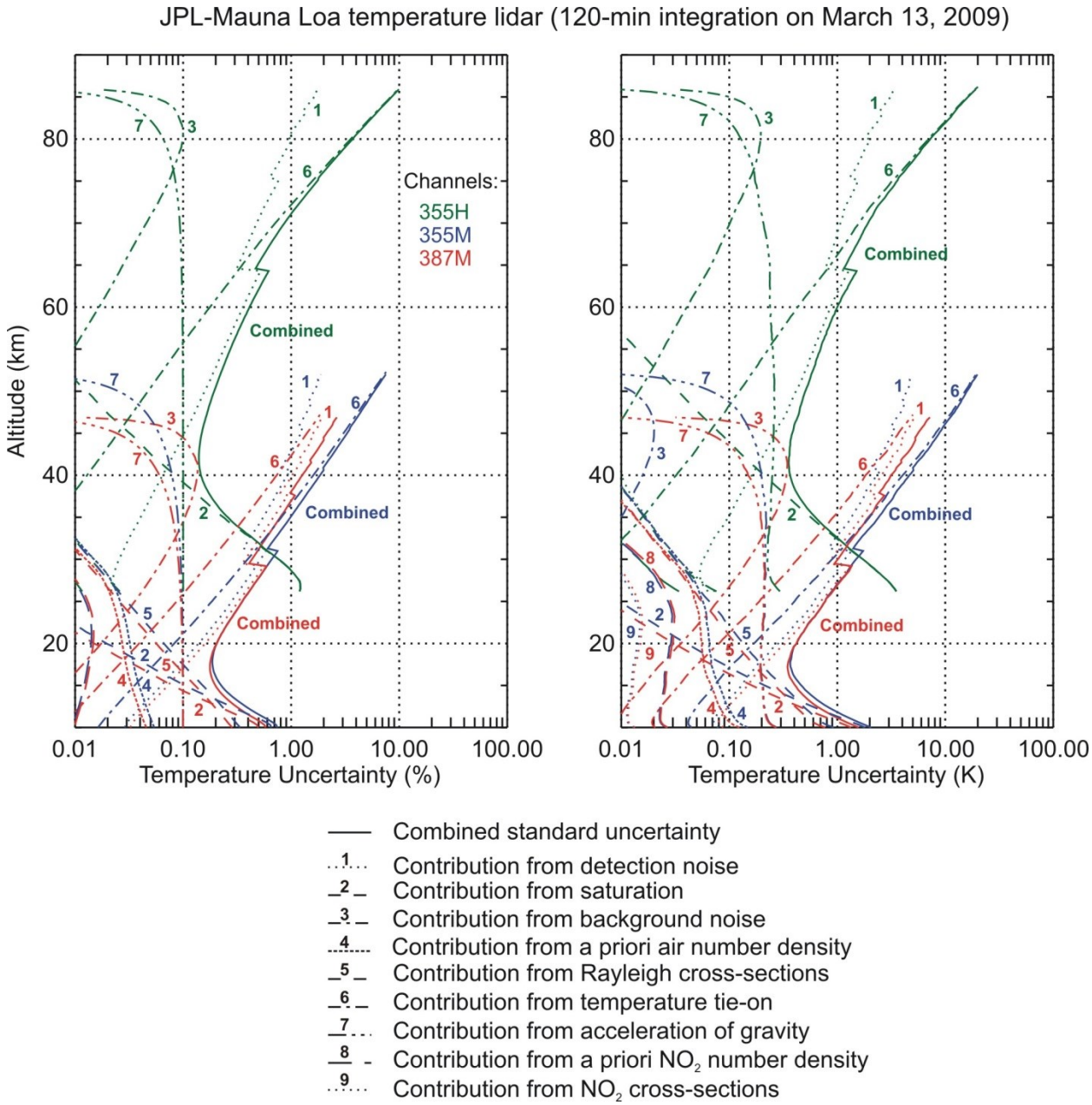


Figure 6.2 Example of full uncertainty budget for the JPL-Mauna Loa Observatory temperature lidar (data taken during 2 hours on March 13, 2009), as computed using the present guidelines

6.3.4 Merging the temperature profiles from multiple channels into one profile

The approach is identical to that presented in **section 4.3.2** for merging the signals. Merging a high intensity channel i_H with a low-intensity channel i_L together can be written:

$$T(k) = m(k)T(i_H,k) + (1-m(k))T(i_L,k) \qquad 0 < m < 1 \qquad \textbf{(6.111)}$$

The uncertainty due to detection noise can be combined assuming that the signals from the high-intensity and low-intensity channels are independent

$$u_{T(DET)}(k) = \sqrt{m^2(k)u^2_{T(DET)}(i_H,k) + (1-m(k))^2 u^2_{T(DET)}(i_L,k)} \quad \textbf{(6.112)}$$

Assuming that the extinction correction of the high- and low-intensity channels' signals are made consistently (i.e., same datasets used for the corrections of the low-intensity and high-intensity channels), the uncertainty components for the extinction correction terms can be propagated as follows:

$$u_{T(\sigma M)}(k) = m(k)u_{T(\sigma M)}(i_H,k) + (1-m(k))u_{T(\sigma M)}(i_L,k) \quad \textbf{(6.113)}$$

$$u_{T(Na)}(k) = m(k)u_{T(Na)}(i_H,k) + (1-m(k))u_{T(Na)}(i_L,k) \quad \textbf{(6.114)}$$

$$u_{T(\sigma O3)}(k) = m(k)u_{T(\sigma O3)}(i_H,k) + (1-m(k))u_{T(\sigma O3)}(i_L,k) \quad \textbf{(6.115)}$$

$$u_{T(NO3)}(k) = m(k)u_{T(NO3)}(i_H,k) + (1-m(k))u_{T(NO3)}(i_L,k) \quad \textbf{(6.116)}$$

$$u_{T(qO3)}(k) = m(k)u_{T(qO3)}(i_H,k) + (1-m(k))u_{T(qO3)}(i_L,k) \quad \textbf{(6.117)}$$

$$u_{T(\sigma NO2)}(k) = m(k)u_{T(\sigma NO2)}(i_H,k) + (1-m(k))u_{T(\sigma NO2)}(i_L,k) \quad \textbf{(6.118)}$$

$$u_{T(NNO2)}(k) = m(k)u_{T(NNO2)}(i_H,k) + (1-m(k))u_{T(NNO2)}(i_L,k) \quad \textbf{(6.119)}$$

$$u_{T(qNO2)}(k) = m(k)u_{T(qNO2)}(i_H,k) + (1-m(k))u_{T(qNO2)}(i_L,k) \quad \textbf{(6.120)}$$

$$u_{T(TTOP)}(k) = m(k)u_{T(TTOP)}(i_H,k) + (1-m(k))u_{T(TTOP)}(i_L,k) \quad \textbf{(6.121)}$$

$$u_{T(g)}(k) = m(k)u_{T(g)}(i_H,k) + (1-m(k))u_{T(g)}(i_L,k) \quad \textbf{(6.122)}$$

$$u_{T(Ma)}(k) = m(k)u_{T(Ma)}(i_H,k) + (1-m(k))u_{T(Ma)}(i_L,k) \quad \textbf{(6.123)}$$

For all other uncertainty components previously introduced, the signals in the low- and high intensity channels may or may not be considered correlated, depending on the design of the instrument and on the consistency of the data processing. If no hardware is shared, the remaining uncertainty components are propagated as follows:

$$u_{T(SAT)}(k) = \sqrt{m^2(k)u^2_{T(SAT)}(i_H,k) + (1-m(k))^2 u^2_{T(SAT)}(i_L,k)} \quad \textbf{(6.124)}$$

$$u_{T(BKG)}(k) = \sqrt{m^2(k)u^2_{T(BKG)}(i_H,k) + (1-m(k))^2 u^2_{T(BKG)}(i_L,k)}$$ **(6.125)**

$$u_{T(OVER)}(k) = \sqrt{m^2(k)u^2_{T(OVER)}(i_H,k) + (1-m(k))^2 u^2_{T(OVER)}(i_L,k)}$$ **(6.126)**

$$u_{T(MERGE)}(k) = \sqrt{m^2(k)u^2_{T(MERGE)}(i_H,k) + (1-m(k))^2 u^2_{T(MERGE)}(i_L,k)}$$ **(6.127)**

If the same hardware is shared, and the data processing is assumed to be consistent for both channels, the remaining uncertainty components are propagated as follows:

$$u_{T(SAT)}(k) = m(k)u_{T(SAT)}(i_H,k) + (1-m(k))u_{T(SAT)}(i_L,k)$$ **(6.128)**

$$u_{T(BKG)}(k) = m(k)u_{T(BKG)}(i_H,k) + (1-m(k))u_{T(BKG)}(i_L,k)$$ **(6.129)**

$$u_{T(OVER)}(k) = m(k)u_{T(OVER)}(i_H,k) + (1-m(k))u_{T(OVER)}(i_L,k)$$ **(6.130)**

$$u_{T(MERGE)}(k) = m(k)u_{T(MERGE)}(i_H,k) + (1-m(k))u_{T(MERGE)}(i_L,k)$$ **(6.131)**

If using number density in the absorption terms, the merged temperature profile combined standard uncertainty can be written:

$$u_T(k) = \sqrt{\begin{aligned} &u^2_{T(DET)}(k) + u^2_{T(SAT)}(k) + u^2_{T(BKG)}(k) + u^2_{T(OVER)}(k) + u^2_{T(MERGE)}(k) \\ &+ u^2_{T(\sigma M)}(k) + u^2_{T(\sigma O3)}(k) + u^2_{T(\sigma NO2)}(k) + u^2_{T(Na)}(k) + u^2_{T(NO3)}(k) + u^2_{T(NNO2)}(k) \\ &+ u^2_{T(g)}(k) + u^2_{T(Ma)}(k) + u^2_{T(TTOP)}(k) \end{aligned}}$$ **(6.132)**

If using mixing ratio in the absorption terms, the merged temperature profile combined standard uncertainty can be written:

$$u_T(k) = \sqrt{\begin{aligned} &u^2_{T(DET)}(k) + u^2_{T(SAT)}(k) + u^2_{T(BKG)}(k) + u^2_{T(OVER)}(k) + u^2_{T(MERGE)}(k) \\ &+ u^2_{T(\sigma M)}(k) + u^2_{T(\sigma O3)}(k) + u^2_{T(\sigma NO2)}(k) + u^2_{T(Na)}(k) + u^2_{T(qO3)}(k) + u^2_{T(qNO2)}(k) \\ &+ u^2_{T(g)}(k) + u^2_{T(Ma)}(k) + u^2_{T(TTOP)}(k) \end{aligned}}$$ **(6.133)**

6.3.5 Derived number density and pressure profiles and their uncertainty

In absence of particulate backscatter and extinction, a pressure profile can be derived after normalizing the corrected lidar signal to an ancillary value of air number density. Providing lidar-derived air number density and pressure profiles together with temperature is often useful, for example when validation of pressure-based satellite measurements is needed. The air number

density and pressure profiles can be derived from the lidar signals and temperature profiles using:

$$N(i,k) = \frac{S_7(i,k)}{S_7(i,k_N)} N_a(k_N) \tag{6.134}$$

$$p(i,k) = \frac{S_7(i,k)}{S_7(i,k_N)} \frac{p_a(k_N)}{T_a(k_N)} T(i,k) \tag{6.135}$$

where $N_a(k_N)$, $p_a(k_N)$ and $T_a(k_N)$ are the ancillary air number density, pressure and temperature values at the lidar signal normalization altitude $z(k_N)$.

A new uncertainty component owed to the normalization to ancillary air number density needs to be introduced. Though the ancillary air number density had been introduced earlier (in the extinction correction), its degree of correlation with the signal S_7 is low. The uncertainty due to normalization propagated to the lidar-derived air number density can be written:

$$u_{N(NORM)}(i,k) = \frac{S_7(i,k)}{S_7(i,k_N)} u_{Na}(k_N) \tag{6.136}$$

The air number density standard uncertainty due to detection noise can be written:

$$u_{N(DET)}(i,k) = N(i,k)\sqrt{\frac{u^2_{S7(DET)}(i,k)}{S_7^2(i,k)} + \frac{u^2_{S7(DET)}(i,k_N)}{S_7^2(i,k_N)}} \qquad \text{if } k \neq k_N \tag{6.137}$$

$$u_{N(DET)}(i,k) = 0 \qquad \text{if } k = k_N \tag{6.138}$$

For all other uncertainty components introduced earlier, the signals at altitude bin k and that at altitude k_N are correlated, and the covariance terms must be taken into account. The uncertainty due to saturation correction can be propagated using:

$$u_{N(SAT)}(i,k) = \frac{1}{S_7(i,k_N)} \left| N_a(i,k_N) u_{S7(SAT)}(i,k) - N(i,k) u_{S7(SAT)}(i,k_N) \right| \tag{6.139}$$

The same approach can be used for all remaining uncertainty components:

$$u_{N(SAT)}(i,k) = \frac{1}{S_7(i,k_N)} \left| N_a(i,k_N) u_{S7(SAT)}(i,k) - N(i,k) u_{S7(SAT)}(i,k_N) \right| \tag{6.140}$$

$$u_{N(BKG)}(i,k) = \frac{1}{S_7(i,k_N)} \left| N_a(i,k_N) u_{S7(BKG)}(i,k) - N(i,k) u_{S7(BKG)}(i,k_N) \right| \tag{6.141}$$

$$u_{N(OVER)}(i,k) = \frac{1}{S_7(i,k_N)} \left| N_a(i,k_N) u_{S7(OVER)}(i,k) - N(i,k) u_{S7(OVER)}(i,k_N) \right|$$

$$u_{N(MERGE)}(i,k) = \frac{1}{S_7(i,k_N)}\left|N_a(i,k_N)u_{S7(MERGE)}(i,k) - N(i,k)u_{S7(MERGE)}(i,k_N)\right| \quad \textbf{(6.142)}$$

$$u_{N(\sigma M)}(i,k) = \frac{1}{S_7(i,k_N)}\left|N_a(i,k_N)u_{S7(\sigma M)}(i,k) - N(i,k)u_{S7(\sigma M)}(i,k_N)\right| \quad \textbf{(6.143)}$$

$$u_{N(Na)}(i,k) = \frac{1}{S_7(i,k_N)}\left|N_a(i,k_N)u_{S7(Na)}(i,k) - N(i,k)u_{S7(Na)}(i,k_N)\right| \quad \textbf{(6.144)}$$

$$u_{N(\sigma O3)}(i,k) = \frac{1}{S_7(i,k_N)}\left|N_a(i,k_N)u_{S7(\sigma O3)}(i,k) - N(i,k)u_{S7(\sigma O3)}(i,k_N)\right| \quad \textbf{(6.145)}$$

$$u_{N(NO3)}(i,k) = \frac{1}{S_7(i,k_N)}\left|N_a(i,k_N)u_{S7(NO3)}(i,k) - N(i,k)u_{S7(NO3)}(i,k_N)\right| \quad \textbf{(6.146)}$$

$$u_{N(qO3)}(i,k) = \frac{1}{S_7(i,k_N)}\left|N_a(i,k_N)u_{S7(qO3)}(i,k) - N(i,k)u_{S7(qO3)}(i,k_N)\right| \quad \textbf{(6.147)}$$

$$u_{N(\sigma NO2)}(i,k) = \frac{1}{S_7(i,k_N)}\left|N_a(i,k_N)u_{S7(\sigma NO2)}(i,k) - N(i,k)u_{S7(\sigma NO2)}(i,k_N)\right| \quad \textbf{(6.148)}$$

$$u_{N(NO2)}(i,k) = \frac{1}{S_7(i,k_N)}\left|N_a(i,k_N)u_{S7(NO2)}(i,k) - N(i,k)u_{S7(NO2)}(i,k_N)\right| \quad \textbf{(6.149)}$$

$$u_{N(qNO2)}(i,k) = \frac{1}{S_7(i,k_N)}\left|N_a(i,k_N)u_{S7(qNO2)}(i,k) - N(i,k)u_{S7(qNO2)}(i,k_N)\right| \quad \textbf{(6.150)}$$

$$u_{N(NNO2)}(i,k) = \frac{1}{S_7(i,k_N)}\left|N_a(i,k_N)u_{S7(NNO2)}(i,k) - N(i,k)u_{S7(NNO2)}(i,k_N)\right| \quad \textbf{(6.151)}$$

If using number density for absorption terms, the lidar-derived air number density combined standard uncertainty is:

$$u_N(i,k) = \sqrt{\begin{aligned}&u^2_{N(NORM)}(i,k) + u^2_{N(DET)}(i,k) + u^2_{N(SAT)}(i,k) + u^2_{N(OVER)}(i,k) + u^2_{N(MERGE)}(i,k) \\ &+ u^2_{N(\sigma M)}(i,k) + u^2_{N(\sigma O3)}(i,k) + u^2_{N(\sigma NO2)}(i,k) + u^2_{N(Na)}(i,k) + u^2_{N(NO3)}(i,k) + u^2_{N(NNO2)}(i,k)\end{aligned}} \quad \textbf{(6.152)}$$

If using mixing ratio for absorption terms, the lidar-derived air number density combined standard uncertainty is:

$$u_N(i,k) = \sqrt{\begin{aligned}&u^2_{N(NORM)}(i,k) + u^2_{N(DET)}(i,k) + u^2_{N(SAT)}(i,k) + u^2_{N(OVER)}(i,k) + u^2_{N(MERGE)}(i,k) \\ &+ u^2_{N(\sigma M)}(i,k) + u^2_{N(\sigma O3)}(i,k) + u^2_{N(\sigma NO2)}(i,k) + u^2_{N(Na)}(i,k) + u^2_{N(qO3)}(i,k) + u^2_{N(qNO2)}(i,k)\end{aligned}} \quad \textbf{(6.153)}$$

The lidar-derived pressure profiles can be re-written:

$$p(i,k) = N(i,k)k_B T(i,k) \tag{6.154}$$

The lidar-derived air number density and temperature profiles are correlated. The covariance terms must be taken into account. The same approach can be used for all uncertainty components propagated to the lidar-derived pressure profile:

$$u_{p(DET)}(i,k) = k_B\left(N(i,k)u_{T(DET)}(i,k) + T(i,k)u_{N(DET)}(i,k)\right) \tag{6.155}$$

$$u_{p(SAT}(i,k) = k_B\left(N(i,k)u_{T(SAT)}(i,k) + T(i,k)u_{N(SAT)}(i,k)\right) \tag{6.156}$$

$$u_{p(BKG)}(i,k) = k_B\left(N(i,k)u_{T(BKG)}(i,k) + T(i,k)u_{N(BKG)}(i,k)\right) \tag{6.157}$$

$$u_{p(OVER)}(i,k) = k_B\left(N(i,k)u_{T(OVER)}(i,k) + T(i,k)u_{N(OVER)}(i,k)\right) \tag{6.158}$$

$$u_{p(MERGE)}(i,k) = k_B\left(N(i,k)u_{T(MERGE)}(i,k) + T(i,k)u_{N(MERGE)}(i,k)\right) \tag{6.159}$$

$$u_{p(\sigma M)}(i,k) = k_B\left(N(i,k)u_{T(\sigma M)}(i,k) + T(i,k)u_{N(\sigma M)}(i,k)\right) \tag{6.160}$$

$$u_{p(Na)}(i,k) = k_B\left(N(i,k)u_{T(Na)}(i,k) + T(i,k)u_{N(Na)}(i,k)\right) \tag{6.161}$$

$$u_{p(\sigma O3)}(i,k) = k_B\left(N(i,k)u_{T(\sigma O3)}(i,k) + T(i,k)u_{N(\sigma O3)}(i,k)\right) \tag{6.162}$$

$$u_{p(NO3)}(i,k) = k_B\left(N(i,k)u_{T(NO3)}(i,k) + T(i,k)u_{N(NO3)}(i,k)\right) \tag{6.163}$$

$$u_{p(qO3)}(i,k) = k_B\left(N(i,k)u_{T(qO3)}(i,k) + T(i,k)u_{N(qO3)}(i,k)\right) \tag{6.164}$$

$$u_{p(\sigma NO2)}(i,k) = k_B\left(N(i,k)u_{T(\sigma NO2)}(i,k) + T(i,k)u_{N(\sigma NO2)}(i,k)\right) \tag{6.165}$$

$$u_{p(NNO2)}(i,k) = k_B\left(N(i,k)u_{T(NNO2)}(i,k) + T(i,k)u_{N(NNO2)}(i,k)\right) \tag{6.166}$$

$$u_{p(qNO2)}(i,k) = k_B\left(N(i,k)u_{T(qNO2)}(i,k) + T(i,k)u_{N(qNO2)}(i,k)\right) \tag{6.167}$$

$$u_{p(NORM)}(i,k) = k_B T(i,k)u_{N(NORM)}(i,k) \tag{6.168}$$

If using number density in the absorption terms, the lidar-derived pressure combined standard uncertainty is:

$$u_p(i,k) = \sqrt{\begin{array}{l} u^2_{p(NORM)}(i,k) + u^2_{p(DET)}(i,k) + u^2_{p(SAT)}(i,k) + u^2_{p(OVER)}(i,k) + u^2_{p(MERGE)}(i,k) \\ + u^2_{p(\sigma M)}(i,k) + u^2_{p(\sigma O3)}(i,k) + u^2_{p(\sigma NO2)}(i,k) + u^2_{p(Na)}(i,k) + u^2_{p(NO3)}(i,k) + u^2_{p(NNO2)}(i,k) \end{array}} \quad \textbf{(6.169)}$$

If using mixing ratio in the absorption terms, the lidar-derived pressure combined standard uncertainty is:

$$u_p(i,k) = \sqrt{\begin{array}{l} u^2_{p(NORM)}(i,k) + u^2_{p(DET)}(i,k) + u^2_{p(SAT)}(i,k) + u^2_{p(OVER)}(i,k) + u^2_{p(MERGE)}(i,k) \\ + u^2_{p(\sigma M)}(i,k) + u^2_{p(\sigma O3)}(i,k) + u^2_{p(\sigma NO2)}(i,k) + u^2_{p(Na)}(i,k) + u^2_{p(qO3)}(i,k) + u^2_{p(qNO2)}(i,k) \end{array}} \quad \textbf{(6.170)}$$

APPENDICES

A Quantitative validation of uncertainty using Monte Carlo experiments

In **chapter 2**, we introduced the metrological concept of measurement model $Y=f(X_1,X_2,\ldots X_N)$, the output quantity Y being a function of the input quantities X_i, i=1,N. In **chapter 3**, we adapted this concept to the lidar measurement of ozone and temperature, and more specifically we proposed to split the measurement model into multiple sub-models through which the input quantities' individual uncertainties can be propagated in parallel until the final product ozone or temperature is obtained. In **chapters 4-6**, we provided expressions of these sub-models and we provided the corresponding propagation expressions for each uncertainty component introduced in the ozone and temperature lidar data processing chain. In the present **appendix A**, we provide an overview of the numerical tools used to simulate and analyze raw lidar signals, and we describe Monte Carlo experiments which, when used with simulated lidar signals, allow the quantification of each uncertainty component propagated to ozone and temperature in the presence of correlated variables. This exercise's objective was not to estimate the magnitude of each uncertainty contribution, but to verify that the expressions used in **chapters 4-6** for the propagation of uncertainty are correct. The quantitative estimates of the input quantities' uncertainty are in many cases arbitrary, yet realistic in order to highlight uncertainty sources that can be neglected in typical ozone or temperature retrievals and those that cannot.

The tools described here comprise a "forward model" which produces simulated lidar signals, and an "inverse model" which analyzes these simulated signals and retrieves ozone and temperature. Their operating principle is as follows: we start from a "known" atmospheric state (referred to as the "true" profile thereafter for brevity), and a set of known instrumental parameters characterizing typical ozone and temperature lidar systems found in NDACC. Using the "true" atmospheric state and the instrumental parameters, we design a measurement model (the "forward model") to simulate the raw lidar signals as if they were acquired by this instrument. We then analyze the simulated signals using a retrieval model (the "inverse model"), and we compare the retrieved ozone (respectively temperature) profile to the true ozone (respectively temperature) profile. When the same instrumental and retrieval parameters are used in the inverse and forward models, the retrieved and true profiles should match perfectly. Once the consistency of the inverse and forward models has been verified (i.e., perfect match of the "true" and retrieved profiles), we can decide to vary any of the instrumental and/or retrieval parameters in the inverse model to study the impact of these changes on the retrieved ozone or temperature profiles. The parameters used in the inverse model correspond to the "input quantities" introduced in **chapter 2** of this report, and must come with uncertainty estimates.

Going one step further, the Monte Carlo experiments consist not only of varying a specific parameter, but of specifically creating a set of normally-distributed values of this parameter with a known mean and standard deviation, and then analyze the simulated lidar signals to produce an ozone or temperature profile for each of these values. If the standard deviation of a parameter's normal distribution is taken as the parameter's standard uncertainty, then the reported ozone (respectively temperature) profile standard uncertainty associated with this parameter should be equal to the calculated ozone (respectively temperature) profile standard deviation. The correct

formulation of the uncertainty propagation equations used in the retrieval model (**chapters 4-6**) is confirmed only after it is verified that the reported ozone (or temperature) standard uncertainty equals the standard deviation obtained from the corresponding dedicated Monte Carlo experiment. Assuming that all input quantities are independent of each other, we can repeat the verification/quantification process described above for each parameter taken separately.

A.1 Producing simulated lidar signals (forward model)

The same forward model is used to simulate the temperature and ozone lidar signals. It is a numerical implementation of **Eq. (1.1)**, with the exception that it reflects the raw lidar signals recorded in the data files instead of the signals collected on the lidar detectors. It therefore includes effects of the data recorders, namely the addition of sky and electronic background noise, and the inclusion of signal saturation (pile-up) for channels operating in photon-counting mode.

Following the recommendations and approach described in **chapter 3**, we start from a well-known atmospheric state defined by a temperature profile T_a, a number density profile N_a, and mixing ratio profiles of ozone, water vapor, NO_2, SO_2 and O_2. Simulations are performed for altitudes ranging roughly between the ground and 120 km, thus covering the typical measurement range of ozone and temperature lidars. An example of atmospheric state ("true" profiles) is shown in **Figure A**.1. The plotted profiles were produced using one or several of the ancillary datasets recommended in **section 3.3**. For this simulation exercise, the actual location and time of the simulated measurements do not matter, a simple climatology or standard atmosphere is sufficient. Here we simply ensured that the boundary layer is assumed to be highly polluted in order to account for interference by NO_2 (Ahmad et al., 2007) and SO_2.

Using this atmospheric state, simulated lidar signals are produced for three different lidar systems (stratospheric ozone, tropospheric ozone, temperature) with instrumental parameters typical of NDACC lidar systems. The general characteristics of the simulated instruments are compiled in **Table A**.1.

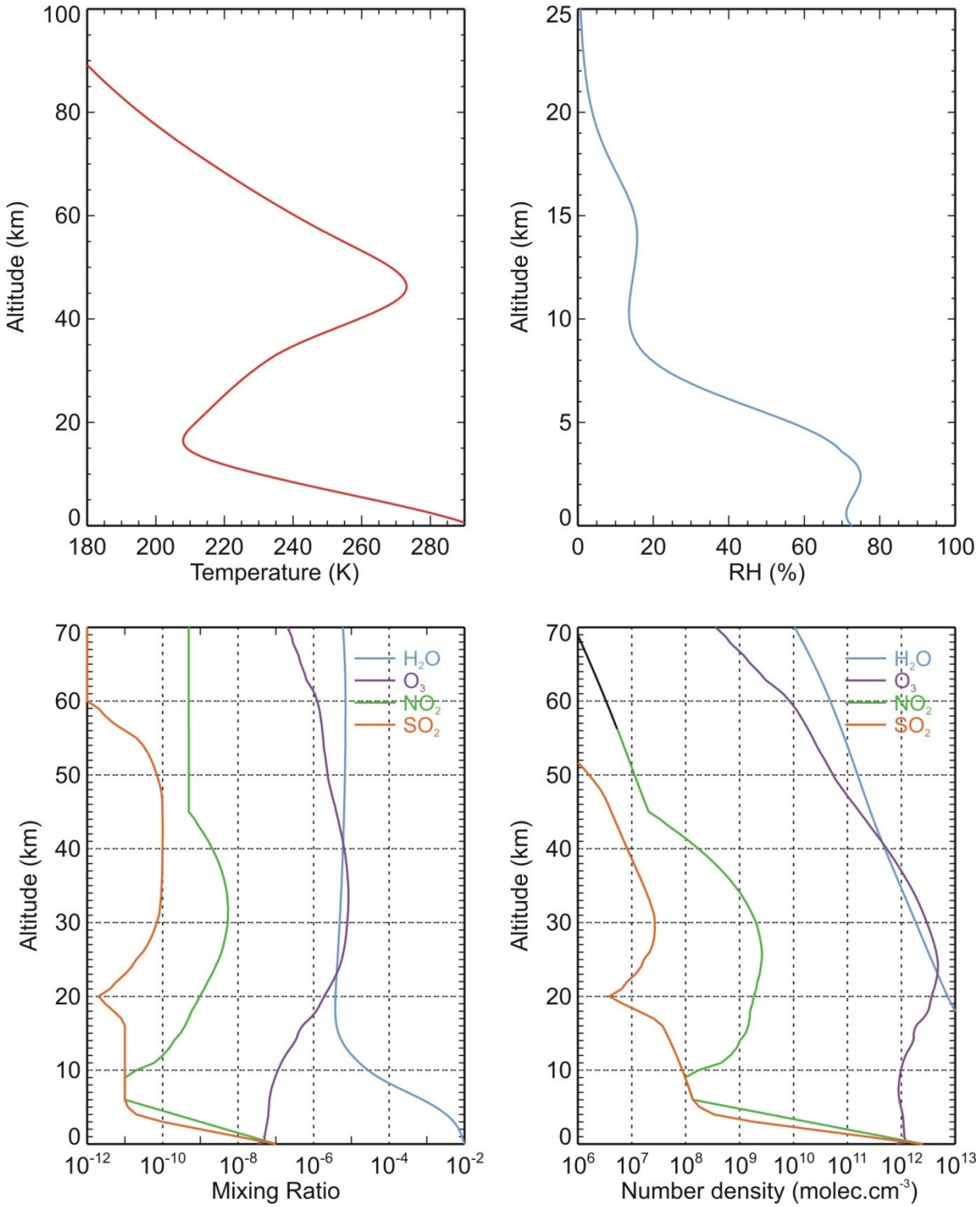


Figure A.1 Typical initial atmospheric profiles ("true" state) used in the forward model to simulate raw lidar signals presented in this work

Table A.1 Values of the forward and inverse models' instrumental parameters

Simulated measurement: Stratospheric ozone	
Name of simulated instrument	O3S
Geolocation of simulated instrument	34.4N, 117.7W, 0 m a.s.l.
DIAL "ON"/"OFF" laser rep. rate (*Hz*)	200/50
Number of simulated channels	6
Sampling resolution (*m*)	75a
Number of data bins	2048
Detection mode	Photon-counting, all channels
DIAL "ON"/"OFF" emitted wavelength (*nm*)	308/355
DIAL "ON"/"OFF" detected wavelength (*nm*)	308/355 high-intensity 308/355 low-intensity 332/387 Raman
Simulated datasets	Multiple datasets of 5 min each
Simulated measurement: Tropospheric ozone	
Name of simulated instrument	O3T
Geolocation of simulated instrument	34.4N, 117.7W, 0 m a.s.l.
Laser rep. rate (*Hz*)	30
Number of simulated channels	8
Sampling resolution (*m*)	30
Number of data bins	2047
Detection mode	Phton-counting, all channels
DIAL "ON"/"OFF" emitted wavelength (*nm*)	299/316 289/299 287/294 266/289
DIAL "ON"/"OFF" detected wavelength (*nm*)	299/316 high-intensity 289/299 high-intensity 287/294 med-intensity 266/289 low-intensity
Simulated datasets	Multiple datasets of 5 min each
Simulated measurement: Temperature	
Name of simulated instrument	TMP
Geolocation of simulated instrument	34.4N, 117.7W, 0 m a.s.l.
Laser rep. rate (*Hz*)	50
Number of simulated channels	6
Sampling resolution (*m*)	75
Number of data bins	2048
Detection mode	Photon-counting, all channels
Emitted wavelength (*nm*)	355 532
Detected wavelength (*nm*)	355 high-intensity 355 low-intensity 387 Raman 532 high-intensity 532 low-intensity 607 Raman
Simulated datasets	Multiple datasets of 5 min each

For all three simulated instruments, **Eq. (1.1)** describing the signals collected on the detectors can be written in its numerical form:

$$P(i,k) = \frac{\kappa(i) N_X(k)}{\left(z(k) - z_L\right)^2} \exp\left[\begin{array}{l} -\sum_{k'=0}^{k} \left(\sigma_{M,UP}(i) + \sigma_{M,DOWN}(i)\right) N_a(k') \delta z \\ -\sum_{k'=0}^{k} \left(\sigma_{O3,UP}(i,k') + \sigma_{O3,DOWN}(i,k')\right) N_{O3}(k) \delta z \\ -\sum_{k'=0}^{k} \left(\sigma_{NO2,UP}(i,k') + \sigma_{NO2,DOWN}(i,k')\right) N_{NO2}(k) \delta z \\ -\sum_{k'=0}^{k} \left(\sigma_{SO2,UP}(i,k') + \sigma_{SO2,DOWN}(i,k')\right) N_{SO2}(k) \delta z \\ -\sum_{k'=0}^{k} \left(\sigma_{O2,UP}(i,k') + \sigma_{O2,DOWN}(i,k')\right) N_{O2}(k) \delta z \end{array}\right] \quad \textbf{(A.1)}$$

Here we have introduced the number density N_X, which can be either air number density N_X=N_a (Rayleigh backscatter) or nitrogen number density N_X=N_{N2} (vibrational Raman backscatter). We also introduced the channel index i, the altitude bin number k, and the term κ, which includes all altitude-independent terms of **Eq. (1.1)** and therefore has no impact on the ozone and temperature retrievals besides detection noise (a higher κ yields a higher altitude range for the same precision, or better precision at the same altitude range).

The simulated lidar signals S_0 to be written in the raw data files must also include the effect of saturation (pulse pile-up) for channels operating in photon-counting mode, and must include background noise in all channels to account for sky light and electronic noise:

If operating in photon-counting mode:

$$S_0(i,k) = \varepsilon(i) \frac{P(i,k)}{1 + \tau(i) P(i,k)} + b_0(i) + b_1(i) z(k) \quad \textbf{(A.2)}$$

ε represents the amplification factor or efficiency of the data recorder system, τ is the dead-time characterizing the speed of the photon-counting system (a longer dead-time will saturate the signals at a lower count rate), and b_0 and b_1 are the coefficient of a linear function of altitude parameterizing the background noise.

If operating in analog detection mode, there is no pile-up effect, but there often is a delay between the acquisition of the PC and AD signals, which is equivalent to a bin shift δk. **Eq. (A**.2**)** therefore becomes:

$$S_0(i,k) = \varepsilon(i) P(i, k + \delta k) + b_0(i) + b_1(i) z(k) \quad \textbf{(A.3)}$$

The parameters defining the atmospheric state and the simulated data acquisition electronics are listed in **Table A**.2. These parameters are also the parameters used in the inverse model, and constitute the input quantities for which Monte Carlo experiments will be performed. In an effort to reproduce realistic lidar signals, the model also includes multiple options to mechanically or electronically gate the signals. This functionality exists only for practical reasons in case the simulated signals need to be analyzed by an existing operational data processing software tailored for gated signals. It does not impact the Monte Carlo experiments. Note that very few

NDACC ozone and temperature lidar instruments comprise analog channels, and only results from PC-only simulated systems will be shown thereafter.

Table A.2 Values or source of the forward model input parameters

	Description	***Source dataset or value***
From atmospheric state		
T_a	Air temperature profile	MSIS-90 at simulated location
N_a	Air number density profile	MSIS-90 at simulated location
N_{O3}	Ozone number density profile	UKMO climatology at simulated location
N_{NO2}	NO_2 number density profile	Fixed single profile
N_{SO2}	SO_2 number density profile	Fixed single profile
N_{O2}	O_2 number density profile	MSIS-90 at simulated location
q_{H2O}	Water vapor mixing ratio profile	Fixed single profile
From theoretical studies		
σ_M	Rayleigh extinction cross-section	Eberhard (see reference)
From laboratory studies		
σ_{O3}	Ozone absorption cross-section	Daumont- Malicet-Brion (see reference)
σ_{NO2}	NO_2 absorption cross-section	Bogumil (see reference)
σ_{SO2}	SO_2 absorption cross-section	Bogumil (see reference)
σ_{O2}	O_2 absorption cross-section	IASB (see reference)
From instrumentation		
τ	PC hardware dead-time	Typically 250 MHz
δk	AD bin shift with PC	Typically < 15 m
b_0	Background noise coefficient 0	Depends on channel
b_1	Background noise coefficient 1	Depends on channel (typically, 0 for flat noise)

Figure A.2 shows an example of simulated raw signals for the tropospheric ozone (O3T), stratospheric ozone (O3S), and temperature (TMP) lidars corresponding to the atmospheric state shown in **Figure A**.1 and for a 2-hour-long simulated measurement (concatenation of 24 5-min datasets). The forward model was built in a way to study the impact of each instrumental or retrieval parameter separately (thereafter referred to as the "target parameter" for brevity). Detection noise, saturation effects, background noise, and any of the extinction terms may therefore be set to zero independently from each other depending on the needs of a particular simulation experiment. For example we can produce simulated signals with no detection noise, no saturation, no background noise, and no absorption by any minor species to study the sole impact of molecular extinction. Similarly we can produce simulated signals with no detection noise, no background noise, no absorption by any minor species, and no molecular extinction to study the sole impact of saturation correction.

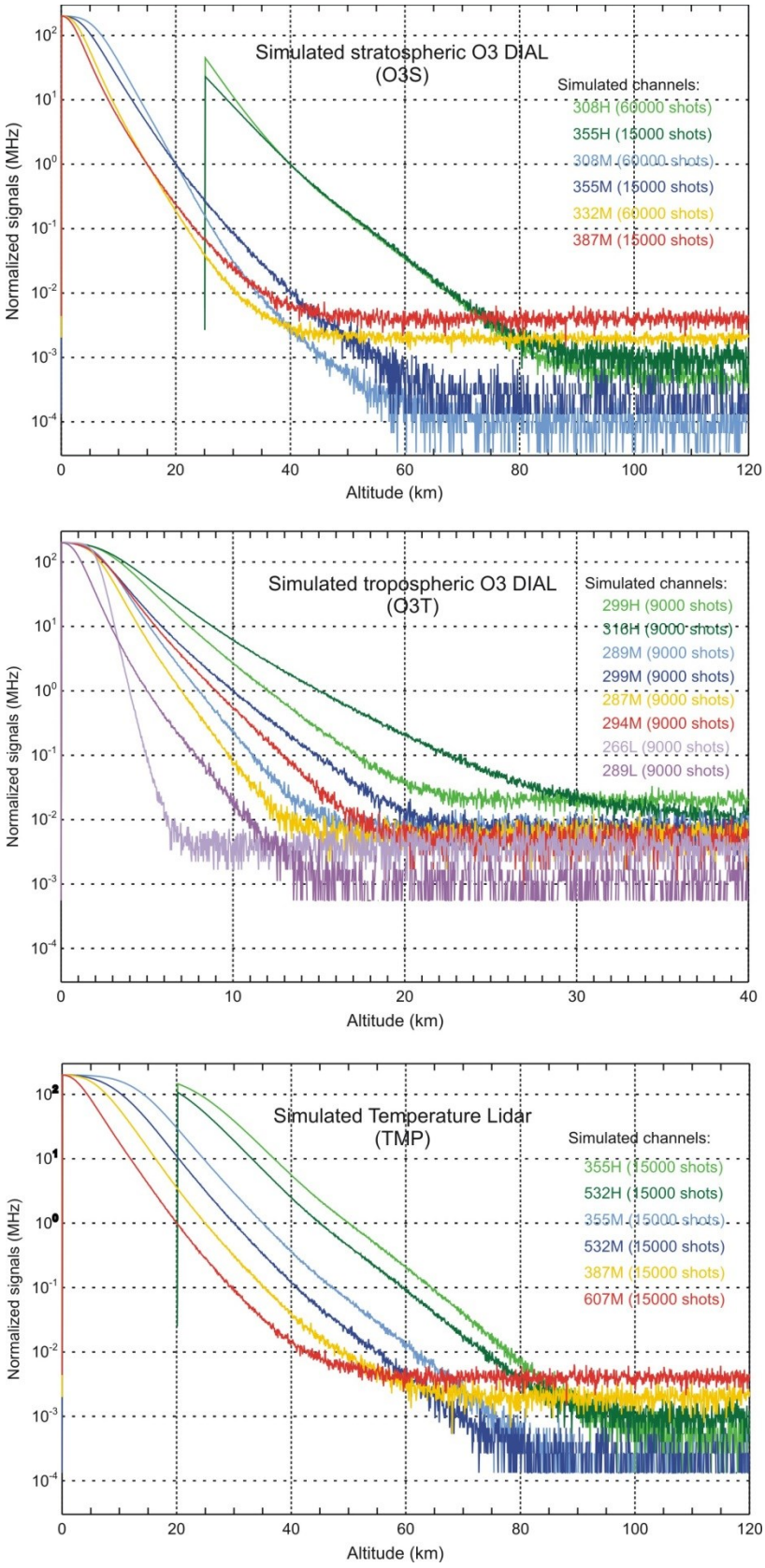


Figure A.2 Example of tropospheric ozone, stratospheric ozone, and temperature lidar signals simulated by the forward model when using the atmospheric state shown in Figure A.1

A.2 Analyzing simulated lidar signals (inverse model)

The inverse model is similar to any data processing algorithm used to retrieve ozone and temperature, and therefore includes all the usual signal corrections such as background noise, saturation, and extinction. In order to ensure consistency, the inverse model was developed jointly with the forward model, with the objective to be fully compliant with the recommendations of the ISSI-Team on vertical resolution and uncertainty. Its structure is similar to that of the lidar data processing software *LidAna* v6.2 used for the routine processing of the Jet Propulsion Laboratory tropospheric ozone, stratospheric ozone, temperature, water vapor and aerosol lidar measurements archived at NDACC. It includes paralyzable and non- paralyzable saturation correction modules, a background correction module that can handle linear and non-linear background noise extraction, an optional overlap correction routine, correction modules for Rayleigh extinction, absorption by ozone, NO_2, SO_2 and O_2, and multiple vertical filtering and channel-merging options. In order to illustrate the consistency between the inverse and forward models, **Figure A**.3 shows the differences between the retrieved and the "true" profiles for the stratospheric ozone (O3S), tropospheric ozone (O3T) and temperature (TMP) simulated instruments in absence of photon-counting noise. The observed differences are negligible to the extent that the true profiles (black curves on the left plots) are not even visible below the retrieved profiles curves. The only apparent differences are at the very top of the profiles and are due to rounding errors owed to the fact that the forward model writes out small integer numbers in the simulated raw data files.

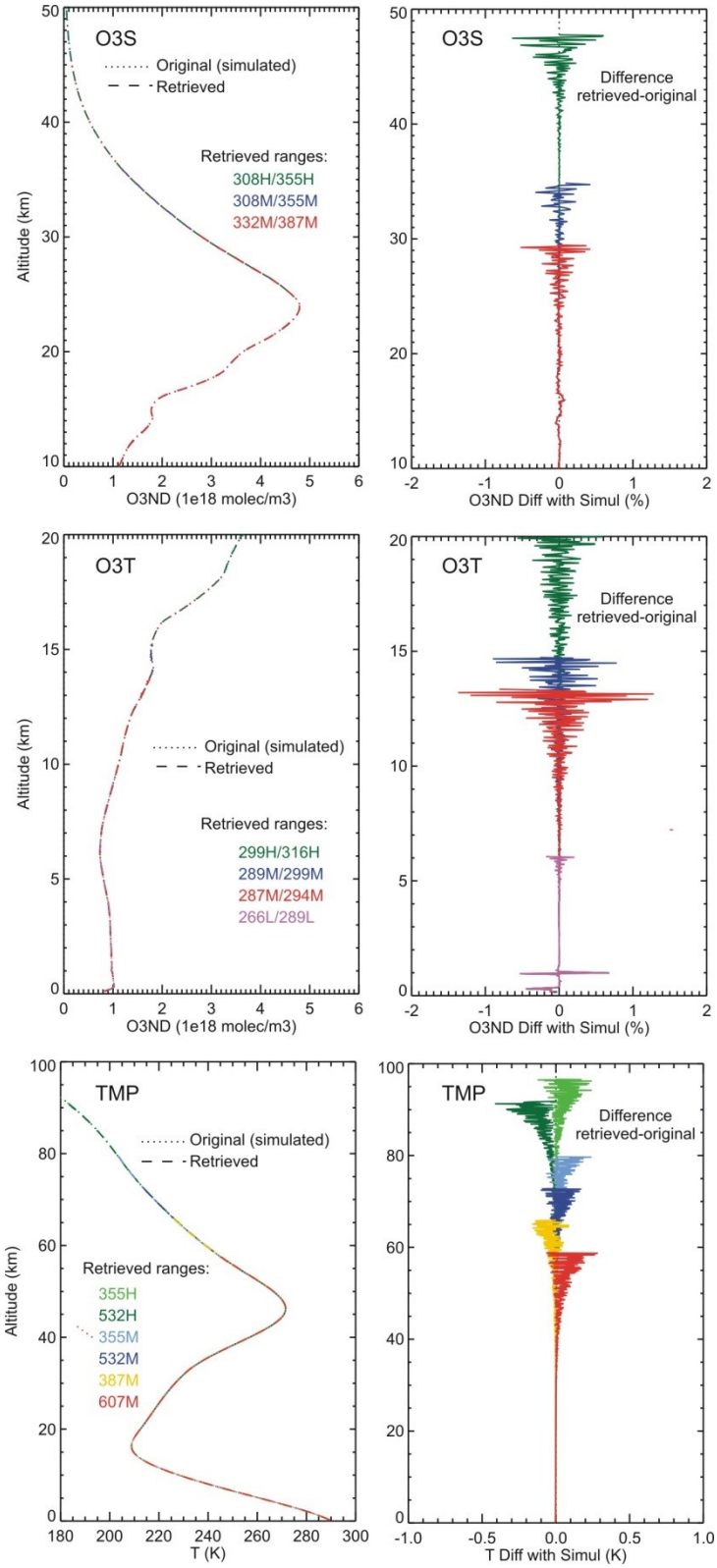


Figure A.3 Left plots: Original ("true") and retrieved profiles obtained from lidar signals simulated for stratospheric ozone O3S (top), tropospheric ozone O3T (middle), and temperature TMP (bottom). Right plots: Difference (%) between them (see text for details). The simulated Rayleigh high-intensity range is plotted in green, the Rayleigh low-intensity range in blue, and the Raman range in red.

A.3 Monte-Carlo experiments operating principle

Once we have ensured that the inverse model and forward model are consistent, we can run the Monte-Carlo experiments specifically dedicated to the quantification of uncertainties. The best way to illustrate these experiments is to describe a specific example. Here we will therefore provide a description of a Monte Carlo experiment dedicated to the quantification of the temperature uncertainty owed to saturation correction. All Monte-Carlo experiments discussed later in this report use the exact same procedure, but applied to other sources of uncertainty.

In our example, the experiment starts with the simulation of raw lidar signals for the simulated temperature lidar system "TMP" with all six channels operating in photon-counting mode. The saturation correction parameter (i.e., the dead-time) is the target parameter of the experiment, therefore all corrections except the saturation correction are turned "OFF" in the forward and inverse models and do not contribute to the temperature uncertainty budget. Referring to **Table A**.2 and **Eqs. (A.**1)**-(A.**3), this is equivalent to taking the following values for the parameters (input quantities) of the forward and inverse models:

- No molecular extinction correction: $\sigma_{M,UP} = \sigma_{M,DOWN} = N_a = 0$
- No correction for absorption: $\sigma_{IG,UP} = \sigma_{IG,DOWN} = N_{IG} = 0$ (IG=O_3, NO_2 SO_2, and O_2)
- No background noise: $b_0 = b_1 = 0$
- An arbitrary non-zero value for the dead-time for all channels: $\tau = 4\,\text{ns}$

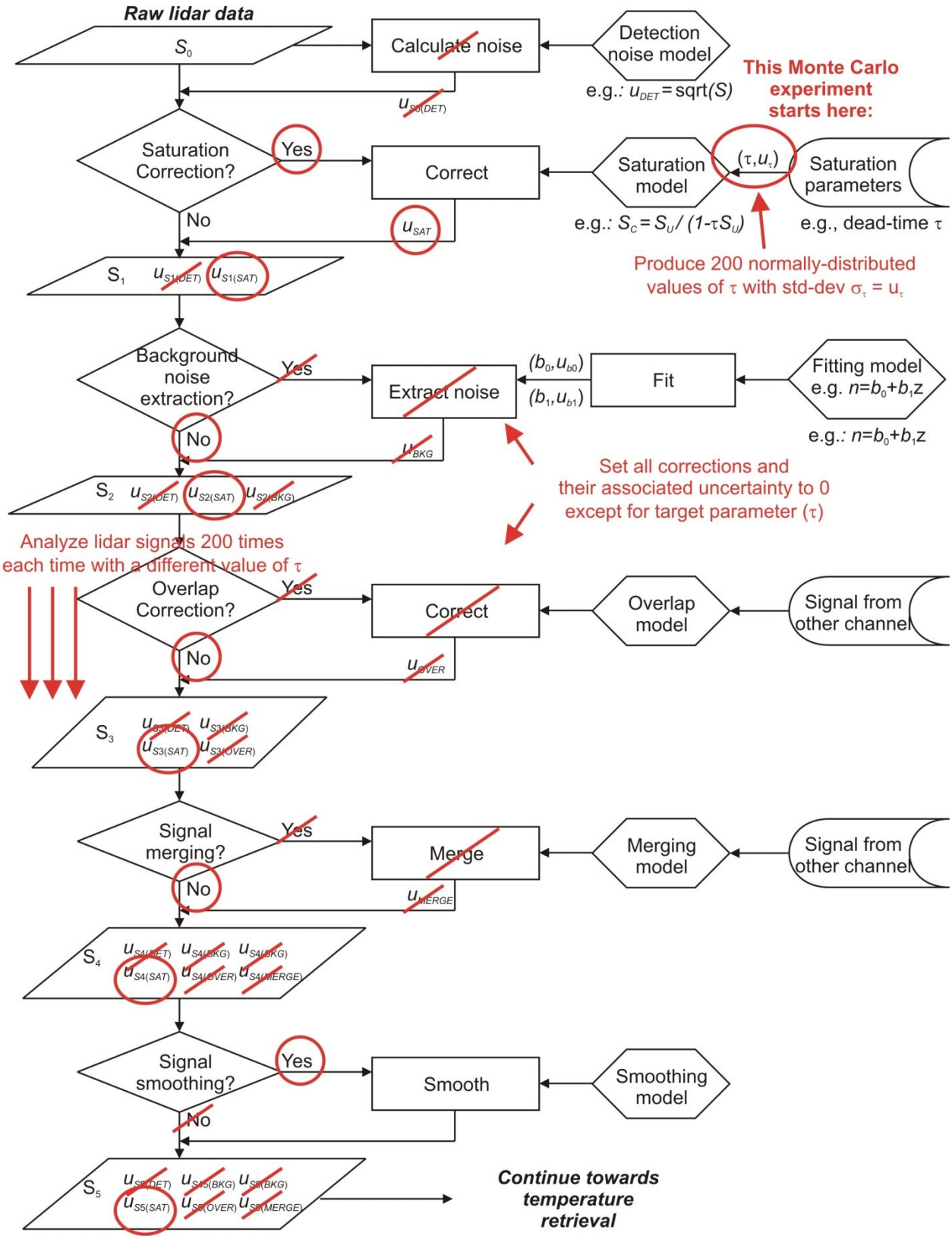
The Monte Carlo experiment consists not only of retrieving temperature from these simulated signals, but of retrieving a large number of temperature profiles (e.g., 200), each time with a different value of dead-time. To do this, we create a normally-distributed population of dead-time values with a mean equal to the dead time expected value (e.g., 4 ns) and a standard deviation equal to the dead-time standard uncertainty (e.g., 10% or 0.4 ns). For each simulated lidar channel we analyze N times the simulated signals (e.g., N=200), each time using a different value of dead-time defined by:

$\tau(0) = \bar{\tau} = 4$ for the control analysis

$\tau(j) = \bar{\tau} + \delta\tau(j)$ for each of the N analysis of this Monte Carlo experiment (j=1,N)

$\delta\tau$ is the normally-distributed array of dead-time perturbations of mean $\overline{\delta\tau} = 0$ and standard deviation $\sigma_{\delta\tau} = u_\tau$ (u_τ is the dead-time standard uncertainty). To minimize numerical errors, the population must contain a large number of samples (at least N=200, depending on the numerical tool used to create the population). We then calculate the temperature profile standard deviation obtained from the N temperature profiles retrieved using the N different dead-time values. **Figure A**.4 summarizes the procedure. This two-page flowchart is a modified version of the flowcharts shown on **Figure 4.1 and Figure 6.1** adapted to the present Monte Carlo experiment example. **Figure A**.5 shows an example of the retrieved temperature profiles. Here we show only 50 (of the 200) retrieved profiles for the sake of clarity in the figure. The effect of the different dead-time values is obvious at the bottom of the profiles where temperature departs significantly from the original ("true") profile. The ultimate purpose of this exercise is to verify that

uncertainty owed to the saturation correction propagated to temperature has been correctly quantified in the inverse model. To do so, the temperature standard deviation obtained from all *N* profiles is calculated and compared to the standard uncertainty calculated by the inverse model. If they match, it means that the expressions used for the propagation of this particular uncertainty component are properly computed in the inverse model. **Figure A**.6 shows the results of this comparison for the present example. The dotted curve depicts the standard deviation calculated from the N retrieved profiles, while the dashed curve depicts the standard uncertainty calculated by the inverse model. As we expected for a successful test, the uncertainty and standard deviation curves perfectly overlap. Note that the actual values of uncertainty and standard deviation do not matter here as they depend on the lidar system considered, whether it is for a simulated or an actual lidar instrument. The key information here is the fact that the uncertainty and the standard deviation curves overlap.



3098

3099 (continues on next page)

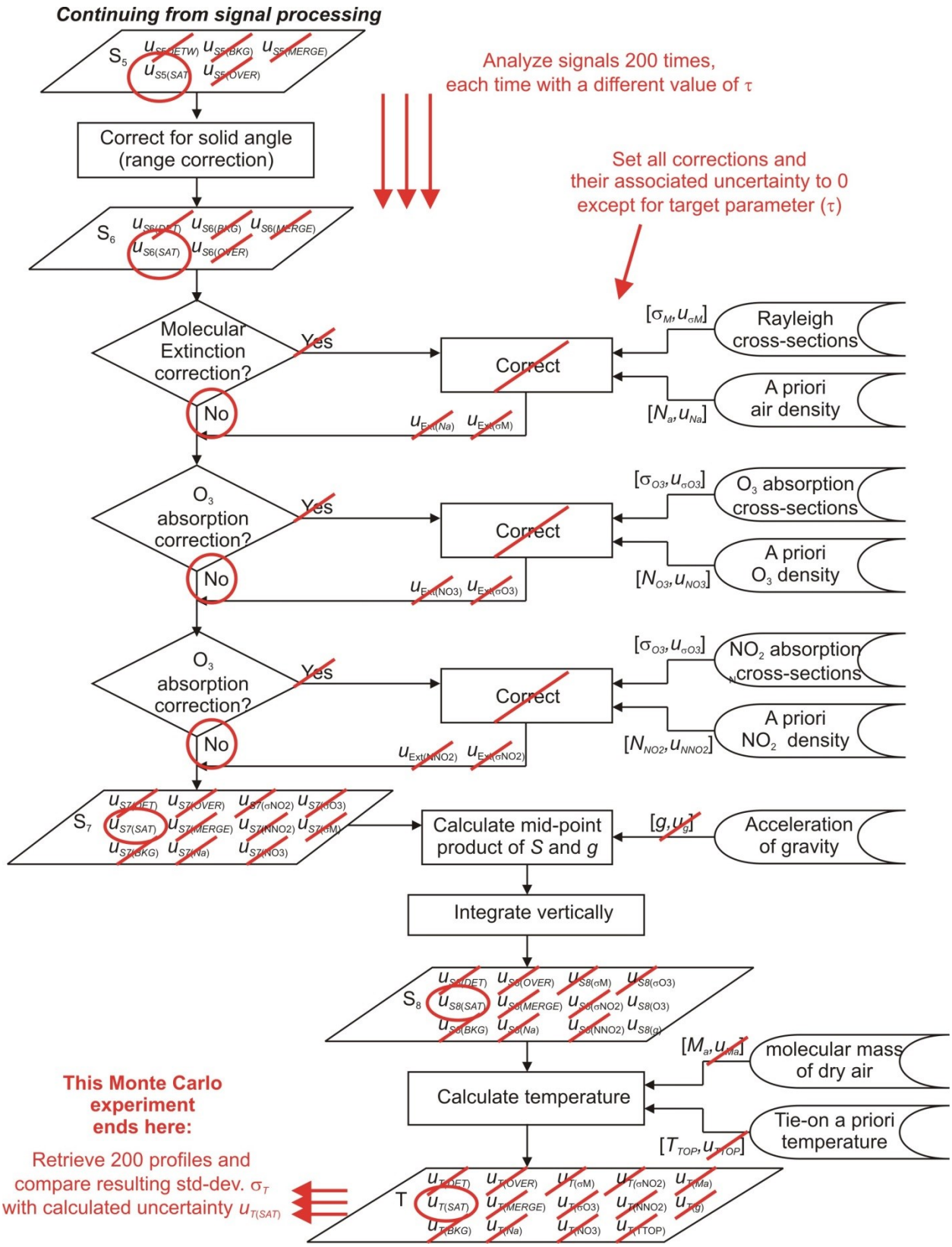


Figure A.4 Flowchart illustrating the signal processing for the retrieval of temperature in the case of the Monte Carlo experiment example described in this chapter

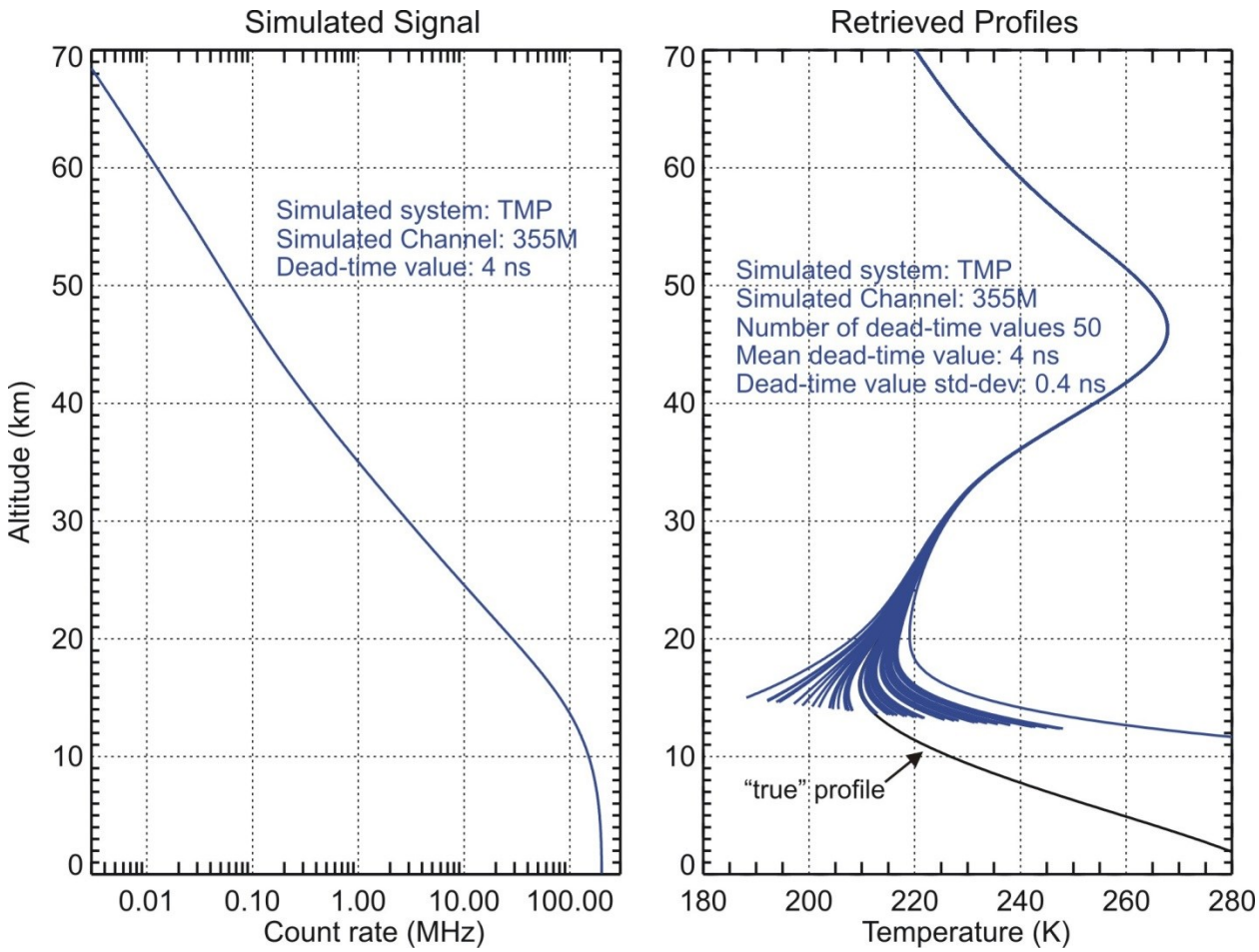


Figure A.5 Left plot: Example of simulated lidar signal including only saturation effect (i.e., no extinction, no background noise, no detection noise). Right plot: Results of a Monte-Carlo experiment with dead-time being the target parameter, as described in this chapter (see text for details)

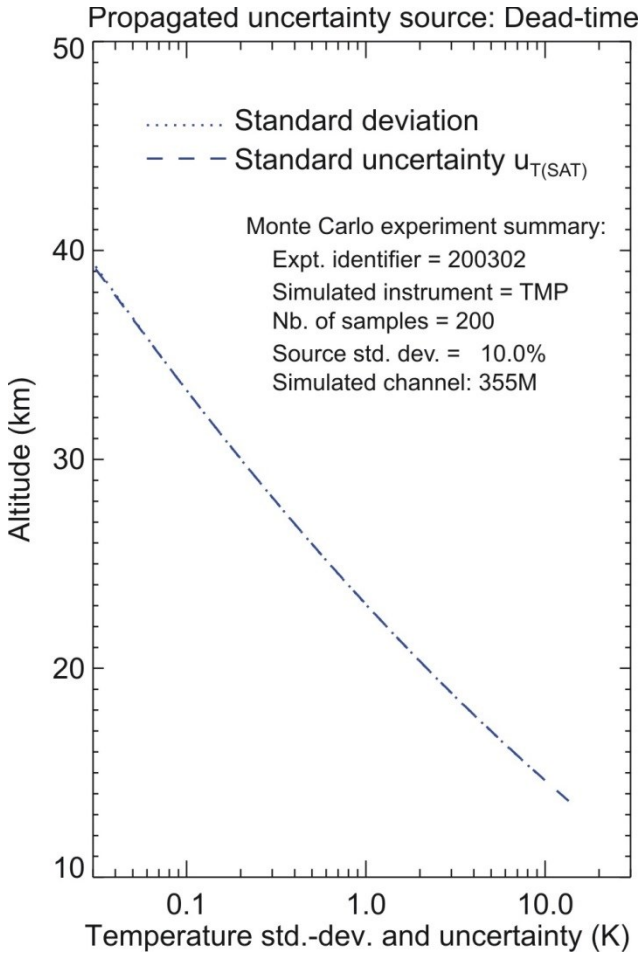


Figure A.6 Temperature standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from the Monte Carlo experiment designed to quantify temperature uncertainty owed to dead-time correction $u_{T(SAT)}$, as described in this chapter (see text for details)

The experiment just described can be repeated for all target parameters, each time by setting all input parameters to zero except the target parameter. This procedure allows a separate quantification of uncertainty owed to each independent input parameter. It can be repeated for all independent sources of uncertainty, and for the simulated stratospheric ozone (O3S), tropospheric ozone (O3T), and temperature (TMP) lidars. **Appendix B** and **appendix C** provide the quantitative validation of the expressions detailed in **chapters 4-6** describing the propagation of uncertainty through the ozone and temperature data processing chains.

B Quantitative validation of uncertainty propagated to ozone

For each uncertainty source introduced in the ozone DIAL measurement sub-models of **chapters 4-5**, we validate the appropriate use of the propagation expressions provided in these **chapters**. We show that the propagated ozone standard uncertainty calculated in parallel for each component quantitatively matches the ozone standard deviation calculated for the corresponding dedicated Monte-Carlo experiment.

In each of the experiments described thereafter, 200 simulated lidar datasets are used, each dataset being equivalent to a 5-min accumulation of photocounts. The instrumental parameters used for the simulations are those listed in **Table A**.1 for the "O3S" (stratospheric ozone) and "O3T" (tropospheric ozone) simulated lidar intruments. A Savitsky-Golay derivative low-pass filter (Savitzky and Golay, 1964) with a vertical width increasing with altitude is applied to the signals. This vertical filtering procedure is typical of ozone DIAL retrievals, and has the effect of not only differentiating but also smoothing. This filtering was used in all Monte Carlo experiments, even experiments producing simulated signals containing no detection noise, in order to verify that uncertainty components associated with variables correlated in altitude are properly propagated. Therefore, for all the examples shown, a matching standard deviation and calculated uncertainty implies the correct use of the equations written in **section 5.2** (differentiation and smoothing). To avoid excessive or unnecessary smoothing, the equations written in **section 4.6** (signal smoothing) were not used in the examples below. Yet they remain valid at any time, even if a derivative low-pass filter is also used.

B.1 Detection noise uncertainty propagated to ozone

Uncertainty associated with detection noise and propagated to ozone number density $u_{O3(DET)}$ is plotted for the stratospheric ozone DIAL system "O3S" in **Figure B**.1 (dashed curves). The results are presented in percent on the left plot and in part-per-million on the right plot. The corresponding propagation equations are listed in **Table B**.1**.** Detection noise was simulated using a Poisson distribution around the mean number of photons detected with no correlation between altitude bins, no correlation between simulated channels, and no correlation between any of the 200 simulated datasets (independent datasets). In **Figure B**.1, the dotted curves show the ozone standard deviation resulting from a Monte-Carlo experiment with detection noise only (i.e., no saturation, no background noise, no extinction terms). The fact that the dotted curves (standard deviation) and dashed curves (calculated uncertainty) match perfectly demonstrates that the set of equations listed in **Table B**.1 for the propagation of uncertainty owed to detection noise are all correct.

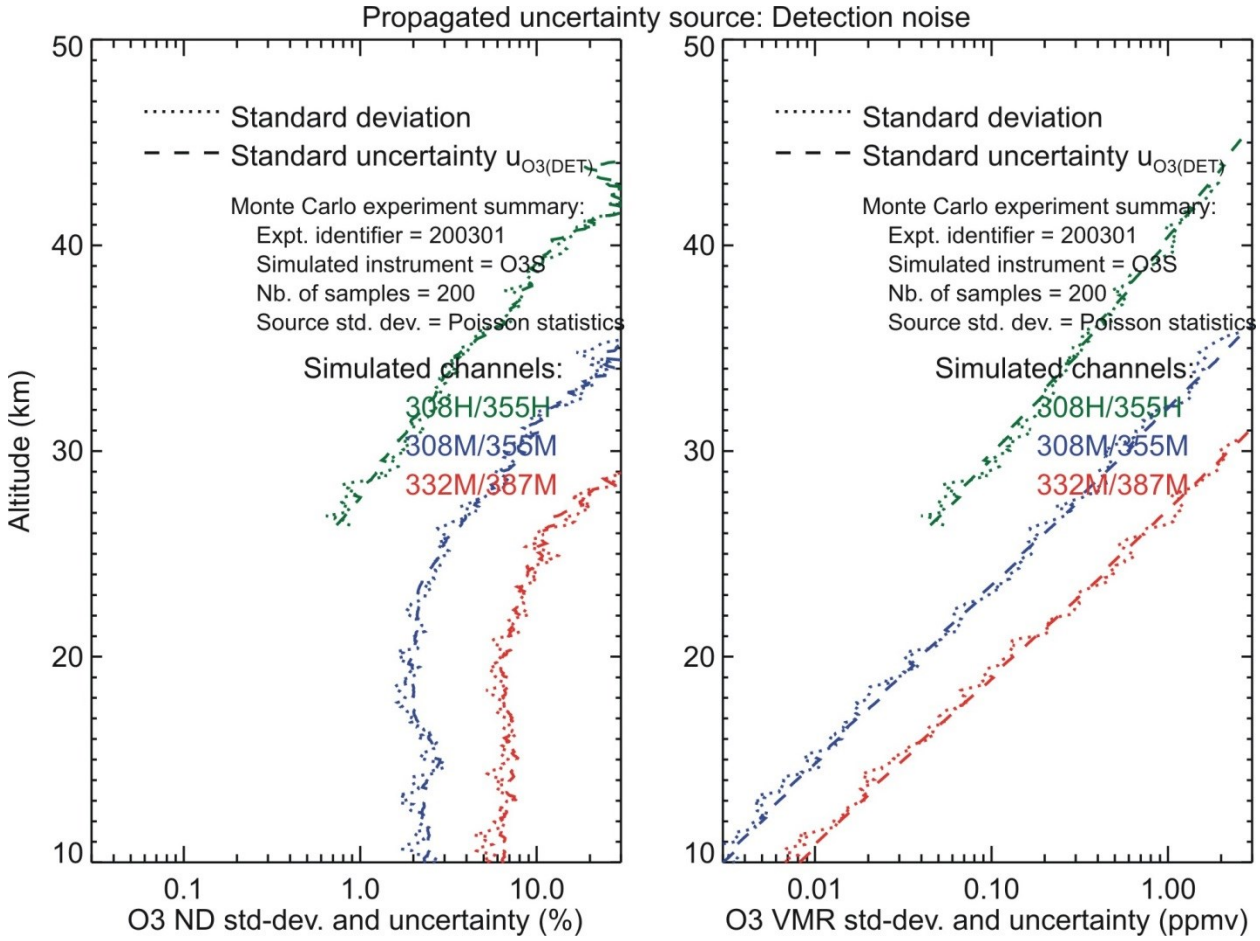


Figure B.1 Stratospheric ozone standard uncertainty owed to detection noise $u_{O3(DET)}$ (long-dash curves) and standard deviation (dotted curves) obtained from the corresponding dedicated Monte Carlo experiment (see text for details)

Table B.1 Expressions used for detection noise uncertainty propagated to ozone

Sub-model description	*Propagated uncertainty*	*Eq. Non-paralyz.*	*Eq. Paralyz.*
Signal detection (PC)	$u_{S0(DET)}$	(4.3)	(4.3)
Saturation correction	$u_{S1(DET)}$	(4.14)	(H.9)
Background extraction	$u_{S2(DET)}$	(4.37)	(4.37)
Merging	$u_{S4(DET)}$	(4.46) (4.52)	(4.46) (4.52)
Smoothing (lin) Smoothing (log)	$u_{S5(DET)}$	(4.66) (4.74)	(4.66) (4.74)
Log(S_{ON}/S_{OFF})	$u_{S6(DET)}$	(5.4)	(5.4)
Differentiation	$u_{S7(DET)}$	(5.18)	(5.18)
DIAL equation	$u_{NO3(DET)}$ $u_{qO3(DET)}$	(5.25) (5.56)	(5.25) (5.56)

B.2 Saturation correction uncertainty propagated to ozone

Uncertainty associated with saturation correction and propagated to ozone number density $u_{O3(SAT)}$ is plotted for the stratospheric ozone DIAL system "O3S" in **Figure B**.2 (dashed curves) for the two configuration cases discussed in **chapter 5**. The corresponding propagation equations are listed in **Table B**.2. In one configuration (left plot) the "ON" and "OFF" channels use independent photon-counting hardware (equations listed in column 3 of **Table B**.2). In the other configuration (right plot) they share the same photon-counting hardware (equations listed in column 4 of **Table B**.2). The dotted curves show the ozone standard deviation resulting from the corresponding two Monte-Carlo experiments in which the lidar signals were simulated with saturation correction only (i.e., no detection noise, no background noise, no extinction terms). In the "independent hardware" case, two independent populations of 200 normally-distributed dead-time values were used in the inverse model for the "ON" and "OFF" channels. In the "shared hardware" case, the same population of 200 normally-distributed dead-time values was used for the "ON" and "OFF" channels. In both cases, the dotted curves (standard deviation) and dashed curves (calculate uncertainty) match perfectly, which confirms that the set of equations listed in **Table B**.2 is correct for both configurations. It is also interesting to note that the magnitude of the calculated uncertainty between one configuration and the other is quite different, as was anticipated in view of the results of **Appendix B**.

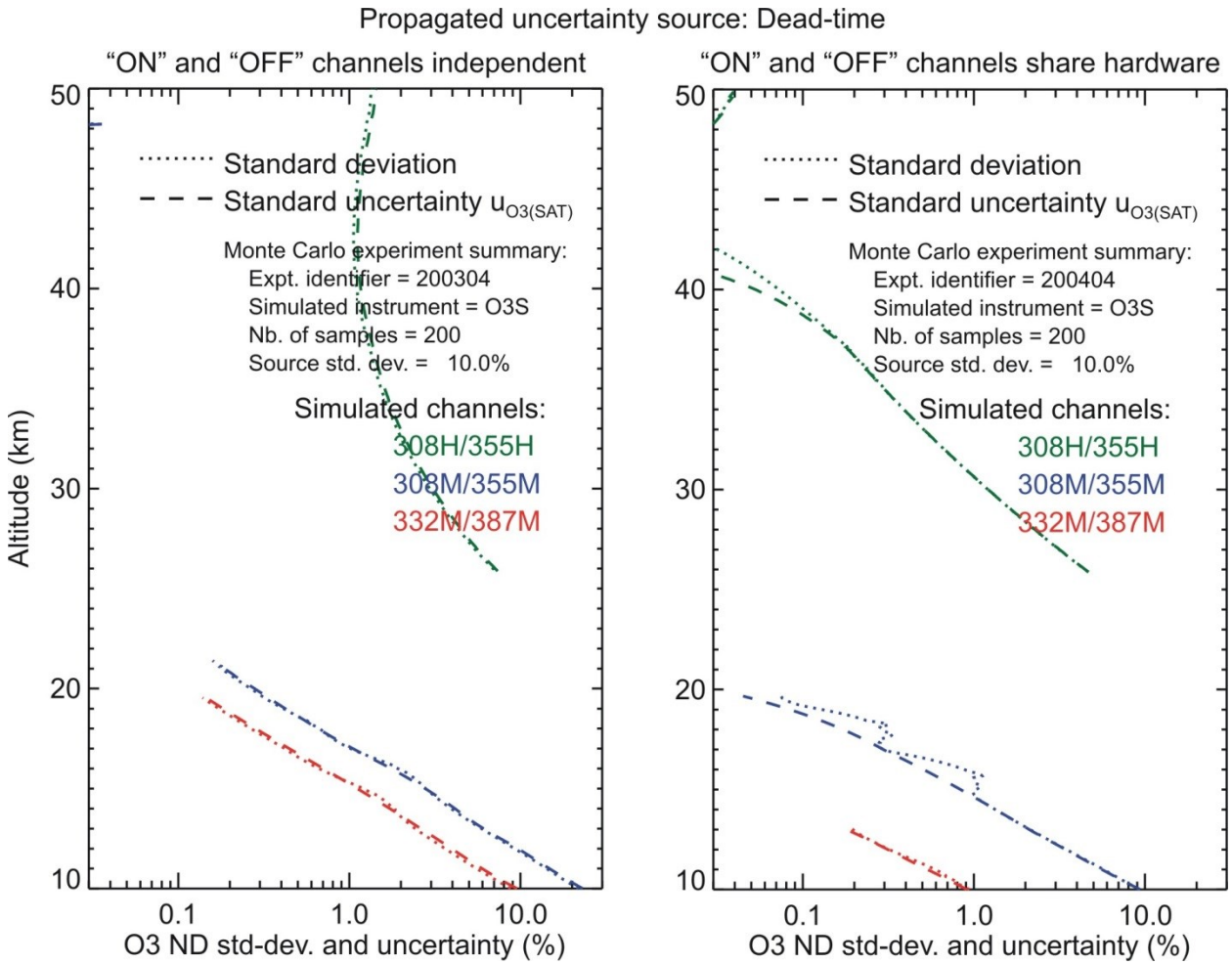


Figure B.2 Stratospheric ozone standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from a Monte Carlo experiment designed to quantify ozone uncertainty owed to dead-time correction $u_{O3(SAT)}$. On the left plot, it is assumed that the "ON" and "OFF" channels use independent counting hardware On the right plot, it is assumed that the "ON" and "OFF" channels share the same counting hardware (see text for details)

Table B.2 Expressions used for saturation correction uncertainty propagated to ozone

Sub-model Description	***Propagated uncertainty***	***Eq. if counting hardware independent Non-paralyz***	***Eq. if counting hardware shared Non-paralyz***	***Eq. if counting hardware independent Paralyz.***	***Eq. if counting hardware shared Paralyz.***
Saturation correction	$u_{S1(SAT)}$	(4.15)	(4.15)	(H.8)	(H.8)
Background extraction	$u_{S2(SAT)}$	(4.38)	(4.38)	(4.38)	(4.38)
Merging (one-point) Merging (with overlap)	$u_{S4(SAT)}$	(4.46) (4.53)	(4.46) (4.56)	(4.46) (4.53)	(4.46) (4.56)
Smoothing (lin) Smoothing (log)	$u_{S5(SAT)}$	(4.67) (4.75)	(4.67) (4.75)	(4.67) (4.75)	(4.67) (4.75)
$\text{Log}(S_{ON}/S_{OFF})$	$u_{S6(SAT)}$	(5.5)	(5.9)	(5.5)	(5.9)
Differentiation	$u_{S7(SAT)}$	(5.19)	(5.19)	(5.19)	(5.19)
DIAL equation	$u_{NO3(SAT)}$ $u_{qO3(SAT)}$	(5.26) (5.57)	(5.26) (5.57)	(5.26) (5.57)	(5.26) (5.57)

Figure B.3 is identical to **Figure B**.2, but for the tropospheric ozone lidar "O3T". In this striking case, the "ON" and "OFF" signals happen to have a similar magnitude for two DIAL pairs, the 289M/299M pair at 5 km altitude, and the 287M/294M pair at 3 km altitude (see **Figure A**.2). As a result, the saturation correction in the "share hardware" configuration is identical for the "ON" and "OFF" channels, and the resulting propagated uncertainty (and standard deviation) tends towards zero. This result is essential and shows that uncertainty owed to saturation correction is not necessarily a monotonically-decreasing function of altitude range. This feature is typically what would be missed in a simple lidar uncertainty budget in which correlation relations are ignored.

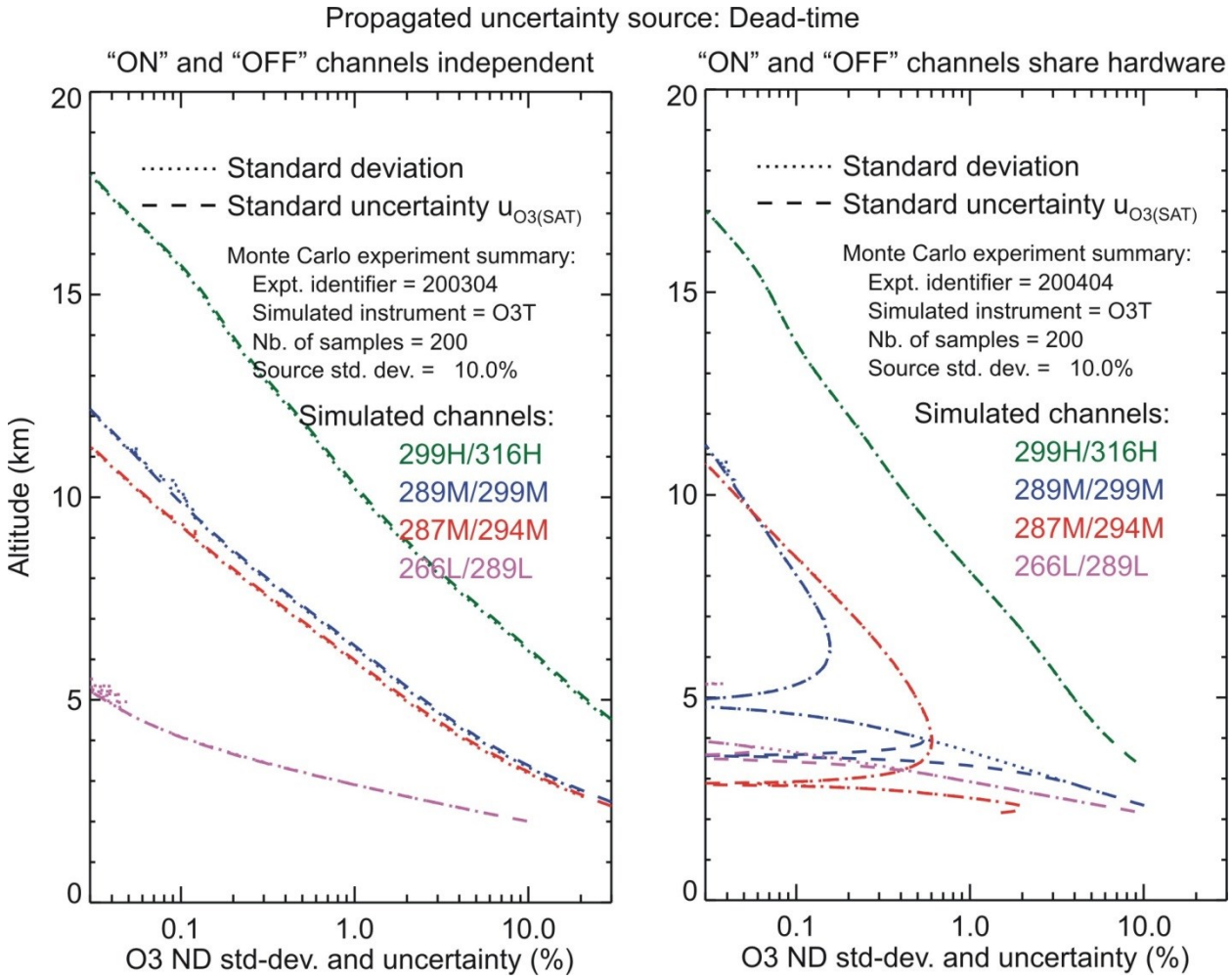


Figure B.3: Same as Figure B.2, but for the tropospheric ozone system O3T

B.3 Background extraction uncertainty propagated to ozone

Uncertainty associated with background correction and propagated to ozone number density $u_{O3(BKG)}$ is plotted for the tropospheric ozone DIAL system "O3T" in **Figure B**.4 (dashed curves). Again the results ar epresente din percent on left plot, and in parts-per-miilion on the right plot. The corresponding propagation equations are listed in **Table B**.3. For this Monte-Carlo experiment, the simulated lidar signals contained detection noise and background noise, but no saturation, and no extinction terms. In the inverse model, background was corrected using the fitting function presented in **section 4.3** (linear function of altitude range). Though detection noise is not the target parameter of this Monte Carlo experiment, it had to be included in the signals in order to produce better fitting results and realistic uncertainty estimates of the fitting function coefficients. In the example shown here, the target parameter was the fitting function coefficient b_1. The magnitude of the coefficient uncertainty was taken directly from the value returned by the fitting routine, and then used as the standard deviation of a normally distributed population of 200 coefficients b_1 subsequently used to correct for background noise and produce 200 ozone profiles. The standard uncertainty reported by the inverse model (long-dash curves) matches again very well the calculated ozone standard deviation for all ranges and all altitudes.

Note that the plotted ozone uncertainty represents the combined uncertainty calculated from both the detection noise and the background correction.

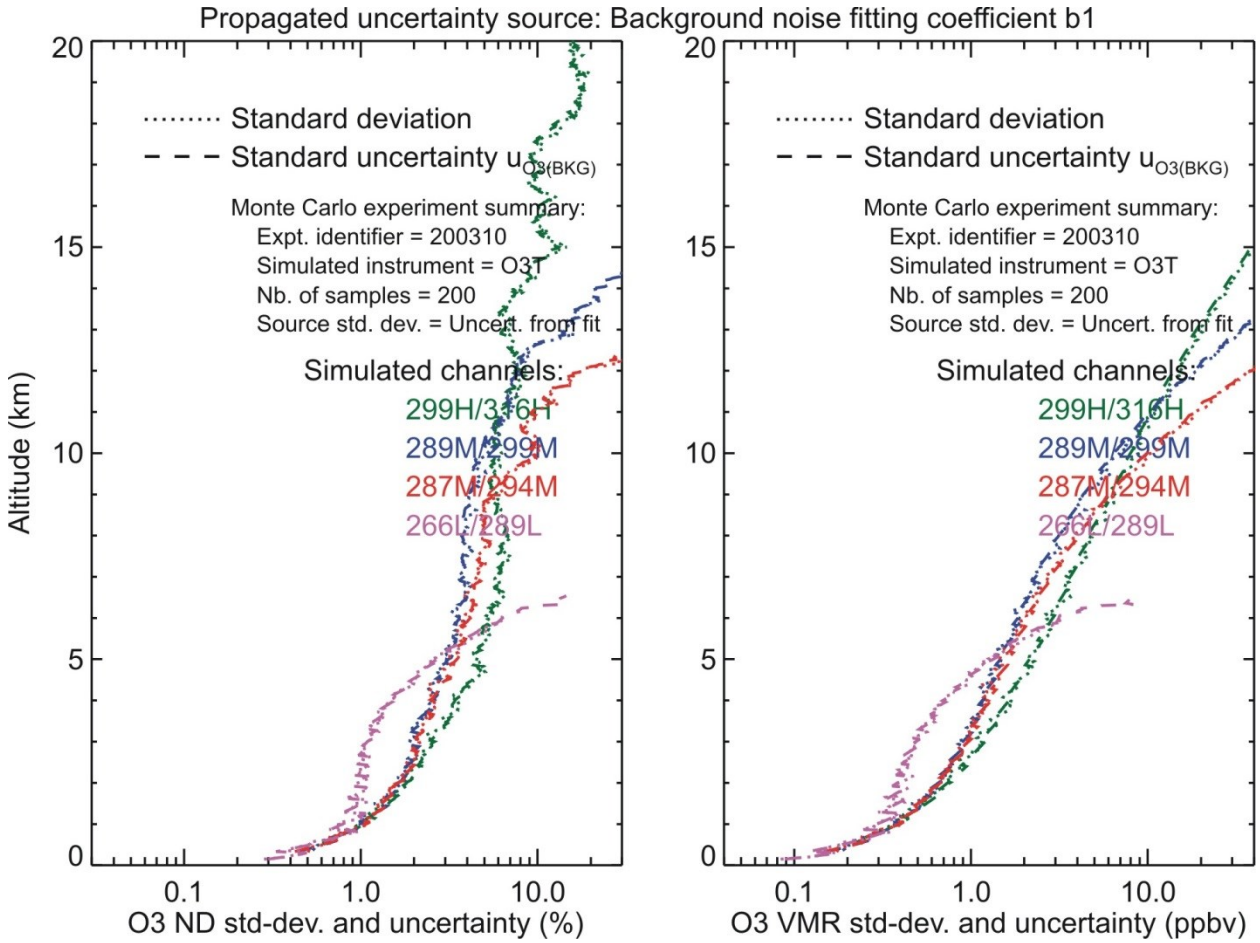


Figure B.4 Tropospheric ozone standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from a Monte Carlo experiment designed to quantify ozone uncertainty owed to background correction $u_{O3(BKG)}$ (see text for details)

Table B.3 Expressions used for background extraction uncertainty propagated to ozone

Sub-model description	*Propagated uncertainty*	*Eq. assuming independent hardware*	*Eq. assuming hardware shared*
Background extraction	$u_{S2(BKG)}$	(4.35)	(4.35)
Merging	$u_{S4(BKG)}$	(4.54)	(4.54)
Smoothing (lin) Smoothing (log)	$u_{S5(BKG)}$	(4.68) (4.76)	(4.68) (4.76)
$Log(S_{ON}/S_{OFF})$	$u_{S6(BKG)}$	(5.6)	(5.6)
Differentiation	$u_{S7(BKG)}$	(5.20)	(5.10)
DIAL equation	$u_{NO3(BKG)}$ $uq_{O3(BKG)}$	(5.27) (5.58)	(5.27) (5.58)

Similar experiments were performed with b_0 being the target parameter instead of b_1, and a perfect match between ozone standard deviation and calculated standard uncertainty was again observed (not shown).

B.4 Ozone cross-section differential uncertainty propagated to ozone

Uncertainty associated with the ozone absorption cross-section differential and propagated to ozone number density $u_{O3((\Delta\sigma O3)}$ is plotted in **Figure B**.5 for the stratospheric ozone system "O3S", and in **Figure B**.6 for the tropospheric ozone system "O3T". The corresponding propagation equations are listed in **Table B**.4**.** For these Monte-Carlo experiments, the signals were simulated with no detection noise, no background noise, no saturation effects, and no extinction terms. In the inverse model, a set of 200 ozone profiles were produced by varying the values of the ozone absorption cross-sections. Using the values and uncertainty estimates provided by University of Reims spectroscopy group (DMB), 200 normally-distributed cross-section perturbation values were used for all channels, with a standard deviation around the mean values of 2% for wavelengths in the Huggins band, 4% for wavelengths in the Hartley band, and 20% for wavelengths in the region of minimum absorption. The results from two configuration cases are shown. On the left hand side, it is assumed that the cross-section values at each wavelength (i.e., 308 nm, 332 nm, 355 nm and 387 nm for "O3S", and 266 nm, 287 nm, 289 nm, 294 nm, 299 nm, and 316 nm for "O3T") are independent from those at all other wavelengths, which corresponds to the propagation equations reported in the third (Rayleigh backscatter) and fourth (Raman backscatter) columns of **Table B**.4. The Monte Carlo experiment in this case consists of using two independent, normally-distributed populations of cross-section perturbations. On the right-hand side, it is assumed that all cross-sections come from the same laboratory measurements, and are assumed fully correlated, which corresponds to the propagation equation reported in the last column of **Table B**.4. The Monte Carlo experiment in this case consists of using the same normally-distributed population of cross-section perturbations for all wavelengths, which simulate a full correlation between the cross-sections at the varoiuos wavelengths. As expected from the DIAL equation, the uncertainty and standard

deviation relative values are constant with height because the ozone absorption cross-section relative perturbations were taken as constant with height. Once again, the standard uncertainty reported by the inverse model (long-dash curves) matches very well the calculated ozone standard deviation (dotted curves) for all ranges and all altitudes.

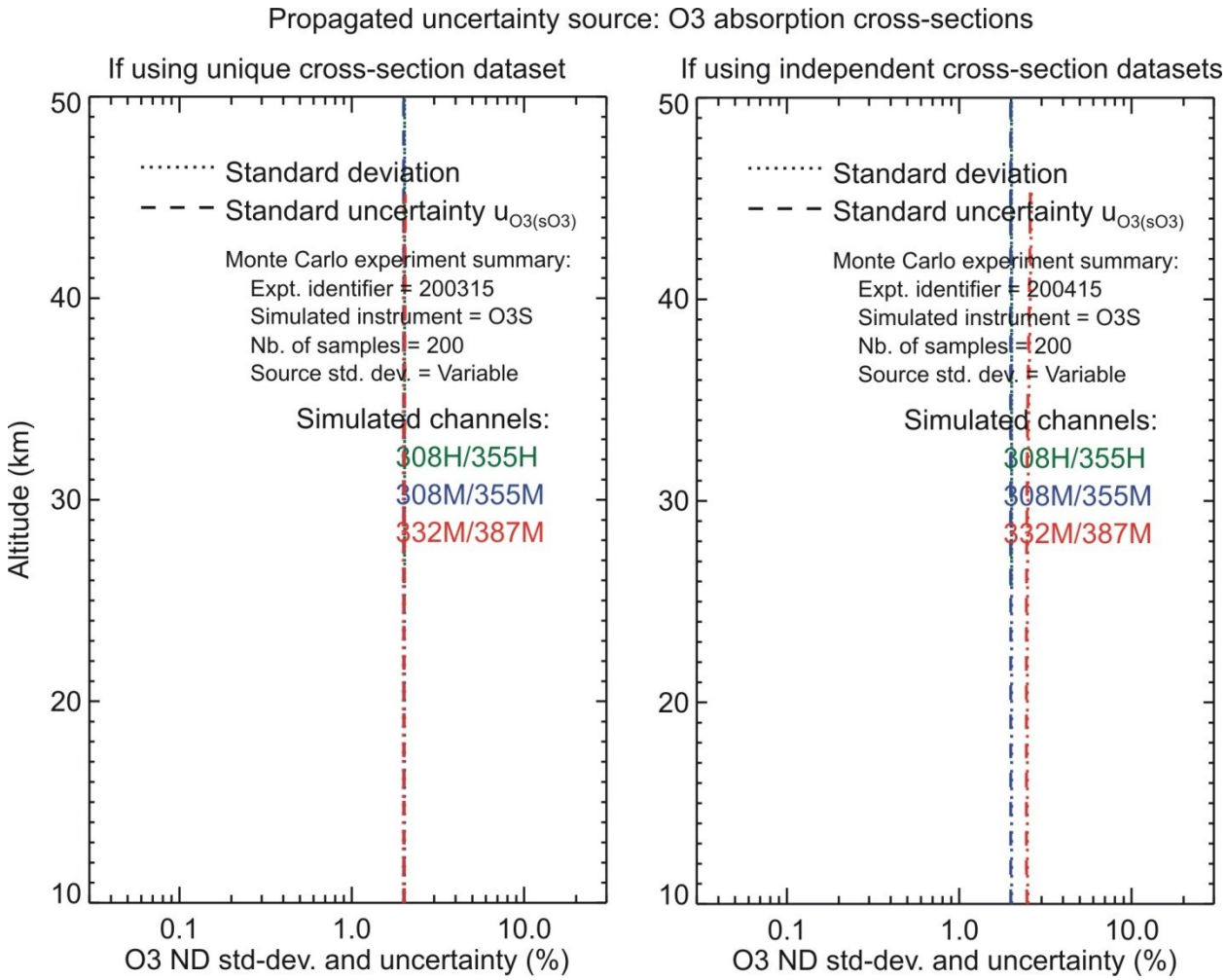


Figure B.5 Stratospheric ozone standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from Monte Carlo experiments designed to quantify ozone uncertainty owed to ozone absorption cross-section differential $u_{O3(\Delta\sigma O3)}$. Right-hand plots: Assuming that cross-sections at all wavelengths are independent; Left-hand plots: Assuming that the cross-sections at all wavelengths are fully correlated (see text for details)

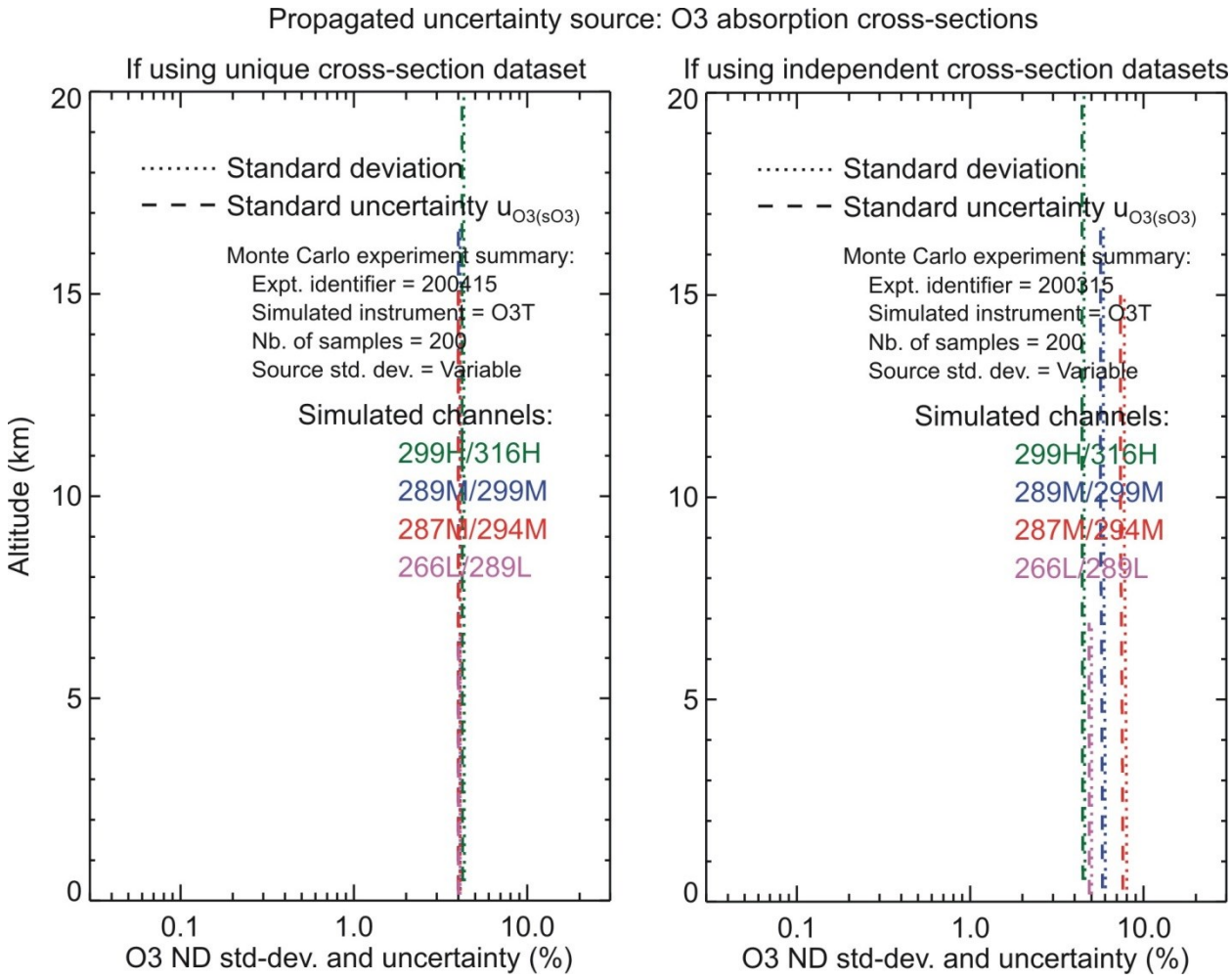


Figure B.6 Same as Figure B.5, but for the tropospheric ozone DIAL (O3T)

Table B.4 Expression used for O_3 absorption cross-section uncertainty propagated to ozone

Sub-model description	***Propagated uncertainty***	***Eq. if using all independent σ***	***Eq. if using all correlated σ***	***Eq. if using correlated σ within independent datasets***
Differential (Rayleigh) Differential (Raman)	$u_{\Delta\sigma O3}$	(5.33) (5.32)	(5.35) (5.34)	(5.37) (5.36)
DIAL equation	$u_{NO3(\Delta\sigma O3)}$ $u_{qO3(\Delta\sigma O3)}$	(5.38) (5.61)	(5.38) (5.61)	(5.38) (5.61)

B.5 Molecular extinction differential uncertainty propagated to ozone

Uncertainty associated with the Rayleigh cross-section differential and propagated to ozone number density $u_{O3((\Delta\sigma M)}$ is plotted in **Figure B**.7 for the stratospheric ozone system "O3S", and

in **Figure B**.8 for the tropospheric ozone system "O3T". The corresponding propagation equations are listed in **Table B**.5. For this Monte-Carlo experiment, the signals were simulated with no detection noise, no background noise, no saturation effects, and no absorption terms. In the inverse model, a set of 200 ozone profiles was produced by varying the values of the Rayleigh cross-sections. It is assumed that the cross-sections at all wavelengths (i.e., 308 nm, 332 nm, 355 nm and 387 nm for "O3S", and 266 nm, 287 nm, 289 nm, 294 nm, 299 nm, and 316 nm for "O3T") come from the same analytical formulae. Using the values and uncertainty estimates provided by Eberhard (2010), a single set of 200 normally-distributed cross-section values was used for all the channels, with a standard deviation around the mean values of 2% at all wavelengths. This time, as expected from the DIAL equation, the uncertainty and standard deviation absotute values are constant with height because the extinction cross-section relative perturbations were taken as constant with height (the extinction correction is an added term to the DIAL equation, not a multiplicative factor). The standard uncertainty reported by the inverse model (long-dash curves) matches again very well the calculated ozone standard deviation (dotted curves) for all ranges and all altitudes.

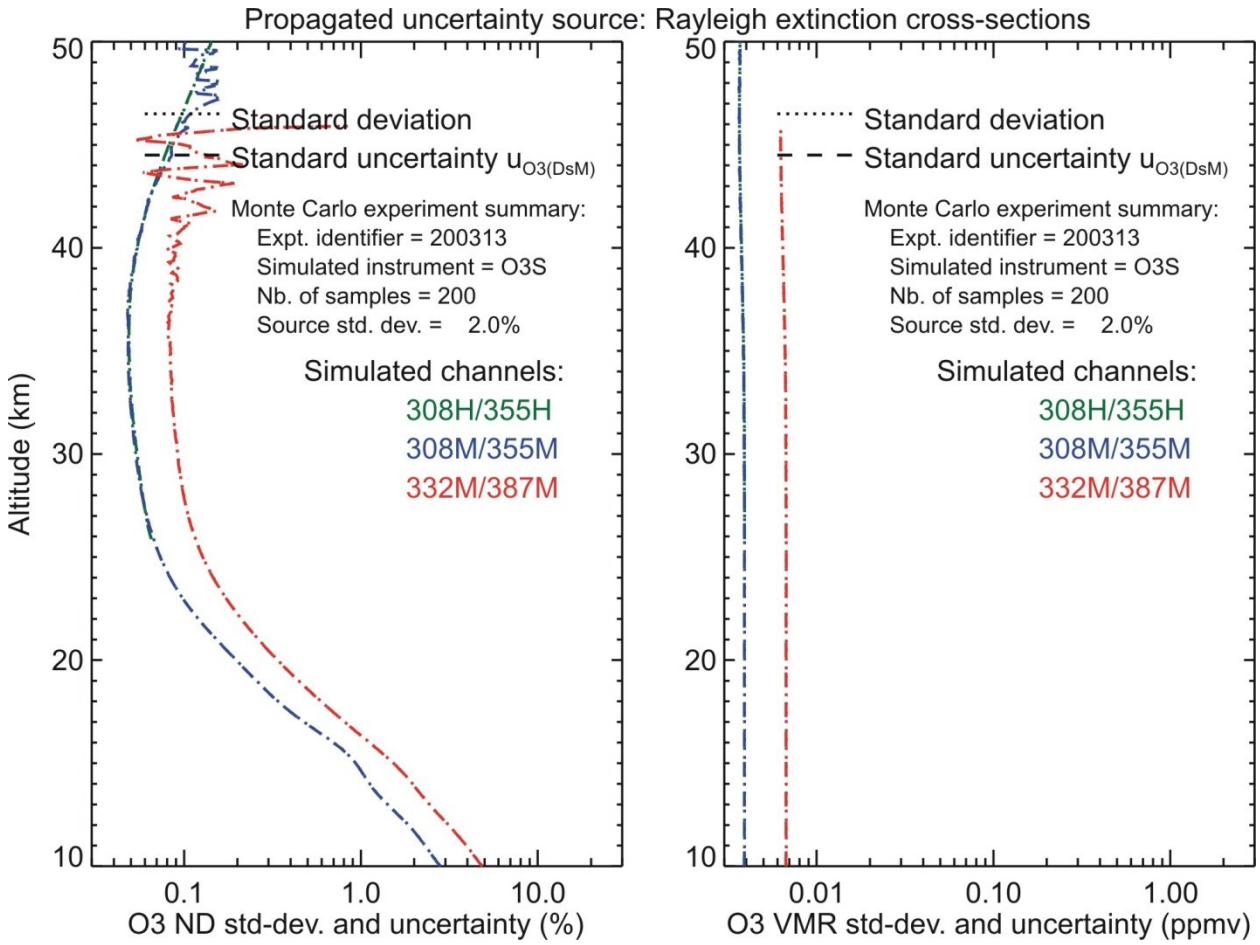


Figure B.7 Stratospheric ozone standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from a Monte Carlo experiment designed to quantify ozone uncertainty owed to Rayleigh extinction cross-sections $u_{O3(\Delta\sigma M)}$ (see text for details)

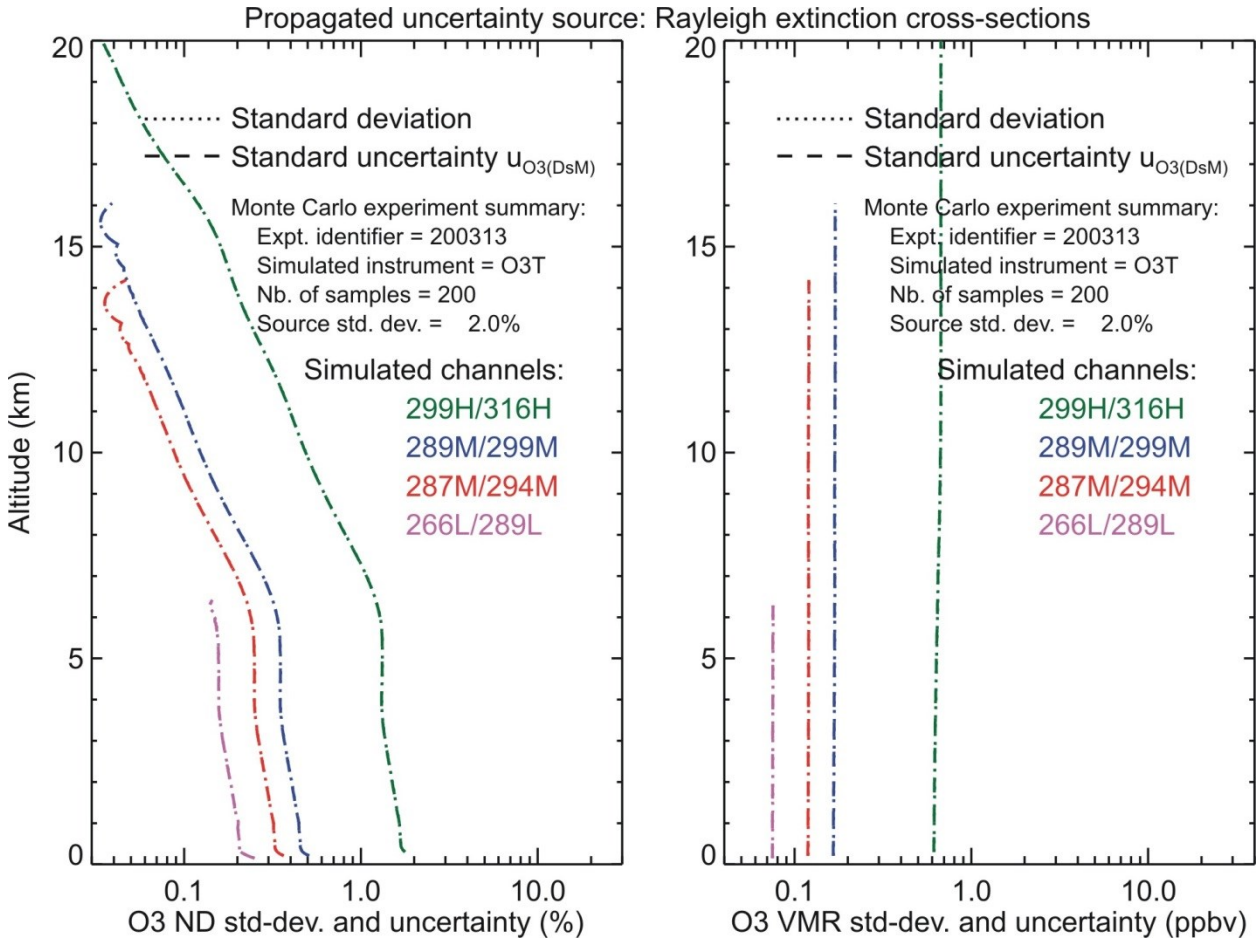


Figure B.8 Same as Figure B.7, but for tropospheric ozone DIAL (O3T)

Table B.5 Expression used for Rayleigh cross-section uncertainty propagated to ozone

Sub-model Description	***Propagated uncertainty***	***Eq. if using all correlated σ***
Differential (Rayleigh) Differnetial (Raman)	$u_{\Delta\sigma M}$	(5.44) (5.43)
DIAL equation	$u_{NO3(\Delta\sigma M)}$ (5.62(5.61)	(5.45) (5.62(5.61)

Uncertainty associated with the ancillary air number density and propagated to ozone number density $u_{O3((Na)}$ is plotted in **Figure B**.9 for the stratospheric ozone system "O3S", and in **Figure B**.10 for the tropospheric ozone system "O3T". The corresponding propagation equations are listed in **Table B**.6. For this Monte-Carlo experiment, the signals were simulated with no detection noise, no background noise, no saturation effects, and no absorption terms. In the inverse model, 200 ozone profiles were produced by varying the values of the ancillary air number density. A set of 200 normally-distributed air number density values was produced, with

a standard deviation around the mean equal to the ancillary air number density uncertainty. This uncertainty can vary significantly depending on the dataset source, which is reflected in **Figure B**.9. If the air number density is computed using pressure and temperature measurements from radiosonde, we should expect a small uncertainty deduced from a typical pressure uncertainty of 0.1 hPa and temperature uncertainty of 0.5 K. This applies to altitudes below 30 km, as shown for O3S in **Figure B**.9, and for O3T on the left hand plot of **Figure B**.10. If the air number density is computed using an analysis or reanalysis model such as NCEP or ECMWF, we should expect an uncertainty of up to 5%. If the air number density is computed using an empirical model such as MSISE-90 or CIRA, we should expect an air number density uncertainty of up to 10%. In the examples shown, this applies to altitudes above 30 km for O3S (**Figure B**.9), and at all altitudes for O3T on the right hand plot of **Figure B**.10. The ozone standard uncertainty reported by the inverse model (long-dash curves) matches again very well the calculated ozone standard deviation (dotted curves) for all ranges and all altitudes.

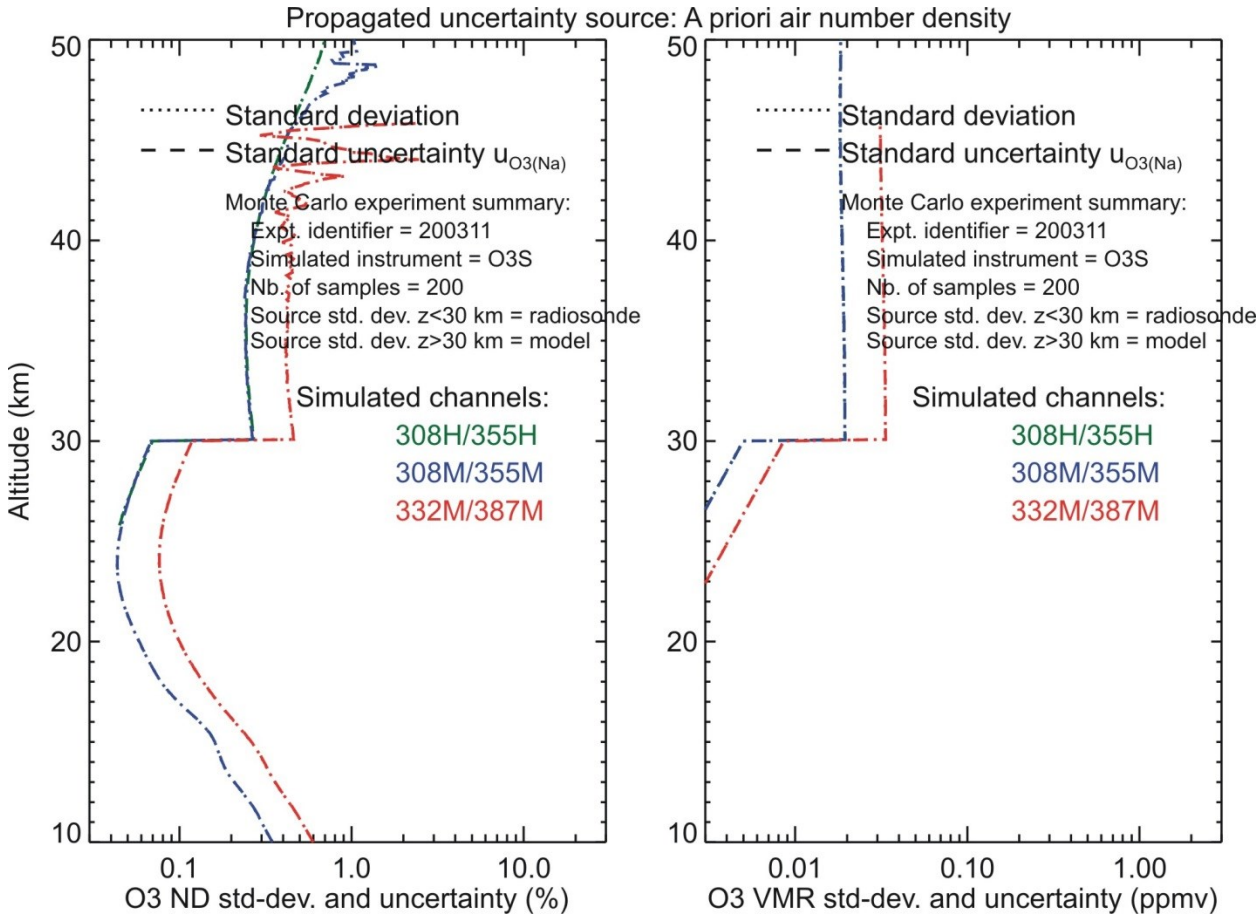


Figure B.9 Stratospheric ozone standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from a Monte Carlo experiment designed to quantify ozone uncertainty owed to air number density $u_{O3(Na)}$ (see text for details)

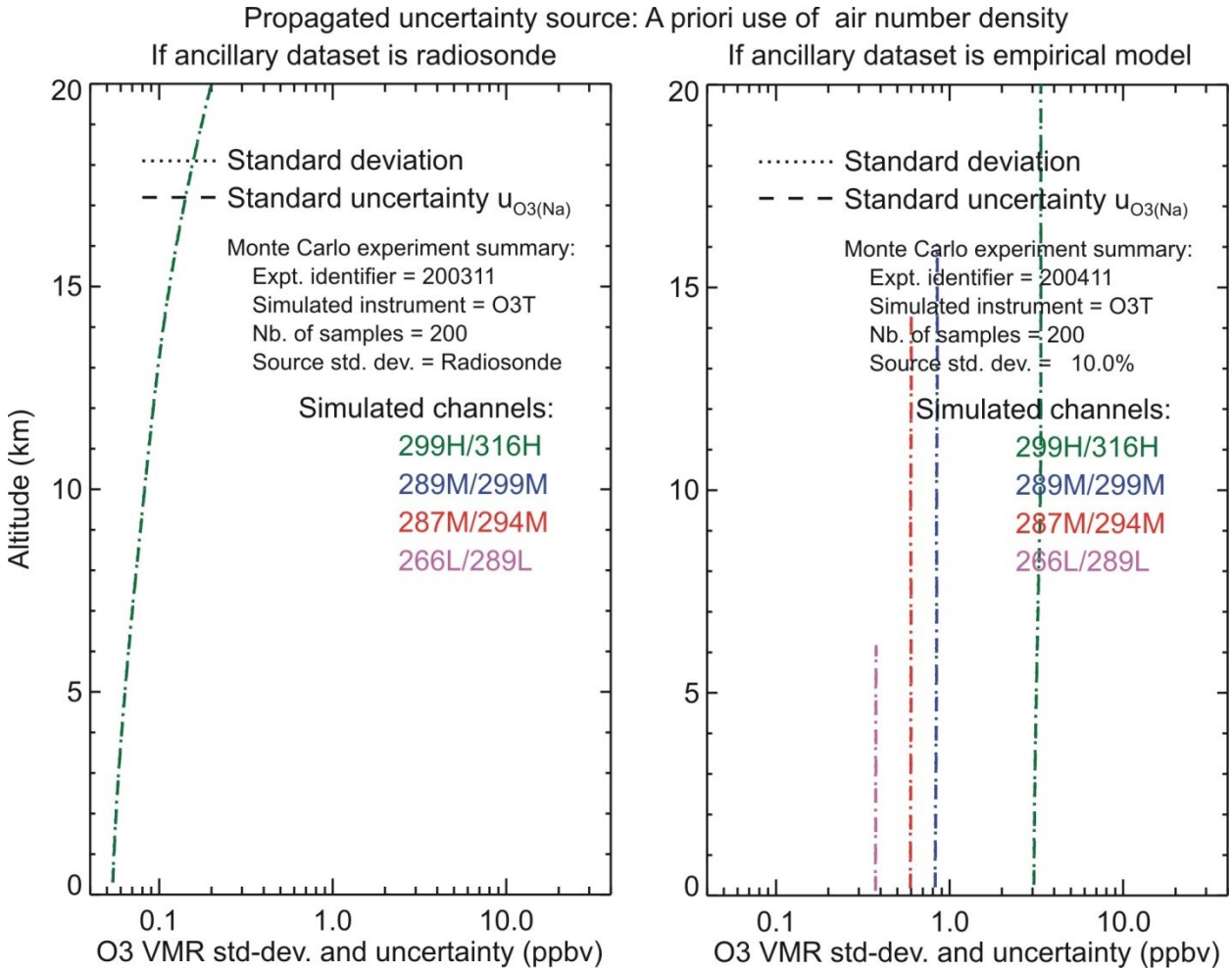


Figure B.10 Tropospheric ozone standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from Monte Carlo experiments designed to quantify ozone uncertainty owed to air number density $u_{O3(Na)}$. Left hand plots: If air number density derived from radiosonde; Right hand plot: if air number density comes from an empirical model (see text for details)

Table B.6 Expression used for ancillary air number density uncertainty propagated to ozone

Sub-model Description	***Propagated Uncertainty***	***Eq. If air pressure and air temperature independent***	***Eq. If air pressure and air temperature correlated***
Air number density	u_{Na}	(5.49)	(5.50)
DIAL equation	$u_{NO3(Na)}$ $u_{qO3(Na)}$	(5.51) (5.66)	(5.51) (5.66)

B.6 NO_2 and SO_2 absorption differential uncertainty propagated to ozone

Uncertainty associated with the NO_2 absorption cross-section differential and propagated to ozone number density $u_{O3((\Delta\sigma NO2)}$ is plotted in **Figure B.11** for the tropospheric ozone system "O3T". The corresponding propagation equations are listed in **Table B**.7. For this Monte-Carlo experiment, the signals were simulated with no detection noise, no background noise, no

saturation effects, no extinction terms except NO_2 absorption. In the inverse model, a set of 200 ozone profiles was produced by varying the values of the NO_2 absorption cross-sections. In the example shown, it is assumed that all cross-sections come from the same laboratory measurements, which corresponds to the propagation equation reported in the last column of **Table B**.7. Using the values and uncertainty estimates provided by University of Bremen spectroscopy group (Bogumil), a set of 200 normally-distributed cross-section values was used for all the channels, with a standard deviation around the mean values of 5% for all wavelengths. The standard uncertainty reported by the inverse model (long-dash curves) matches again very well the calculated ozone standard deviation (dotted curves) for all ranges and all altitudes.

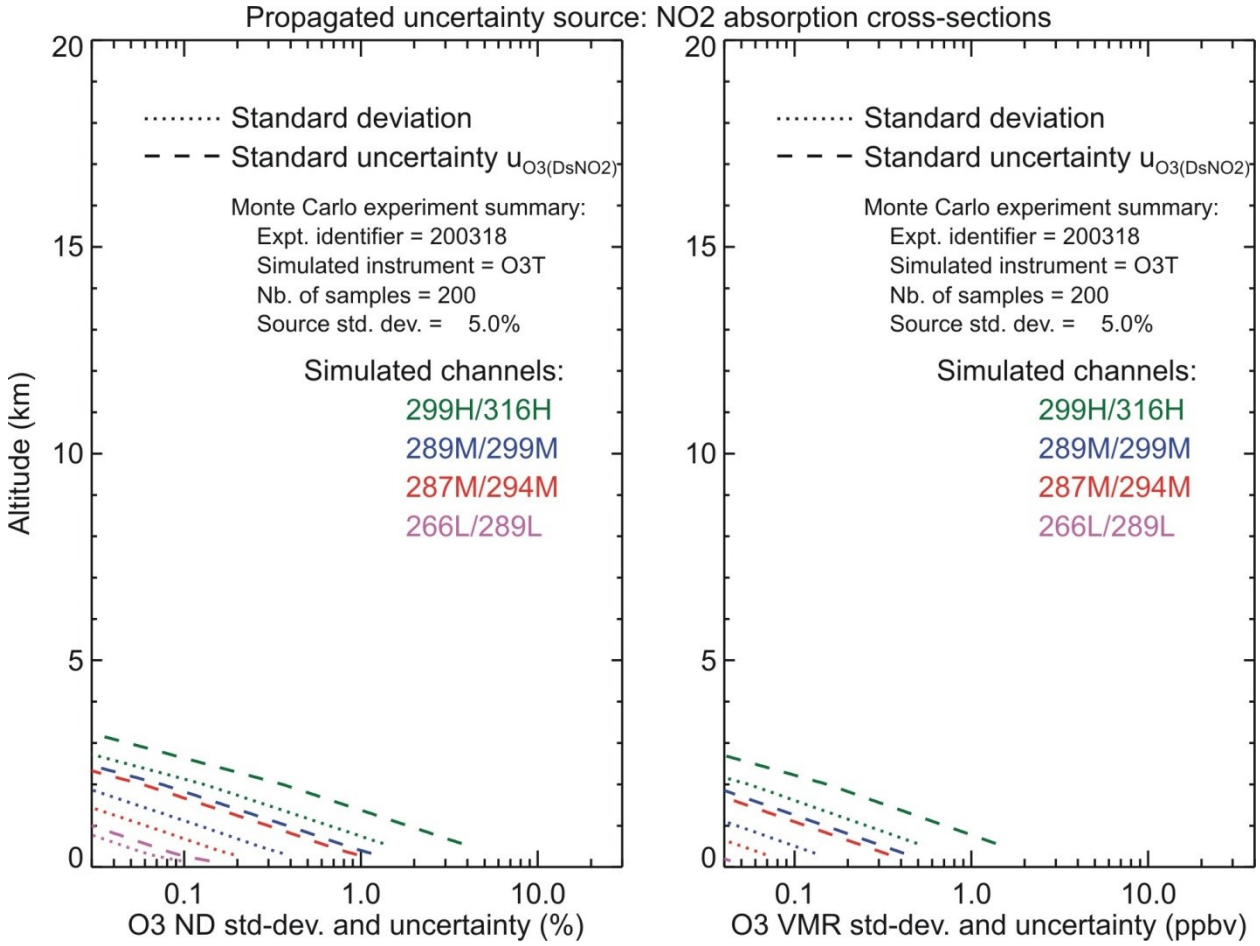


Figure B.11 Tropopsheric ozone standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from a Monte Carlo experiment designed to quantify ozone uncertainty owed to NO_2 absorption cross-section differential $u_{O3(\square\square NO2)}$ (see text for details)

Table B.7 Expression used for NO_2 absorption cross-section uncertainty propagated to ozone

Sub-model description	*Propagated uncertainty*	*Eq. if using all independent σ*	*Eq. if using all correlated σ*	*Eq. if using correlated σ within independent datasets*
Differential (Rayleigh) Differential (Raman)	$u_{\Delta\sigma NO2}$	(5.33)* (5.32)*	(5.35)* (5.34)*	(5.37)* (5.36)*
DIAL equation	$u_{NO3(\Delta\sigma NO2)}$ $u_{qO3(\Delta\sigma NO2)}$	(5.39) (5.63)	(5.39) (5.63)	(5.39) (5.63)

* Same equation as for O_3 but applied to NO_2

Uncertainty associated with the ancillary NO_2 number density and propagated to ozone number density $u_{O3((NO2)}$ is plotted in **Figure B**.12 for the stratospheric ozone system "O3S", and in **Figure B**.13 for the tropospheric ozone system "O3T". The corresponding propagation equation is listed in **Table B**.8. For this Monte-Carlo experiment, the signals were simulated with no detection noise, no background noise, no saturation effects, and no extinction terms except NO_2. In the inverse model, 200 ozone profiles were produced by varying the values of the ancillary NO_2 number density. A set of 200 normally-distributed air number density values was produced, with a standard deviation around the mean equal to the ancillary NO_2 number density uncertainty. This uncertainty can vary significantly depending on the dataset source. In our example, the number density profiles are taken from SCIAMACHY measurements in the stratosphere (Bracher et al., 2005) and from the "worst-case scenario" of an heavily-polluted boundary layer (0-3 km) (Cao et al., 2006; Miyazaki et al., 2012a; 2012b). The associated uncertainty is 10%. The ozone standard uncertainty reported by the inverse model (long-dash curves) matches again very well the calculated ozone standard deviation (dotted curves) for all ranges and all altitudes.

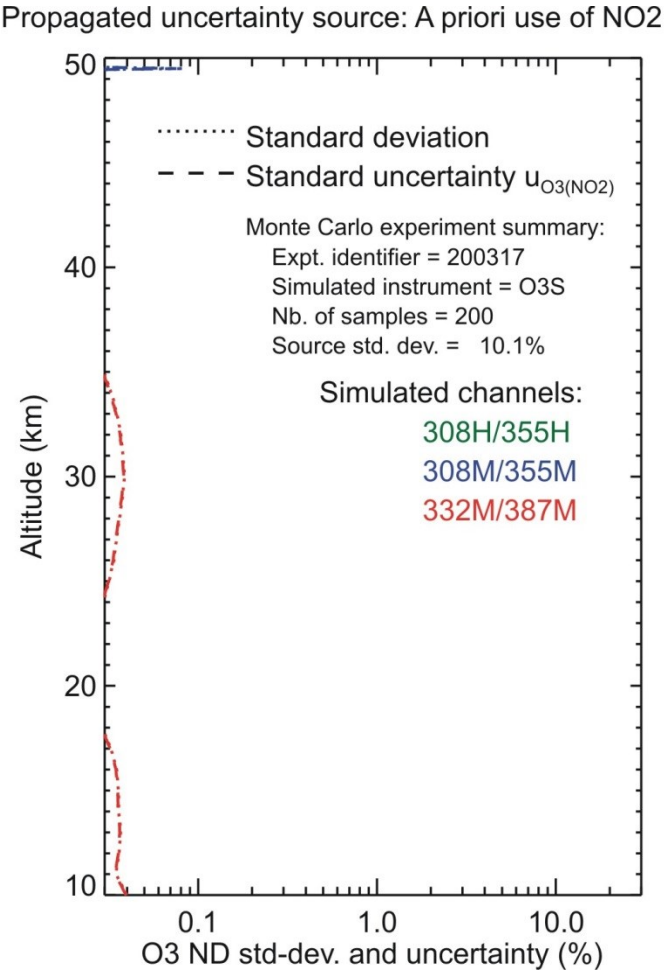


Figure B.12 Stratospheric ozone standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from a Monte Carlo experiment designed to quantify ozone uncertainty owed to NO_2 number density $u_{O3(NO2)}$ (see text for details)

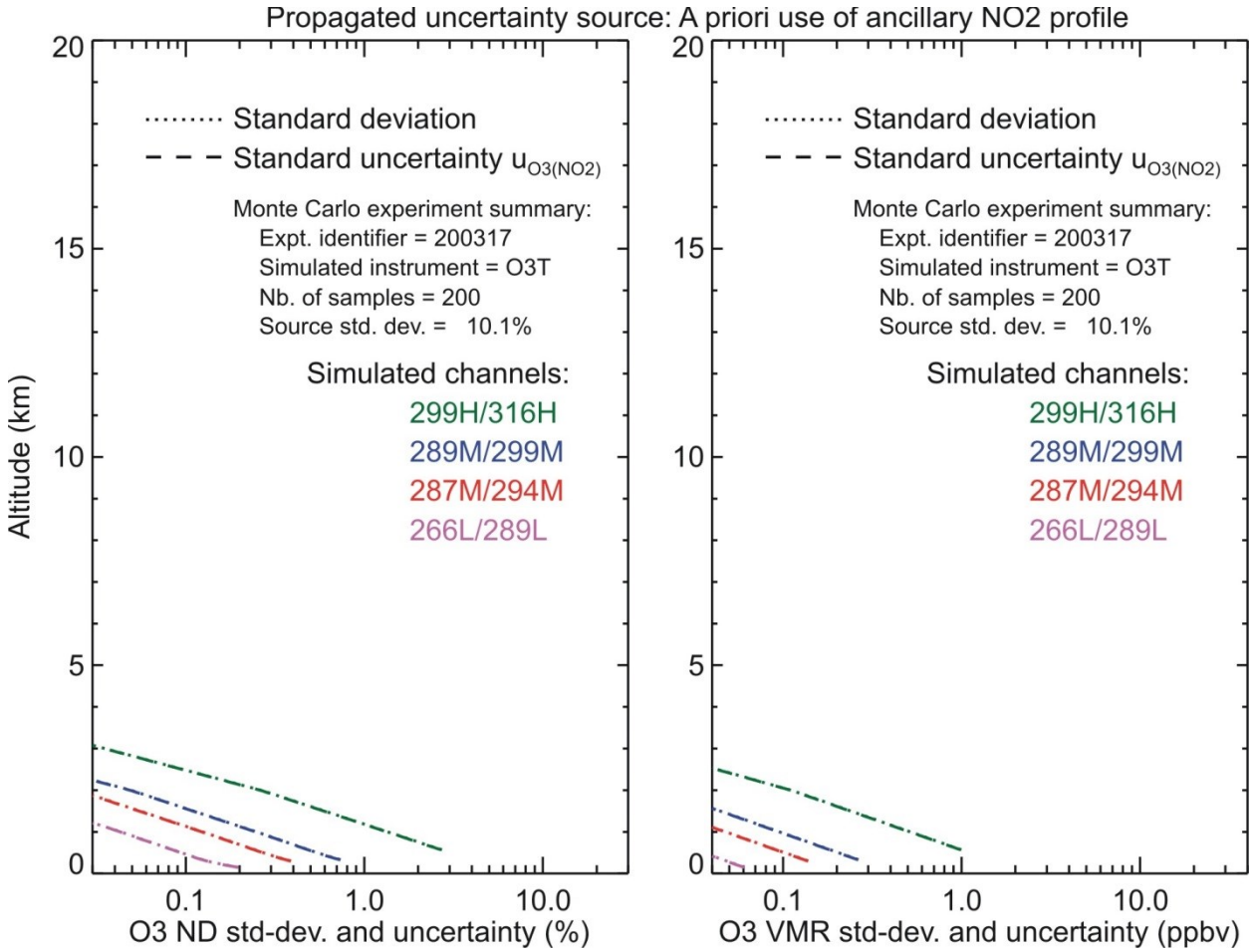


Figure B.13 Same as Figure B.12, but for tropospheric ozone DIAL (O3T)

Table B.8 Expression used for ancillary NO_2 number density or mixing ratio uncertainty propagated to ozone

Sub-model Description	***Propagated uncertainty***	***Eq.***
DIAL equation (using number density)	$u_{NO3(NNO2)}$ $u_{NO3(qNO2)}$	(5.52) (5.72)
DIAL equation (using mixing ratio)	$uq_{O3(NNO2)}$ $uq_{O3(qNO2)}$	(5.67) (5.76)

Uncertainty associated with the SO_2 absorption cross-section differential and propagated to ozone number density $u_{O3((\Delta\sigma SO2)}$ is plotted in **Figure B**.14 for the tropospheric ozone system "O3T". The corresponding propagation equations are listed in **Table B**.9. For this Monte-Carlo experiment, the signals were simulated with no detection noise, no background noise, no saturation effects, no extinction terms except SO_2 absorption. In the inverse model, a set of 200 ozone profiles was produced by varying the values of the SO_2 absorption cross-sections. In the example shown, it is assumed that all cross-sections come from the same laboratory

measurements, which corresponds to the propagation equation reported in the last column of **Table B**.9. Using the values and uncertainty estimates provided by University of Bremen spectroscopy group (Bogumil), a set of 200 normally-distributed cross-section values was used for all the channels, with a standard deviation around the mean values of 5% for all wavelengths. The standard uncertainty reported by the inverse model (long-dash curves) matches again very well the calculated ozone standard deviation (dotted curves) for all ranges and all altitudes.

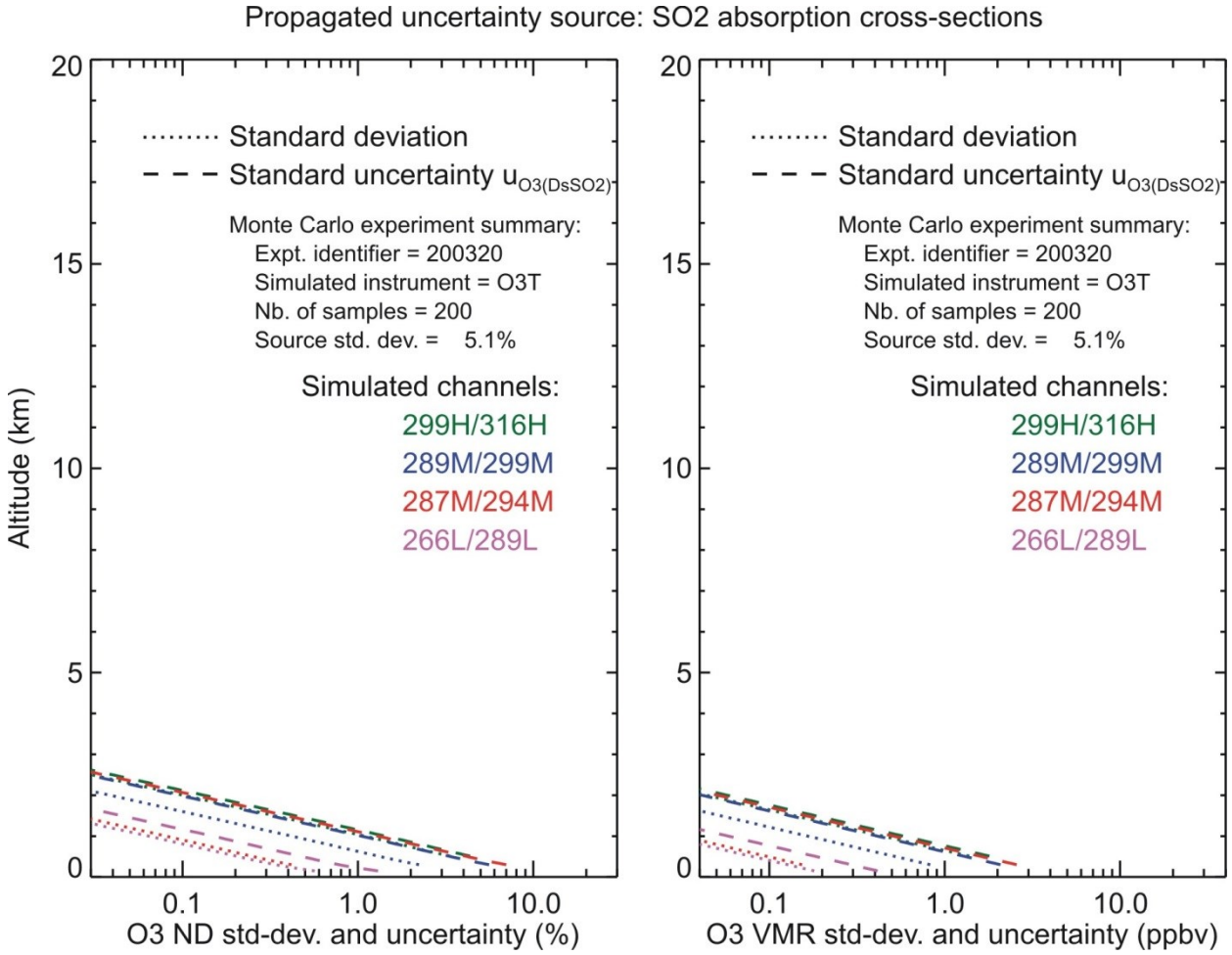


Figure B.14 Tropospheric ozone standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from a Monte Carlo experiment designed to quantify ozone uncertainty owed to SO2 absorption cross-section differential $u_{O3(\Delta\sigma SO2)}$ (see text for details)

Table B.9 Expression used for SO_2 absorption cross-section uncertainty propagated to ozone

Sub-model description	*Propagated uncertainty*	*Eq. if using all independent σ*	*Eq. if using all correlated σ*	*Eq. if using correlated σ within independent datasets*
Differential (Rayleigh) Differential (Raman)	$u_{\Delta\sigma SO2}$	(5.33)* (5.32)*	(5.35)* (5.34)*	(5.37)* (5.36)*
DIAL equation	$u_{NO3(\Delta\sigma SO2)}$ $u_{qO3(\Delta\sigma SO2)}$	(5.40) (5.64)	(5.40) (5.64)	(5.40) (5.64)

* Same equation as for O_3 but applied to SO_2

Uncertainty associated with the ancillary SO_2 number density and propagated to ozone number density $u_{O3((\Delta\sigma NO2)}$ is plotted in **Figure B**.15 for the tropopsheric ozone system "O3T". The corresponding propagation equation is listed in **Table B**.10. For this Monte-Carlo experiment, the signals were simulated with no detection noise, no background noise, no saturation effects, and no extinction terms except SO_2. In the inverse model, 200 ozone profiles were produced by varying the values of the ancillary SO_2 number density. A set of 200 normally-distributed air number density values was produced, with a standard deviation around the mean equal to the ancillary SO_2 number density uncertainty. This uncertainty can vary significantly depending on the dataset source. In our example, the number density profiles are taken from a MIPAS climatology in the stratosphere (Hopfner et al., 2013) and from the "worst-case scenario" of an heavily-polluted boundary layer (0-3 km) (McLinden et al., 2014). The corresponding uncertainty is 10%. The ozone standard uncertainty reported by the inverse model (long-dash curves) matches again very well the calculated ozone standard deviation (dotted curves) for all ranges and all altitudes.

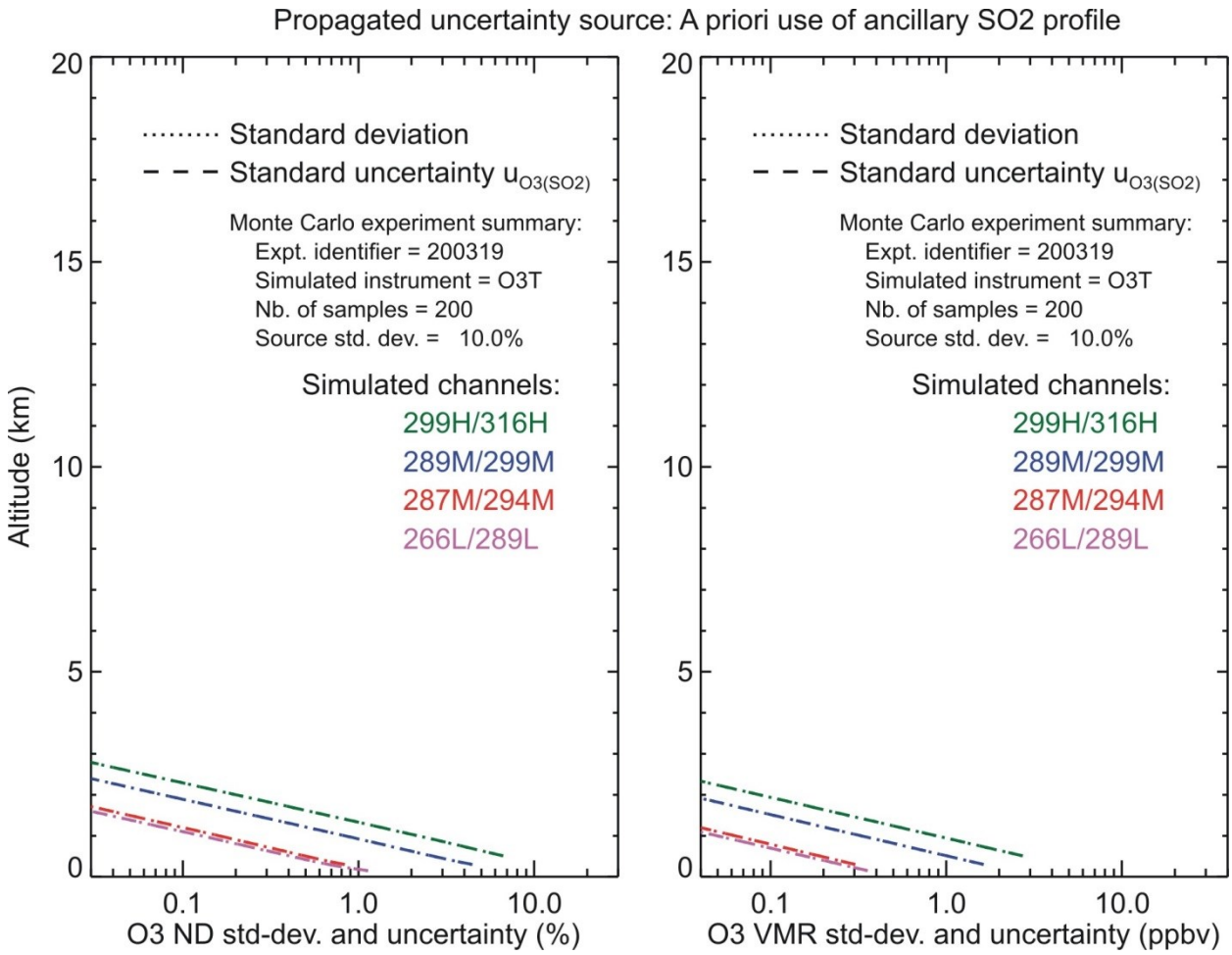


Figure B.15 Tropospheric ozone standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from a Monte Carlo experiment designed to quantify ozone uncertainty owed to SO2 number density $u_{O3(SO2)}$ (see text for details)

Table B.10 Expression used for ancillary SO_2 number density or mixing ratio uncertainty propagated to ozone

Sub-model Description	***Propagated uncertainty***	***Eq.***
DIAL equation (using number density)	$u_{NO3(NSO2)}$ $u_{NO3(qSO2)}$	(5.53) (5.73)
DIAL equation (using mixing ratio)	$uq_{O3(NSO2)}$ $uq_{O3(qSO2)}$	(5.68) (5.77)

B.7 O_2 absorption differential uncertainty propagated to ozone

Uncertainty associated with the O_2 absorption cross-section differential and propagated to ozone number density $u_{O3((\Delta\sigma O2)}$ is plotted in **Figure B**.16 for the tropospheric ozone system "O3T". The corresponding propagation equations are listed in **Table B**.11. For this Monte-Carlo

experiment, the signals were simulated with no detection noise, no background noise, no saturation effects, no extinction terms except O_2 absorption. In the inverse model, a set of 200 ozone profiles were produced by varying the values of the O_2 absorption cross-sections. In the example shown, it is assumed that all cross-sections come from the same laboratory measurements, which corresponds to the propagation equation reported in the last column of **Table B**.11. Using the values and uncertainty estimates provided by the Institut d'Aéronomie Spatiale de Belgique (IASB) spectroscopy group, a single set of 200 normally-distributed cross-section values was used for all the channels, with a standard deviation around the mean values of 5% for all wavelengths. No Monte Carlo experiment was needed for the stratospheric ozone DIAL system O3S because this absorption occurs at wavelengths shorter than 294 nm. The standard uncertainty reported by the inverse model (long-dash curves) matches again very well the calculated ozone standard deviation (dotted curves) for all ranges and all altitudes.

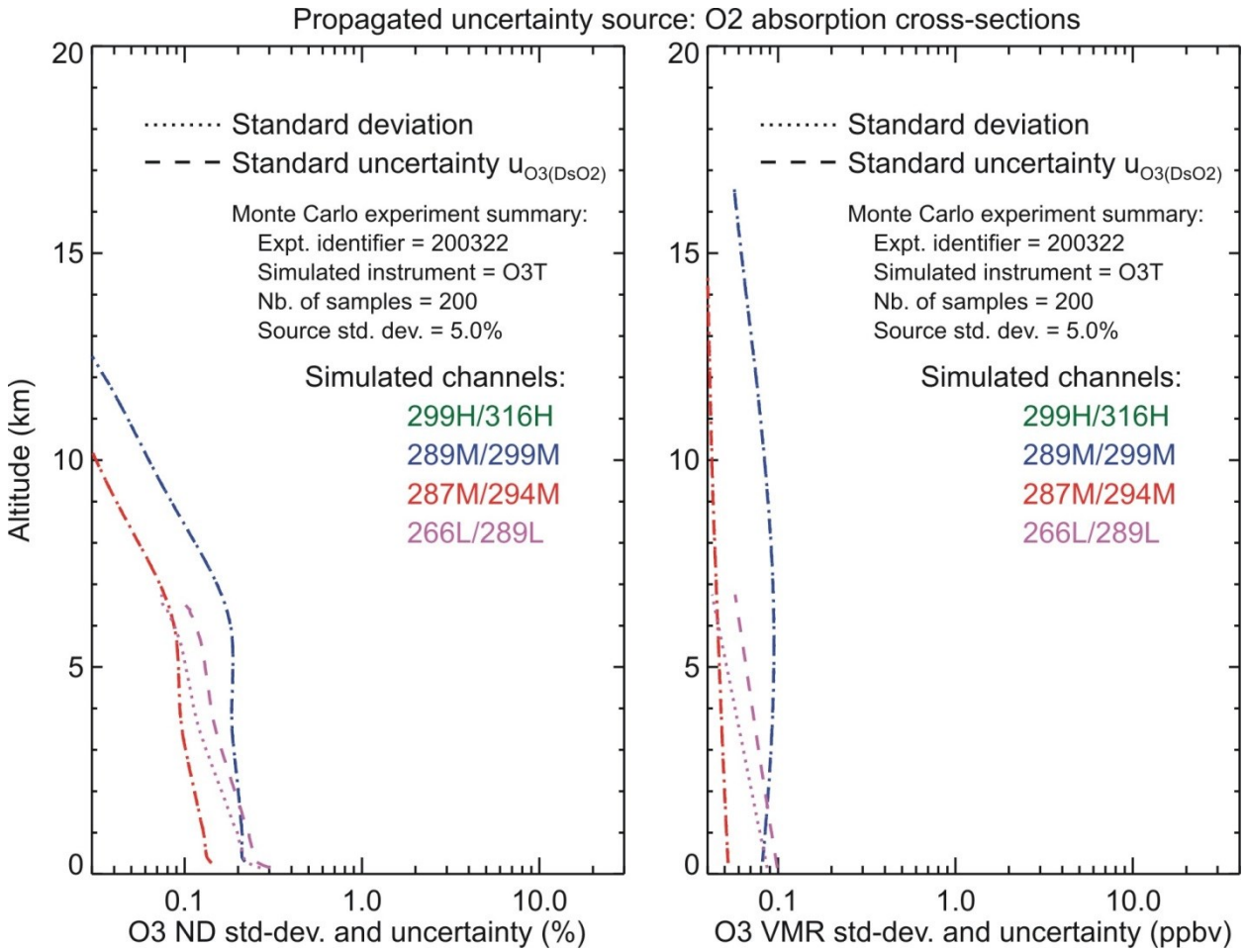


Figure B.16 Tropospheric ozone standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from a Monte Carlo experiment designed to quantify ozone uncertainty owed to O2 absorption cross-section differential $u_{O3(\Delta\sigma O2)}$ (see text for details)

3509 **Table B.11 Expression used for O_2 absorption cross-section uncertainty propagated to ozone**

Sub-model description	*Propagated uncertainty*	*Eq. if using all independent σ*	*Eq. if using all correlated σ*	*Eq. if using correlated σ within independent datasets*
Differential (Rayleigh) Differential (Raman)	$u_{\Delta\sigma O2}$	(5.33)* (5.32)*	(5.35)* (5.34)*	(5.37)* (5.36)*
DIAL equation	$u_{NO3(\Delta\sigma O2)}$ $u_{qO3(\Delta\sigma O2)}$	(5.41) (5.65)	(5.41) (5.65)	(5.41) (5.65)

3510

3511

3512 ## B.8 Joint propagation of all ozone uncertainty components

3513 Now that all individual components have been properly quantified using their propagation
3514 expressions listed in **Table B**.1**-Table B**.11 it is time to verify that each individual component is
3515 independent from the others, so that we can derive the ozone combined standard uncertainty by
3516 computing the quadratic sum of the individual components. To do this, we generalize the Monte
3517 Carlo experiments presented so far for an individual component to an experiment in which all
3518 sources of uncertainty are applied and propagated simultaneously. For each input quantity, we
3519 compute a population of 200 normally-distributed values of this quantity with a standard
3520 deviation equal to the quantity's uncertainty. Each produced population is "orthogonal" to the
3521 others, i.e., every set of normally-distributed values is independent from the others (correlation
3522 coefficient=0). The ozone number density combined uncertainty u_{O3} as computed using **Eq.**
3523 **(5.54)** is plotted in **Figure B**.17 for the stratospheric ozone system "O3S", and in **Figure B**.18
3524 for the tropospheric ozone system "O3T". The ozone combined standard uncertainty reported by
3525 the inverse model (long-dash curves) matches again very well the calculated ozone standard
3526 deviation (dotted curves) for all ranges and all altitudes, demonstrating that the input quantities
3527 introduced in the sub-models of **chapters 4-5** are uncorrelated and can be propagated in parallel
3528 throughout the ozone data processing chain.

3529

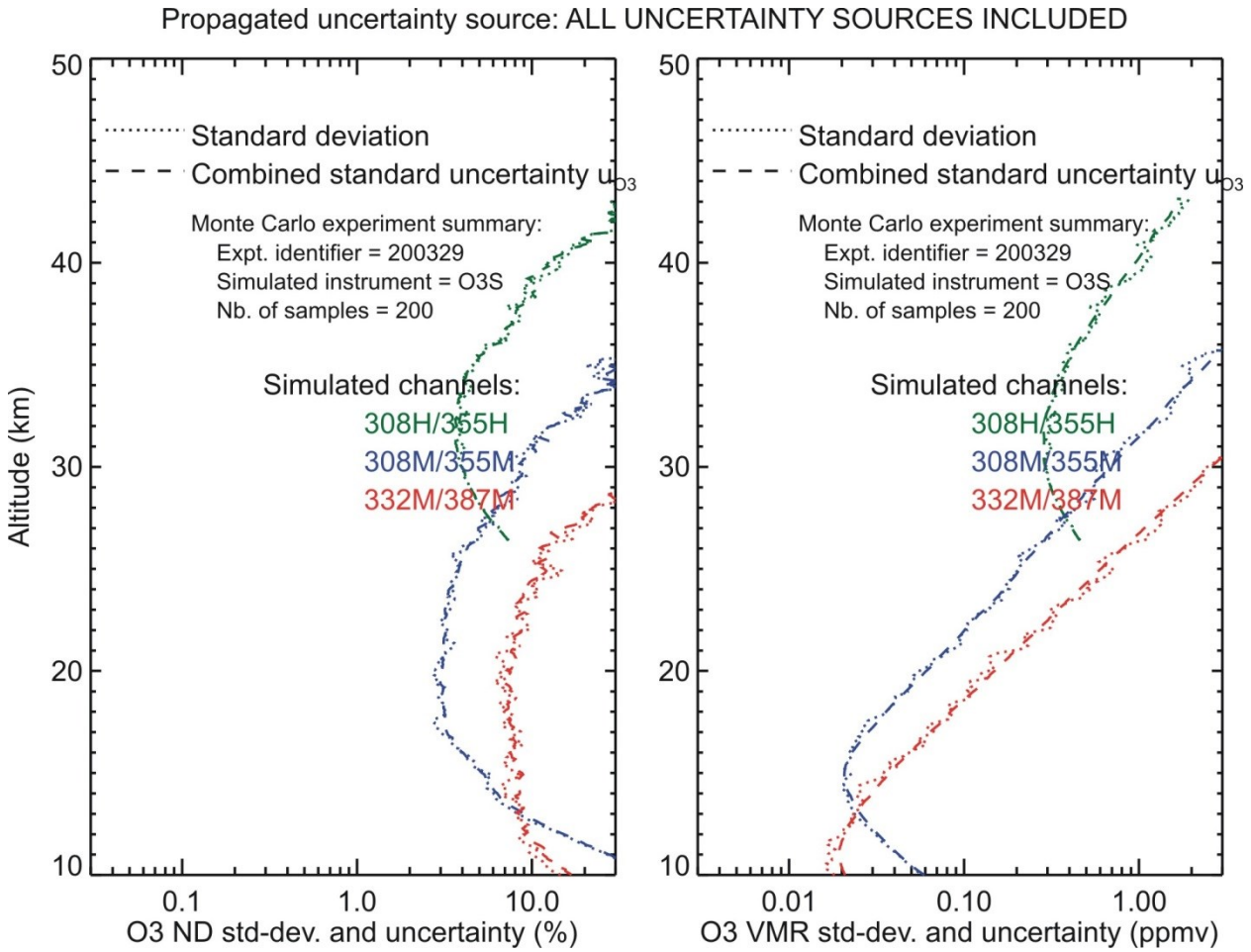


Figure B.17 Stratospheric ozone standard deviation (dotted curves) and combined standard uncertainty (long-dash curves) obtained from a Monte Carlo experiment designed to quantify ozone uncertainty owed to all the sources included in this appendix (see text for details)

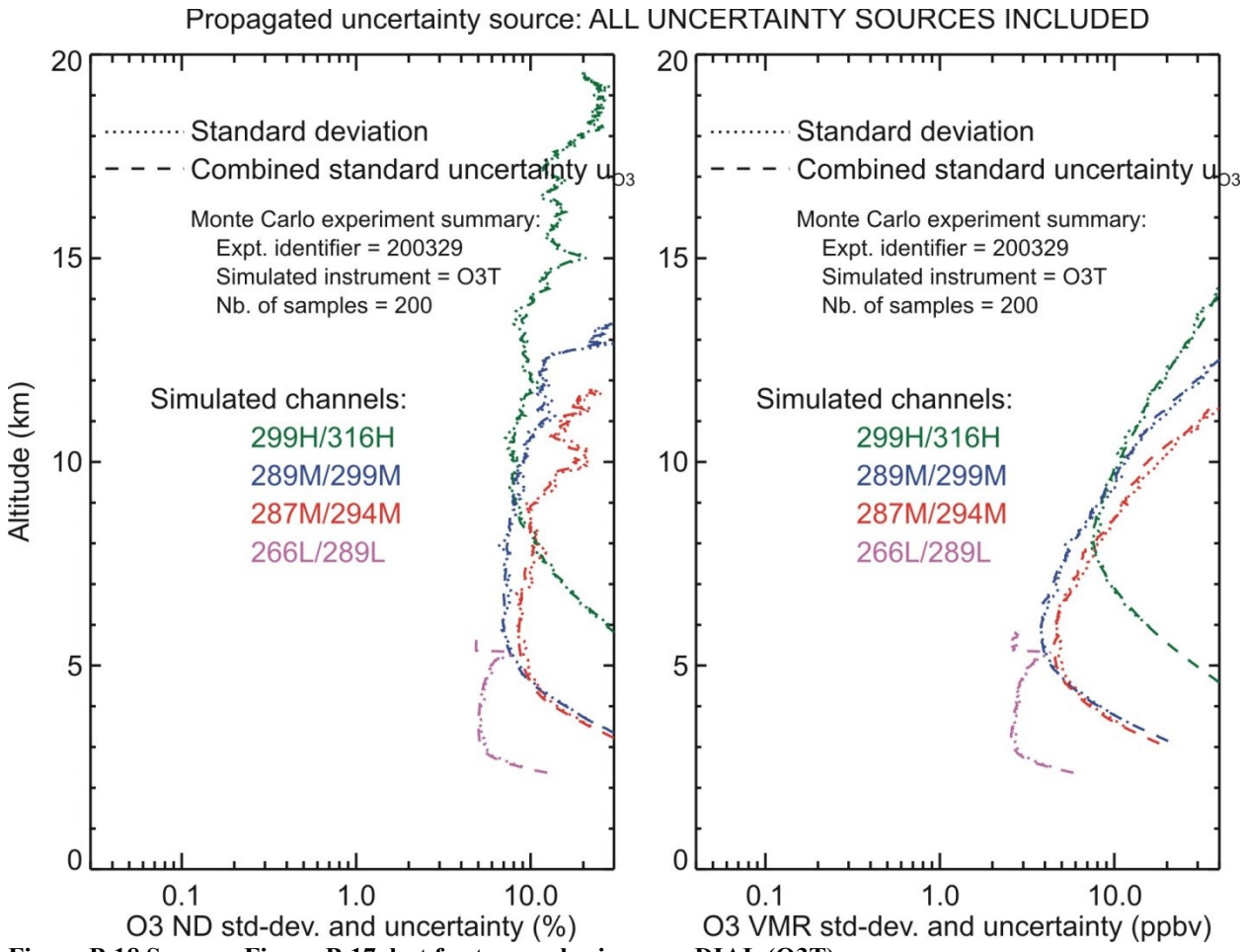


Figure B.18 Same as Figure B.17, but for tropospheric ozone DIAL (O3T)

C Quantitative validation of uncertainty propagated to temperature

For each uncertainty source introduced in the temperature measurement sub-models of **chapter 4 and chapter 6**, we now validate the appropriate use of the propagation expressions provided in these **chapters**. We show that the propagated temperature standard uncertainty calculated in parallel for each component quantitatively matches the temperature standard deviation calculated for the corresponding dedicated Monte-Carlo experiment.

In each of the experiments described thereafter, 200 simulated lidar datasets are used, each dataset being equivalent to a 5-min accumulation of photocounts. The instrumental parameters used for the simulations are those listed in **Table A**.1 for the "TMP" simulated lidar intrument. A boxcar low-pass filter with a vertical width increasing with altitude is applied to the signals. This filtering procedure is typical of temperature retrievals. It was applied to all Monte Carlo experiments, even those producing simulated signals without detection noise, essentially to verify that uncertainty components that yield correlated terms in altitude are properly propagated. Therefore, for all the examples shown, a matching standard deviation and calculated uncertainty implies the correct use of the equations written in **section 4.6** (smoothing).

C.1 Detection noise uncertainty propagated to temperature

Uncertainty associated with detection noise and propagated to temperature $u_{T(DET)}$ is plotted for the simulated system "TMP" in **Figure C**.1 (dashed curves). The corresponding propagation equations are listed in **Table C**.1. The dotted curves show the temperature standard deviation resulting from a Monte-Carlo experiment with detection noise only (i.e., no saturation, no background noise, no extinction terms). Detection noise was simulated using a typical Poisson distribution around the mean number of photons detected with no correlation between altitude bins, no correlation between simulated channels, and no correlation between any of the 200 simulated datasets (independent datasets). The results are presented in percent (left) and in degree Kelvin (right) as a function of signal-to-noise ratio. Choosing signal-to-noise ratio (specifically detection noise) instead of altitude as the independent variable allows to bring the curves from all six channels together into a consistent behavior. The fact that the dotted curves (standard deviation) and dashed curves (calculated uncertainty) match perfectly demonstrates that the set of equations listed in **Table C**.1 for the propagation of uncertainty owed to detection noise are all correct.

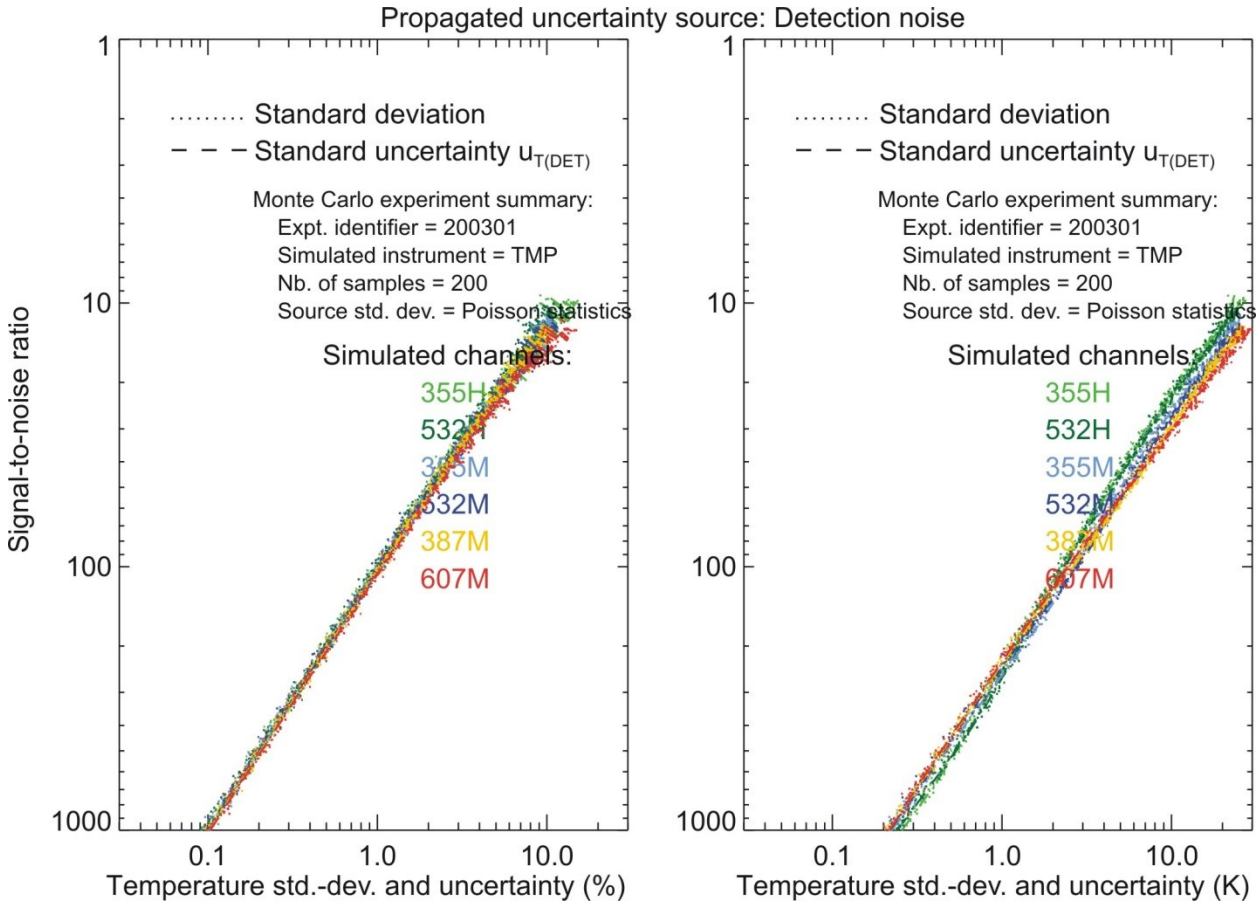


Figure C.1 Temperature standard uncertainty owed to detection noise $u_{T(DET)}$ (long-dash curves) and standard deviation (dotted curves) obtained from the corresponding dedicated Monte Carlo experiment (see text for details)

Table C.1 Expressions used for detection noise uncertainty propagated to temperature

Sub-model description	*Propagated uncertainty*	*Eq. Non-paralyz.*	*Eq. Paralyz.*
Signal detection (PC)	$u_{S0(DET)}$	(4.3)	(4.3)
Saturation correction	$u_{S1(DET)}$	(4.14)	(H.9)
Background extraction	$u_{S2(DET)}$	(4.37)	(4.37)
Merging	$u_{S4(DET)}$	(4.52)	(4.52)
Smoothing (lin) Smoothing (log)	$u_{S5(DET)}$	(4.66) (4.74)	(4.66) (4.74)
Range correction	$u_{S6(DET)}$	(6.2)	(6.2)
Extinction correction	$u_{S7(DET)}$	(6.10)	(6.10)
Layer-averaged signal and gravity	$u_{S8(DET)}$	(6.45) (6.64)	(6.45) (6.64)
Density integration	$u_{S9(DET)}$	(6.78)	(6.78)
Temperature equation	$u_{T(DET)}$	(6.95)	(6.95)

C.2 Saturation correction uncertainty propagated to temperature

Figure C.2 is similar to **Figure C**.1, but for temperature uncertainty associated with saturation correction $u_{T(SAT)}$. The corresponding propagation equations are listed in **Table C**.2. The dotted curves show the ozone standard deviation resulting from the corresponding two Monte-Carlo experiments in which the lidar signals were simulated with saturation correction only (i.e., no detection noise, no background noise, no extinction terms).

The dotted curves show the ozone standard deviation resulting from the corresponding Monte-Carlo experiment in which the lidar signals were simulated with saturation correction only (i.e., no detection noise, no background noise, no extinction terms). The same population of 200 normally-distributed dead-time values was used in the inverse model for all channels. The results are presented as a function of normalized signal (counts per pulse per microsecond). Choosing normalized signal instead of altitude as the independent variable allows to bring the curves from all six channels together into a consistent behavior. The dotted curves (standard deviation) and dashed curves (calculate uncertainty) match perfectly, which confirms that the set of equations listed in **Table C**.2 is correct.

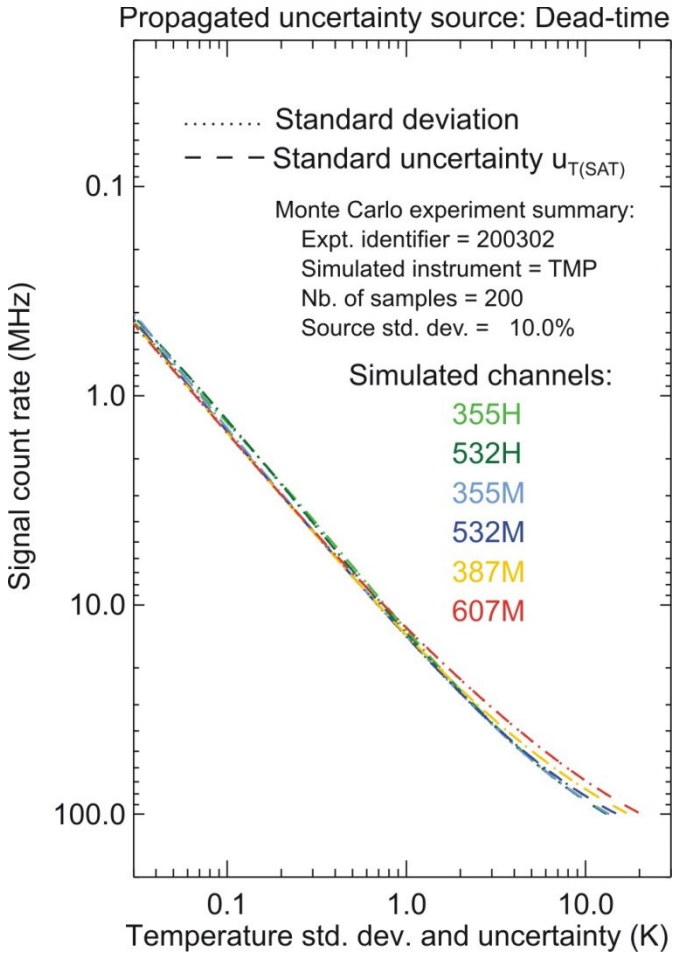


Figure C.2 Temperature standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from a Monte Carlo experiment designed to quantify temperature uncertainty owed to dead-time correction $u_{T(SAT)}$. (see text for details)

Table C.2 Expressions used for saturation correction uncertainty propagated to temperature

Sub-model Description	*Propagated uncertainty*	*Eq. Non-paralyz*	*Eq. Paralyz.*
Saturation correction	$u_{S1(SAT)}$	(4.15)	(H.8)
Background extraction	$u_{S2(SAT)}$	(4.38)	(4.38)
Merging	$u_{S4(SAT)}$	(4.53)	(4.53)
Smoothing (lin) Smoothing (log)	$u_{S5(SAT)}$	(4.67) (4.75)	(4.67) (4.75)
Range correction	$u_{S6(SAT)}$	(6.3)	(6.3)
Extinction correction	$u_{S7(SAT)}$	(6.11)	(6.11)
Layer-averaged signal and gravity	$u_{S8(SAT)}$	(6.46) (6.65)	(6.46) (6.65)
Density integration	$u_{S9(SAT)}$	(6.79)	(6.79)
Temperature equation	$u_{T(SAT)}$	(6.96)	(6.96)

C.3 Background extraction uncertainty propagated to temperature

Uncertainty associated with background correction and propagated to temperature $u_{T(BKG)}$ is plotted in **Figure C**.3 (dashed curves). The corresponding propagation equations are listed in **Table C**.3. For this Monte-Carlo experiment, the simulated lidar signals contained detection noise and background noise, but no saturation, and no extinction terms. In the inverse model, the background was corrected using the fitting function presented in **section 4.3** (linear function of altitude range). Though detection noise is not the target parameter of this Monte Carlo experiment, it had to be included in the signals in order to produce better fitting results and realistic uncertainty estimates of the fitting function coefficients. In the example shown here, the target parameter was the fitting function coefficient b_1. The magnitude of the coefficient uncertainty was taken directly from the value returned by the fitting routine, and then used as the standard deviation of a normally distributed population of 200 coefficients b_1 subsequently used to correct for background noise, and thus producing 200 temperature profiles, each of which produced using a different value of b_1. The results are presented as a function of normalized signal (counts per pulse per microsecond). Choosing normalized signal instead of altitude (left) and as a function of signal-to-noise ratio (right). The right plot shows that all six channels have a consistent behavior, i.e., uncertainty decreasing with height at a constant rate in log-log-scale. Not surprisingly the uncertainty and standard deviation increase with height, but then collapse to zero just below the tie-on altitude. This is due to the increased impact of the ancillary measurement when we approach the top. Nevertheless, the temperature standard uncertainty reported by the inverse model (long-dash curves) again matches well the calculated temperature standard deviation for all ranges and all altitudes.

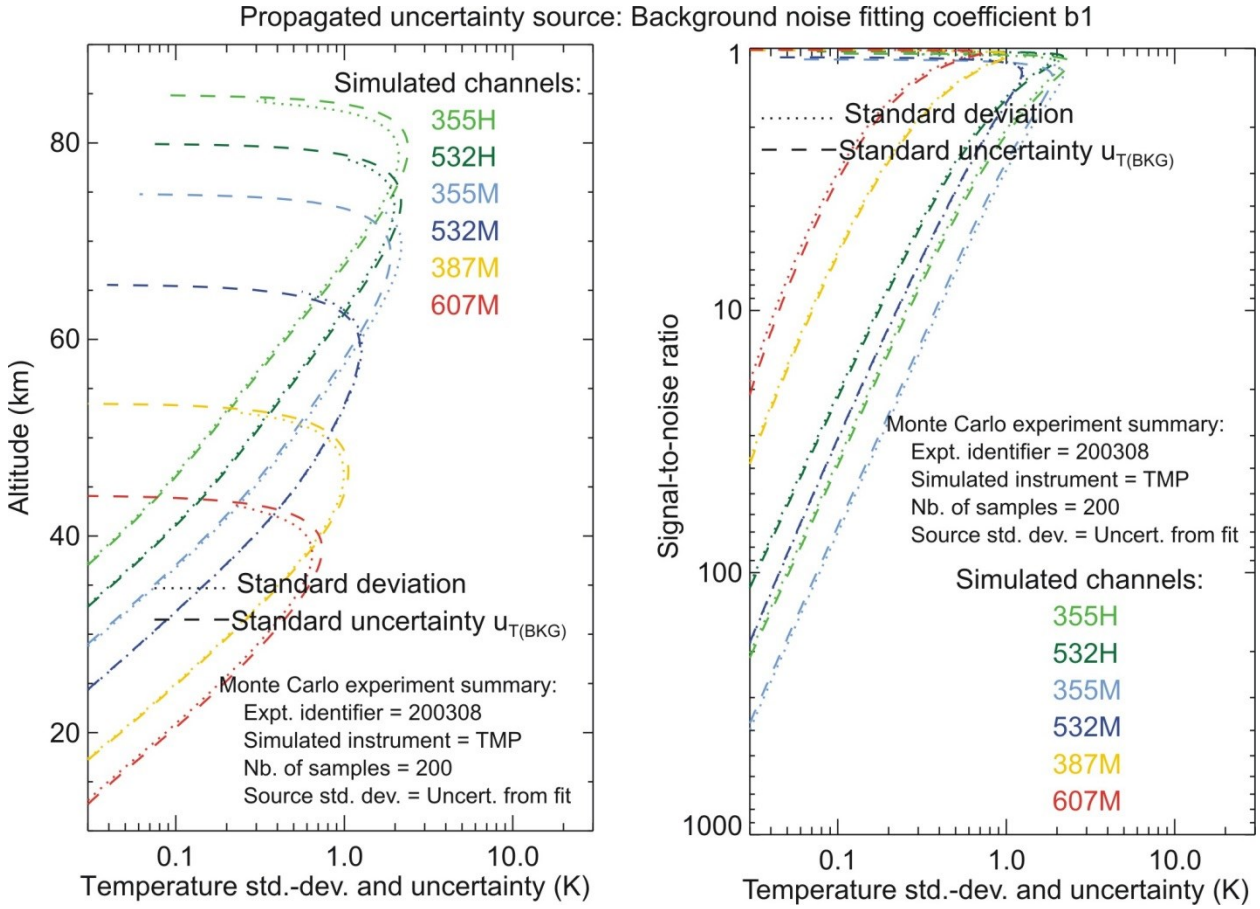


Figure C.3 Temperature standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from a Monte Carlo experiment designed to quantify temperature uncertainty owed to background correction $u_{T(BKG)}$ (see text for details)

Table C.3 Expressions used for background correction uncertainty propagated to temperature

Sub-model description	***Propagated uncertainty***	***Eq.***
Background extraction	$u_{S2(BKG)}$	(4.35)
Merging	$u_{S4(BKG)}$	(4.54)
Smoothing (lin) Smoothing (log)	$u_{S5(BKG)}$	(4.68) (4.76)
Range correction	$u_{S6(BKG)}$	(6.4)
Extinction correction	$u_{S7(BKG)}$	(6.12)
Layer-averaged signal and gravity	$u_{S8(BKG)}$	(6.47) (6.66)
Density integration	$u_{S9(BKG)}$	(6.80)
Temperature equation	$u_{T(BKG)}$	(6.97)

Similar experiments were performed with b_0 being the target parameter instead of b_1, and a perfect match between temperature standard deviation and calculated standard uncertainty was again observed (not shown). Additionally, similar results would be obtained with the temperature standard uncertainty associated with overlap correction $u_{T(OVER)}$ and signal merging $u_{T(MERGE)}$ (not shown).

C.4 Ozone absorption uncertainty propagated to temperature

Uncertainty associated with the ozone absorption cross-section and propagated to temperature $u_{T((\sigma O3)}$ is plotted in **Figure C**.4. The corresponding propagation equations are listed in **Table C**.4. For this Monte-Carlo experiment, the signals were simulated with no detection noise, no background noise, no saturation effects, no absorption terms by minor species except ozone, and no extinction terms. In the inverse model, a set of 200 temperature profiles were produced by varying the values of the ozone absorption cross-sections. Two configuration cases are shown. On the left-hand side, it is assumed that the cross-sections at all wavelengths are independent from each other (uncorrelated), which corresponds to the propagation equation reported in the third column (Rayleigh backscatter) and fourth column (Raman backscatter) of **Table C**.4. Using the values and uncertainty estimates provided by University of Reims spectroscopy group (DMB), two independent sets of 200 normally-distributed cross-section perturbation values were used for the emitted and received wavelengths of each channel, with a standard deviation around the mean values of 5% for wavelengths in the Chappuis band and 20% for wavelengths in the region of minimum ozone absorption. On the right-hand side, it is assumed that the cross-sections at all wavelengths are fully correlated, which corresponds to the propagation equation reported in the last column of **Table C**.4. Using again the values and uncertainty estimates provided by DMB, the same set of 200 normally-distributed cross-section perturbation values was used for the emitted and received wavelengths of each channel. Again, for both configurations, and both Rayleigh and Raman cases, the standard uncertainty reported by the inverse model (long-dash curves) matches again very well the calculated ozone standard deviation (dotted curves) for all ranges and all altitudes. Note that in the Rayleigh backscatter case, the results are identical for the correlated cross-sections and uncorrelated cross-sections configurations because a unique wavelength is used for emission and reception.

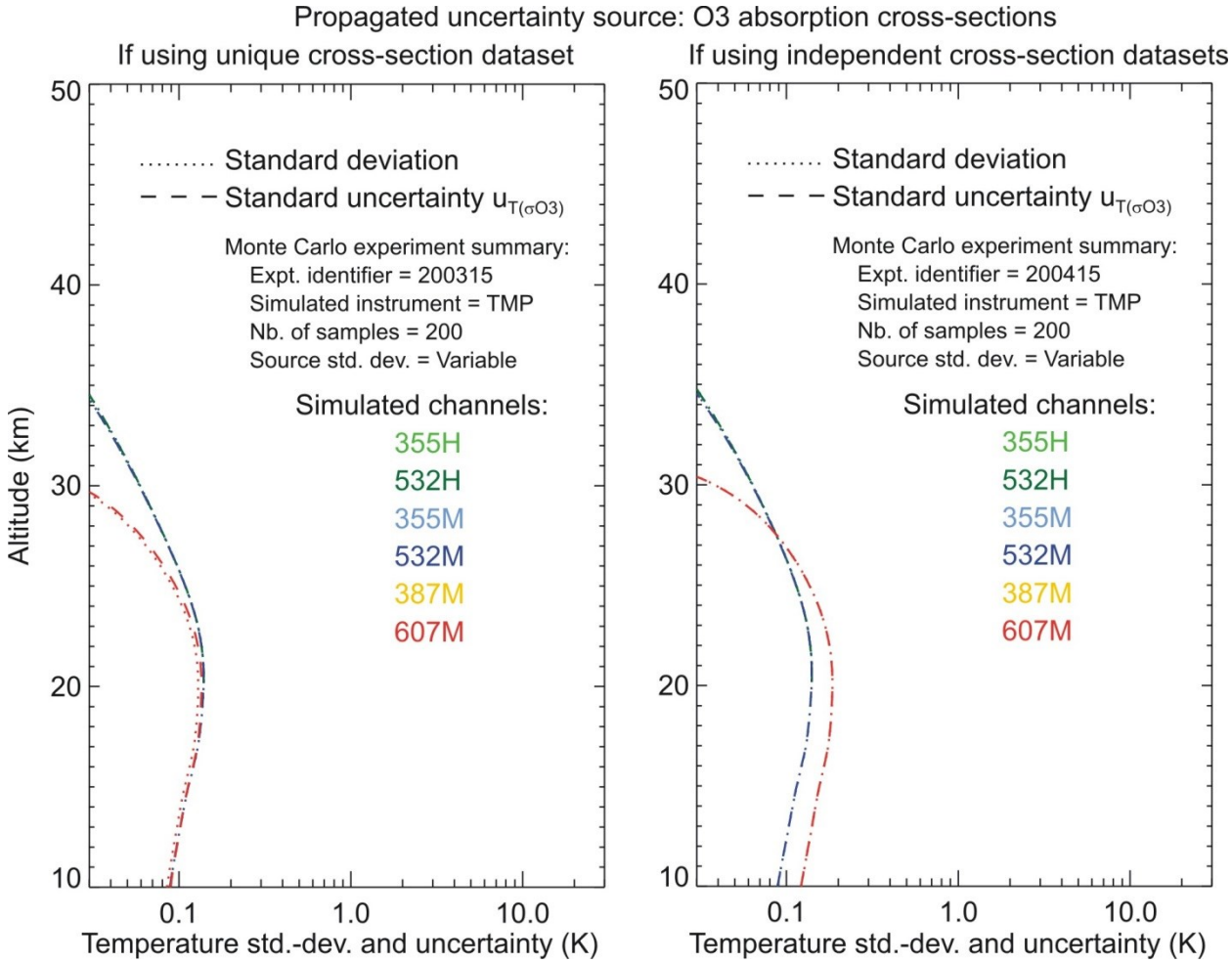


Figure C.4 Temperature standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from Monte Carlo experiments designed to quantify temperature uncertainty owed to ozone absorption cross-section $u_{T(\sigma O3)}$. Left-hand plot: Assuming that the cross-sections at all wavelengths are fully correlated; Right-hand plot: Assuming that cross-sections at all wavelengths are independent (see text for details)

Table C.4 Expression used for O_3 absorption cross-section uncertainty propagated to ozone

Sub-model description	***Propagated uncertainty***	***Eq. if using independent σ***	***Eq. if using correlated σ***
Extinction correction (Rayleigh) Extinction correction (Raman)	$u_{S7(\sigma O3)}$	(6.21) (6.20)	(6.23) (6.22)
Layer-averaged signal and gravity	$u_{S8(\sigma O3)}$	(6.52) (6.71)	(6.52) (6.71)
Density integration	$u_{S9(\sigma O3)}$	(6.85)	(6.85)
Temperature equation	$u_{T(\sigma O3)}$	(6.102)	(6.102)

Uncertainty associated with the ancillary O_3 number density and propagated to temperature $u_{T((O3)}$ is plotted in **Figure C**.5. The corresponding propagation equations are listed in **Table C**.5. For this Monte-Carlo experiment, the signals were simulated with no detection noise, no background noise, no saturation effects, and no extinction terms except ozone absorption. In the inverse model, 200 temperature profiles were produced by varying the values of the ancillary O_3 number density. A set of 200 normally-distributed O_3 number density values was produced, with a standard deviation around the mean equal to the ancillary O_3 number density uncertainty. This uncertainty can vary significantly depending on the dataset source. In the example shown, this value was set to 10%. The temperature standard uncertainty reported by the inverse model (long-dashed curves) matches again very well the calculated temperature standard deviation (dotted curves) for all ranges and all altitudes.

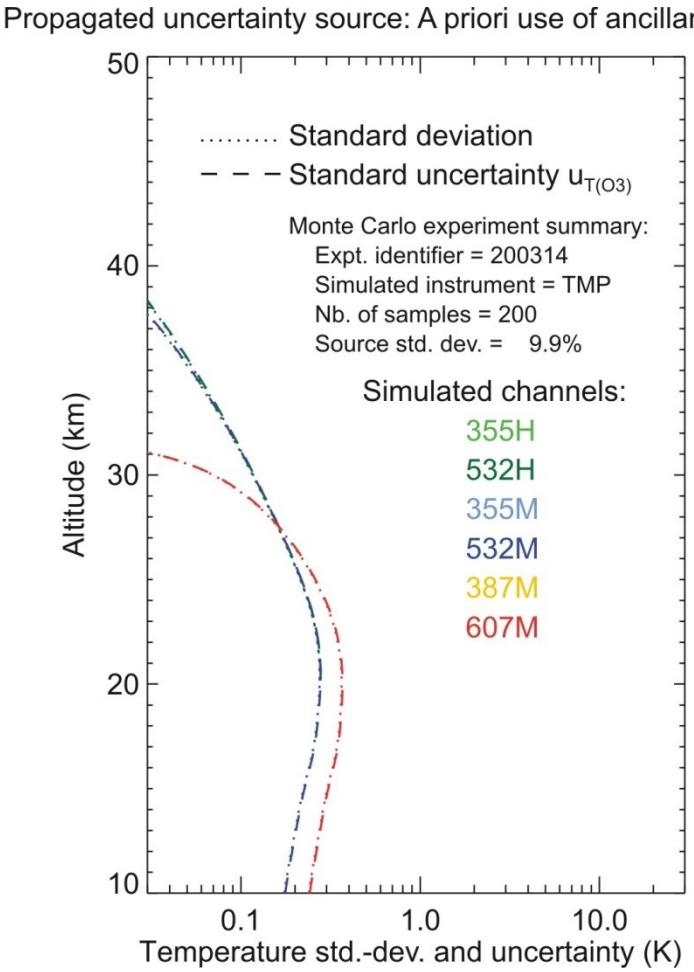


Figure C.5 Temperature standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from a Monte Carlo experiment designed to quantify temperature uncertainty owed to O3 number density $u_{T(O3)}$ (see text for details)

Table C.5 Expression used for ancillary O_3 number density or mixing ratio uncertainty propagated to ozone

Sub-model Description	*Propagated uncertainty*	*Eq.*
Extinction correction (Rayleigh channels)	$u_{S7(NO3)}$ $u_{S7(qO3)}$	(6.31) (6.39)
Extinction correction (Raman channels)	$u_{S7(NO3)}$ $u_{S7(qO3)}$	(6.29) (6.37)
Layer-averaged signal and gravity (using number density)	$u_{S8(NO3)}$	(6.53) (6.72)
Layer-averaged signal and gravity (using mixing ratio)	$u_{S8(qO3)}$	(6.54) (6.73)
Density integration (using number density) Density integration (using mixing ratio)	$u_{S9(NO3)}$ $u_{S9(qO3)}$	(6.86) (6.87)
Temperature equation (using number density) Temperature equation (using mixing ratio)	$u_{T(NO3)}$ $u_{T(qO3)}$	(6.103) (6.104)

C.5 NO_2 absorption uncertainty propagated to temperature

Uncertainty associated with the NO_2 absorption cross-section and propagated to temperature $u_{T((\sigma NO2)}$ is plotted in **Figure C**.6. The corresponding propagation equations are listed in **Table C**.6. For this Monte-Carlo experiment, the signals were simulated with no detection noise, no background noise, no saturation effects, no Rayleigh extinction, and no absorption terms by minor species except NO_2. In the inverse model, a set of 200 temperature profiles were produced by varying the values of the NO_2 absorption cross-sections. Here it is assumed that the cross-sections at all wavelengths come from the same dataset and are fully correlated, which corresponds to the propagation equation reported in the third column (Rayleigh backscatter) and fourth column (Raman backscatter) of **Table C**.6. Using the values and uncertainty estimates provided by University of Bremen spectroscopy group (Bogumil et al., 2003), two independent sets of 200 normally-distributed cross-section perturbation values were used for the emitted and received wavelengths of each channel, with a standard deviation around the mean values of 5% at all wavelengths. For both Rayleigh and Raman cases, the standard uncertainty reported by the inverse model (long-dash curves) matches again very well the calculated ozone standard deviation (dotted curves) for all ranges and all altitudes.

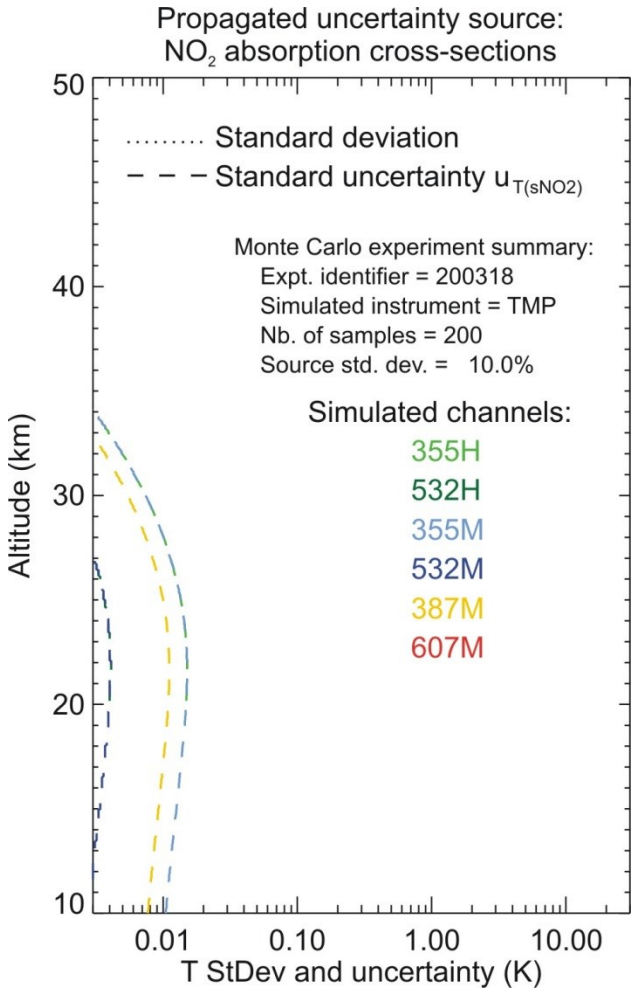


Figure C.6 Temperature standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from Monte Carlo experiments designed to quantify temperature uncertainty owed to NO_2 absorption cross-section $u_{T(\sigma NO2)}$, assuming that the cross-sections at all wavelengths are fully correlated (see text for details)

Table C.6 Expression used for NO_2 absorption cross-section uncertainty propagated to temperature

Sub-model description	***Propagated uncertainty***	***Eq. if using independent σ***	***Eq. if using correlated σ***
Extinction correction (Rayleigh channels) Extinction correction (Raman channels)	$u_{S7(\sigma NO2)}$	(6.25) (6.24)	(6.27) (6.26)
Layer-averaged signal and gravity	$u_{S8(\sigma NO2)}$	(6.55) (6.74)	(6.55) (6.74)
Density integration	$u_{S9(\sigma NO2)}$	(6.88)	(6.88)
Temperature equation	$u_{T(\sigma NO2)}$	(6.105)	(6.105)

Uncertainty associated with the ancillary NO_2 number density and propagated to temperature $u_{T((NNO2)}$ is plotted in **Figure C**.7. The corresponding propagation equations are listed in **Table**

C.7. For this Monte-Carlo experiment, the signals were simulated with no detection noise, no background noise, no saturation effects, and no extinction terms except NO_2 absorption. In the inverse model, 200 temperature profiles were produced by varying the values of the ancillary NO_2 number density. A set of 200 normally-distributed NO_2 number density values was produced, with a standard deviation around the mean equal to the ancillary NO_2 number density uncertainty. This uncertainty can vary significantly depending on the dataset source. In the example shown, this value was set to 10%. The temperature standard uncertainty reported by the inverse model (long-dash curves) matches again very well the calculated temperature standard deviation (dotted curves) for all ranges and all altitudes.

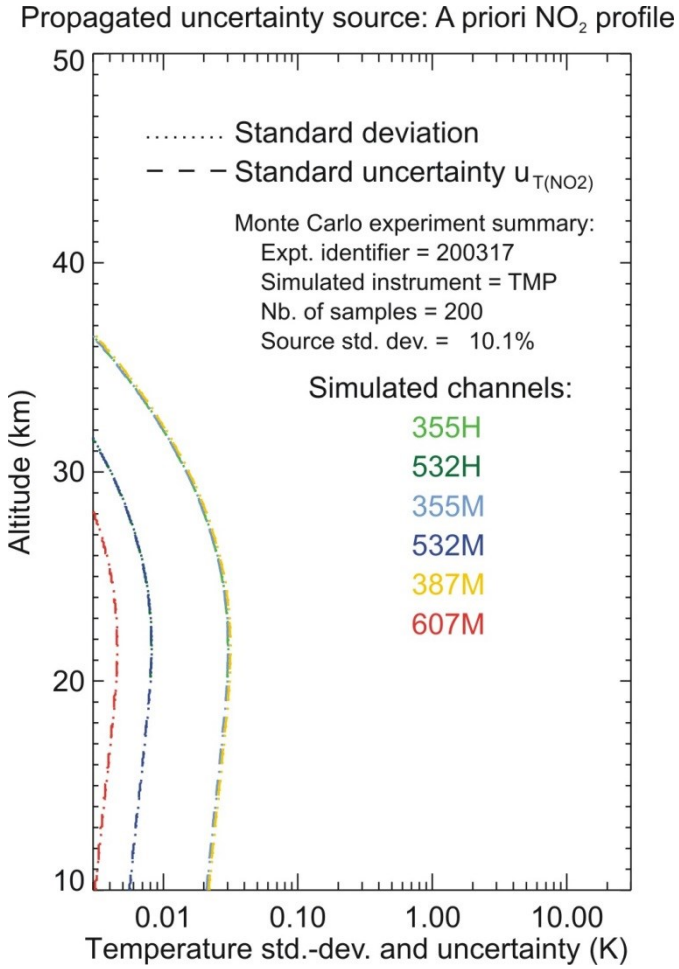


Figure C.7 Temperature standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from a Monte Carlo experiment designed to quantify temperature uncertainty owed to NO_2 number density $u_{T(NNO2)}$ (see text for details)

3765 **Table C.7 Expression used for ancillary NO_2 number density or mixing ratio uncertainty propagated to**
3766 **temperature**

Sub-model ***Description***	***Propagated uncertainty***	***Eq.***
Extinction correction (Rayleigh channels)	$u_{S7(\mathrm{NN}O2)}$ $u_{S7(\mathrm{q}NO2)}$	(6.32) (6.40)
Extinction correction (Raman channels)	$u_{S7(\mathrm{NNO2})}$ $u_{S7(\mathrm{q}NO2)}$	(6.30) (6.38)
Layer-averaged signal and gravity (using number density)	$u_{S8(\mathrm{NNO2})}$	(6.56) (6.75)
Layer-averaged signal and gravity (using mixing ratio)	$u_{S8(\mathrm{q}NO2)}$	(6.57) (6.76)
Density integration (using number density) Density integration (using mixing ratio)	$u_{S9(\mathrm{N}NO2)}$ $u_{S9(\mathrm{q}NO2)}$	(6.89) (6.90)
Temperature equation (using number density) Temperature equation (using mixing ratio)	$u_{T(\mathrm{N}NO2)}$ $u_{T(\mathrm{q}NO2)}$	(6.106) (6.107)

3767

3768 ## C.6 Molecular extinction uncertainty propagated to temperature

3769 Uncertainty associated with the Rayleigh cross-section and propagated to temperature $u_{T((\sigma M)}$ is
3770 plotted in **Figure C**.8. The corresponding propagation equations are listed in **Table C**.8. For this
3771 Monte-Carlo experiment, the signals were simulated with no detection noise, no background
3772 noise, no saturation effects, and no ozone absorption. In the inverse model, a set of 200
3773 temperature profiles were produced by varying the values of the Rayleigh cross-sections. It is
3774 assumed that the cross-sections at all wavelengths (i.e., 355 nm, 387 nm, 532 nm, and 607 nm)
3775 come from the same analytical formulae. Using the values and uncertainty estimates provided by
3776 Eberhard (2010), a single set of 200 normally-distributed cross-section values was used for all
3777 the channels, with a standard deviation around the mean values of 2% at all wavelengths and all
3778 altitudes. The resulting behavior of the temperature uncertainty and standard deviation is a
3779 constant slope with height. The temperature standard uncertainty reported by the inverse model
3780 (long-dash curves) matches again very well the calculated temperature standard deviation (dotted
3781 curves) for all ranges and all altitudes.

3782

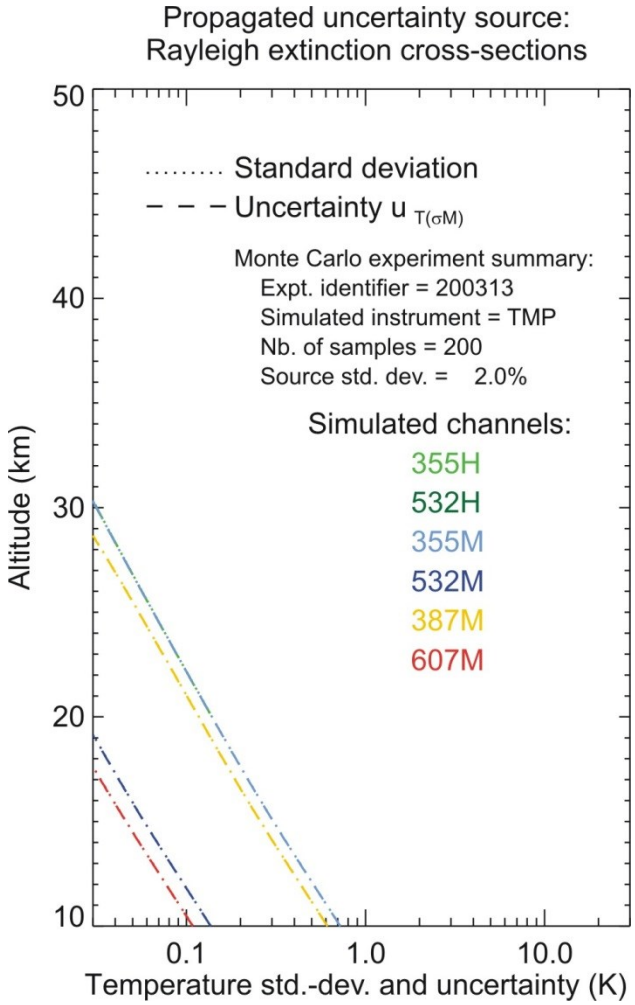


Figure C.8 Temperature standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from a Monte Carlo experiment designed to quantify temperature uncertainty owed to Rayleigh extinction cross-sections $u_{T(\sigma M)}$ (see text for details)

Table C.8 Expression used for Rayleigh cross-section uncertainty propagated to temperature

Sub-model description	***Propagated uncertainty***	***Eq. if using correlated σ***
Extinction correction (Rayleigh) Extinction correction (Raman)	$u_{S7(\sigma M)}$	(6.17) (6.16)
Layer-averaged signal and gravity	$u_{S8(\sigma M)}$	(6.50) (6.69)
Density integration	$u_{S9(\sigma M)}$	(6.83)
Temperature equation	$u_{T(\sigma M)}$	(6.100)

Uncertainty associated with the ancillary air number density and propagated to temperature $u_{T((Na)}$ is plotted in **Figure C**.9. The corresponding propagation equations are listed in **Table C**.9.

For this Monte-Carlo experiment, the signals were simulated with no detection noise, no background noise, no saturation effects, and no absorption terms. In the inverse model, 200 temperature profiles were produced by varying the values of the ancillary air number density. A set of 200 normally-distributed air number density values was produced, with a standard deviation around the mean equal to the ancillary air number density uncertainty. This uncertainty can vary significantly depending on the dataset source. If the air number density is computed using pressure and temperature measurements from radiosonde, one should expect a small uncertainty deduced from a typical pressure uncertainty of 0.1 hPa and temperature uncertainty of 0.5 K. If the air number density is computed using an analysis or reanalysis model such as NCEP or ECMWF, we should expect an uncertainty of up to 5%. If the air number density is computed using an empirical model such as MSISE-90 or CIRA, we should expect a larger uncertainty of up to 10%. In the example shown, estimates from radiosonde were used below 30 km, and estimates from empirical model were used above 30 km, which explains the different behavior above and below 30 km. The temperature standard uncertainty reported by the inverse model (long-dash curves) matches again very well the calculated temperature standard deviation (dotted curves) for all ranges and all altitudes.

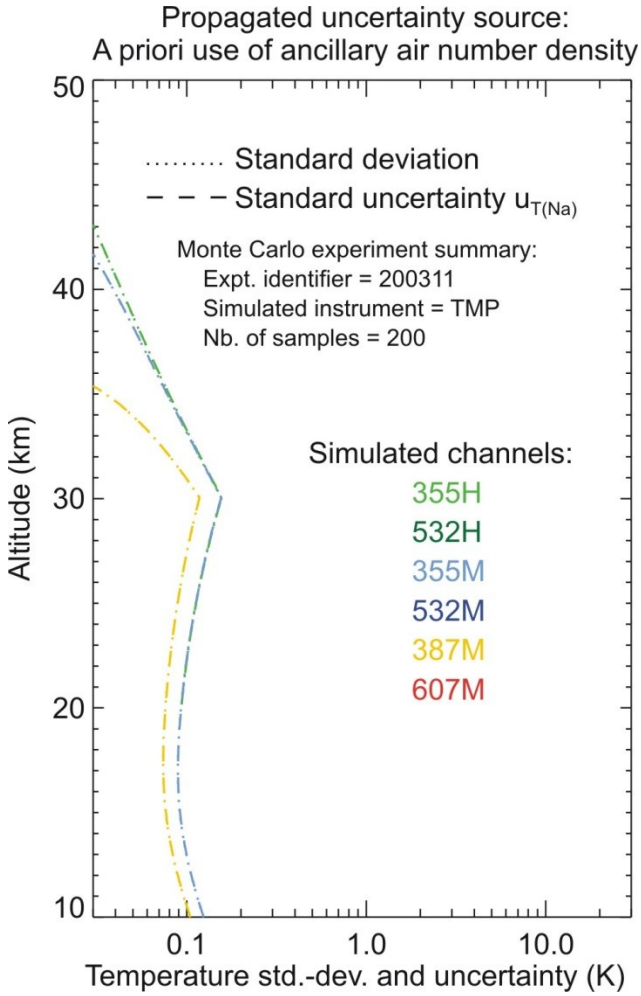


Figure C.9 Temperature standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from a Monte Carlo experiment designed to quantify temperature uncertainty owed to air number density $u_{T(Na)}$ (see text for details)

Table C.9 Expression used for ancillary air number density uncertainty propagated to temperature

Sub-model Description	***Propagated uncertainty***	***Eq.***
Extinction correction (Rayleigh channels)	$u_{S7(Na)}$	(6.19)
Extinction correction (Raman channels)	$u_{S7(Na)}$	(6.18)
Layer-averaged signal and gravity	$u_{S8(Na)}$	(6.51) (6.70)
Density integration	$u_{S9(NO3)}$	(6.84)
Temperature equation	$u_{T(NO3)}$	(6.101)

C.7 Acceleration of gravity uncertainty propagated to temperature

Uncertainty associated with the acceleration of gravity and propagated to temperature $u_{T((g)}$ is plotted in **Figure C**.10. The results are presented as a function of the distance from the top of the

profile (tie-on altitude) in order to show the consistent behavior of all six channels. This behavior is a direct consequence of the density downward integration and is identical for all the channels. The corresponding propagation equations are listed in **Table C**.10. For this Monte-Carlo experiment, the signals were simulated with no detection noise, no background noise, no saturation effects, and no extinction terms. In the inverse model, a set of 200 temperature profiles were produced by varying the values of the acceleration of gravity. Two configurations cases are shown. In one case (results shown on the left hand side), a set of 200 normally-distributed acceleration of gravity values was used for all the channels, with a standard deviation around the mean values of 0.1% corresponding to a very conservative uncertainty estimate from the WGS84 latitude- and altitude-dependent gravity model. In the other case (results shown on the right hand side), the 200 normally-distributed values of acceleration of gravity have a standard deviation around the mean of 3% which is equivalent to using a constant value of acceleration by gravity at all latitudes and all altitudes (for example 9.8065 ms^{-2}). The temperature standard uncertainty reported by the inverse model (long-dash curves) matches again very well the calculated temperature standard deviation (dotted curves) for all ranges and all altitudes.

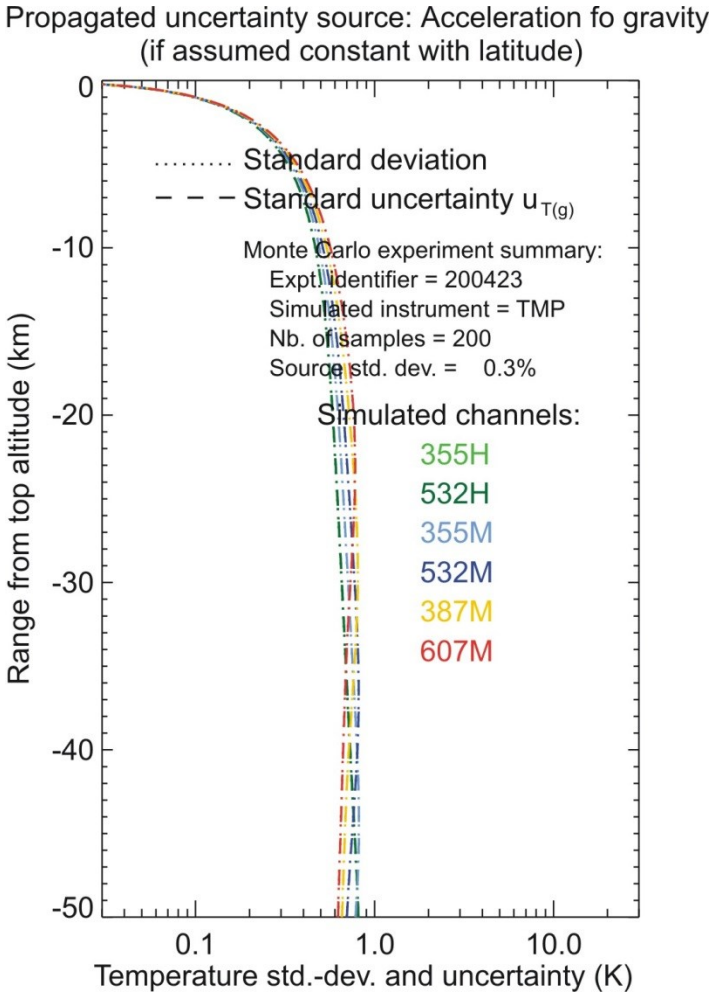


Figure C.10 Temperature standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from a Monte Carlo experiment designed to quantify temperature uncertainty owed to the acceleration of gravity $u_{T(g)}$. Left hand plots: using the altitude and latitude-dependent values of the WGS84 model; Right-hand plots: using a constant value for all latitudes and all altitudes (see text for details)

Table C.10 Expression used for acceleration of gravity uncertainty propagated to temperature

Sub-model Description	*Propagated uncertainty*	*Eq.*
Layer-averaged signal and gravity	$u_{S8(g)}$	(6.63)
Density integration	$u_{S9(g)}$	(6.91)
Temperature equation	$u_{T(g)}$	(6.108)

C.8 Molecular mass of air uncertainty propagated to temperature

Uncertainty associated with the molecular mass of air and propagated to temperature $u_{T(Ma)}$ is plotted in **Figure C**.11. Once again, and for the reasons explained in the previous section, the results are presented as a function of the distance from the top of the profile (tie-on altitude). The corresponding propagation equation is listed in **Table C**.11. For this Monte-Carlo experiment, the signals were simulated with no detection noise, no background noise, no saturation effects, and no extinction terms. In the inverse model, a set of 200 temperature profiles were produced by varying the values of the molecular mass of air. In this case, a set of 200 normally-distributed values was used for all the channels, with a standard deviation around the mean values of 0.02% corresponding to the relative difference between a rounded value of 0.02896 $Kg.mol^{-1}$ and the CIPM-2007 value of 0.0289654 $Kg.mol^{-1}$ (Picard et al., 2008). The temperature standard uncertainty reported by the inverse model (long-dash curves) matches again very well the calculated temperature standard deviation (dotted curves) for all ranges and all altitudes.

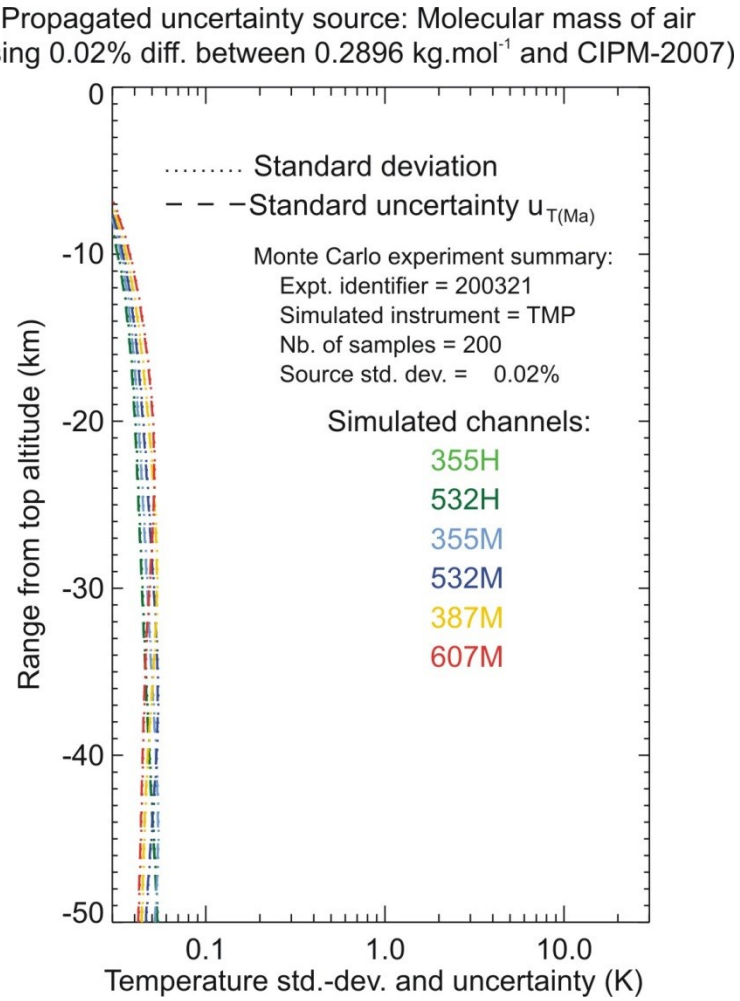


Figure C.11 Temperature standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from a Monte Carlo experiment designed to quantify temperature uncertainty owed to the rounding of the molecular mass of dry air $u_{T(Ma)}$ (see text for details)

Table C.11 Expression used for the molecular mass of dry air uncertainty propagated to temperature

Sub-model Description	***Propagated uncertainty***	***Eq.***
Temperature equation	$u_{T(Ma)}$	(6.94)

C.9 Temperature tie-on uncertainty propagated to temperature

Temperature uncertainty associated with the tie-on procedure at the top of the temperature profile $u_{T((TTOP)}$ is plotted in **Figure C**.12. The corresponding propagation equation is listed in **Table C**.12. For this Monte-Carlo experiment, the signals were simulated with no detection noise, no background noise, no saturation effects, and no absorption terms. In the inverse model, 200 temperature profiles were produced by varying the values of the ancillary temperature used to initialize the top of the profile. A set of 200 normally-distributed temperature values was produced, with a standard deviation around the mean equal to the ancillary temperature uncertainty. This uncertainty can vary significantly depending on the dataset source. It ranges

between 10 K for datasets such as Aura-MLS or SABER, to 20 K if an empirical model such as MSISE-90 or CIRA is used, as is the case in our example. The results are presented in function of the distance from the tie-on altitude. As expected, the uncertainty and standard deviation rate of increase with altitude is identical for all channels. The absolute value depends only on the value of the tie-on temperature uncertainty. In the present case, this value is 20 K for the 355H, 355M and 532H (using the empirical model MSIS in the mesosphere), and 10 K for the other channels (tie-on in the upper stratosphere). The temperature standard uncertainty reported by the inverse model (long-dash curves) matches again very well the calculated temperature standard deviation (dotted curves) for all ranges and all altitudes.

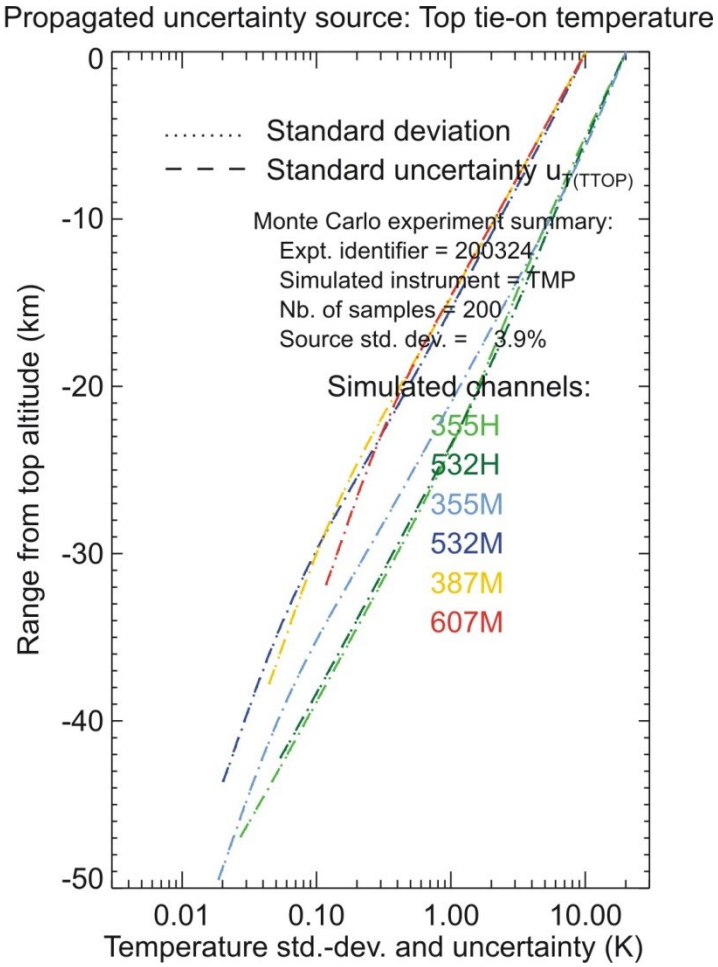


Figure C.12 Temperature standard deviation (dotted curves) and standard uncertainty (long-dash curves) obtained from a Monte Carlo experiment designed to quantify temperature uncertainty owed to tie-on at the top of the profile $u_{T(TTOP)}$ (see text for details)

Table C.12 Expression used for ancillary top tie-on temperature uncertainty propagated to temperature

Sub-model Description	***Propagated uncertainty***	***Eq.***
Temperature equation	$u_{T(TTOP)}$	(6.93)

C.10 Joint propagation of all temperature uncertainty components

Now that all individual components have been properly quantified using their propagation expressions listed in **Table C**.1**-Table C**.12, it is time to verify that each individual component is independent from the others, so that we can derive the temperature combined standard uncertainty by computing the quadratic sum of the individual components. To do this, we generalize the Monte Carlo experiments presented so far for an individual component to an experiment in which all sources of uncertainty are applied and propagated simultaneously. For each input quantity, we compute a population of 200 normally-distributed values of this quantity with a standard deviation equal to the quantity's uncertainty. Each produced population is "orthogonal" to the others, i.e., every set of normally-distributed values is independent from the others (correlation coefficient=0). The temperature combined uncertainty u_T as computed using **Eq. (6.109)** is plotted in **Figure C**.13 for our simulated temperature system "TMP". The temperature combined standard uncertainty reported by the inverse model (long-dash curves) matches again very well the calculated temperature standard deviation (dotted curves) for all ranges and all altitudes, demonstrating that the individual components are uncorrelated and can be propagated in parallel throughout the temperature data processing chain.

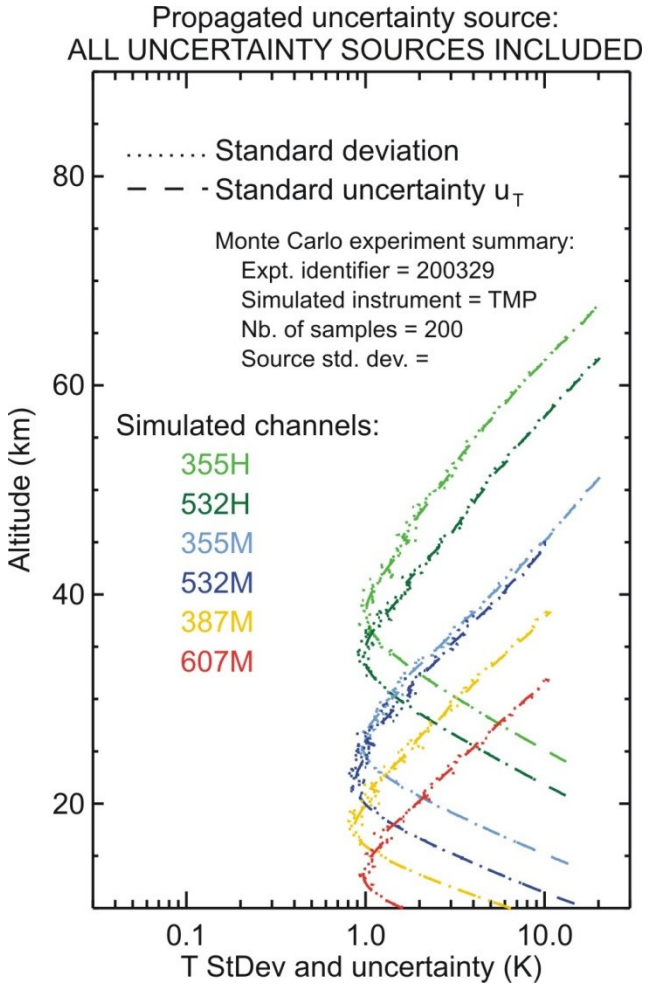


Figure C.13 Temperature standard deviation (dotted curves) and combined standard uncertainty (long-dash curves) obtained from a Monte Carlo experiment designed to quantify temperature combined uncertainty u_T (see text for details)

D Published approximations of the expression of molecular scattering

Our definition of "molecular extinction" throughout this report includes scattering and absorption by the air molecules excluding absorption by active or interfering minor species, as defined from **Eqs. (1.1)-(1.6)**. In this appendix we will use the term "Rayleigh scattering" rather than "extinction".

D.1 Theoretical derivation of the air Rayleigh scattering coefficient

Scattering here typifies attenuation caused by all elastic and inelastic linear scattering processes, which typically produces the non-shifted Landau-Placzek line, the weakly-shifted Rayleigh-Brillouin lines, weakly-shifted pure rotational Raman lines, and strongly-shifted rotational-vibrational Raman lines. The Raman lines are sometimes ignored, hence the historical use of an ambiguous term "Rayleigh extinction" and erroneous underestimation of the total scattering cross-section (Young, 1980; 1981). Because air is a mixture of gases, we cannot define the cross-section in the same straightforward manner as we did for absorption by individual species. Theoretically, it should be the volumetric Rayleigh scattering coefficient α_{Ray} (unit m^{-1}) as a whole that we would need to consider rather than the product of the air cross-section σ_{Ray} (unit m^2) by the air number density N_a (unit m^{-3}). However, we can derive a relatively self-consistent expression of the air mean Rayleigh cross-section based on relatively robust assumptions on the composition and properties of the air mixture.

Light scattering by molecules originates from the creation of an electric dipole induced by the electromagnetic field associated with the incident light. A number of assumptions on the properties of the molecules are made before molecular scattering can be treated analytically: 1) the incident wavelengths are much longer than the size of the scattering molecules, 2) the molecules are not intrinsically dipolar, are randomly distributed, and are randomly oriented within the scattering volume considered, and 3) the optical properties of the molecules are not altered by the presence of neighboring molecules. Under these conditions, the Rayleigh scattering coefficient (Strutt, 1899), where the air is represented by a mixture of gases G_i (i=1,N_G) with volume mixing ratios ν_i, can be written as a function of the each gas' refractive index n_i, their polarizability a_i, their anisotropy factor γ_i, the macroscopic air refractive index n_a, permittivity ε_0, the air total number density N_a, and the wavelength of the incident light λ (Eberhard, 2010):

$$\alpha_{Ray}(\lambda) = \frac{8\pi^3 N_a}{3\varepsilon_0^2 \lambda^4} L^2(\lambda) \sum_i \nu_i a_i(\lambda) F_i(\lambda) \qquad \textbf{(D.1)}$$

In this equation, the term F_i is the King factor:

$$F_i(\lambda) = 1 + \frac{2\gamma_i^2(\lambda)}{9a_i^2(\lambda)} = \frac{6+3\delta_i}{6-7\delta_i} \qquad \textbf{(D.2)}$$

with:

$$\delta_i = 1 + \frac{6\gamma_i^2(\lambda)}{45a_i^2(\lambda) + 7\gamma_i^2(\lambda)} \quad \textbf{(D.3)}$$

The King factor is introduced for a specific gas G_i, to take into account the molecules' anisotropic effects of scattering (King, 1923). N_a and λ should be in compatible units (e.g., m^{-3} and m, or cm^{-3} and cm). The term:

$$L(\lambda) = \frac{n_a^2(\lambda) + 2}{3} \quad \textbf{(D.4)}$$

is the Lorentz factor, equivalent to the proportionality factor between the energy of the electric field induced within the molecules and that of the incident electric field. This term is absent when dealing with pure gases. The Lorentz-Lorenz equation provides, for each "pure" gas G_i, a relationship between the gas' microscopic property a_i (which is not easy to measure) and its macroscopic refractive index n_i at density N_i:

$$a_i(\lambda) = 3\varepsilon_0^2 \frac{1}{N_i} \frac{n_i^2(\lambda) - 1}{n_i^2(\lambda) + 2} \quad \textbf{(D.5)}$$

It also yields the following invariance relationship:

$$\frac{1}{N_i} \frac{n_i^2 - 1}{n_i^2 + 2} = cst = \frac{1}{N_{iS}} \frac{n_{iS}^2 - 1}{n_{iS}^2 + 2} \quad \textbf{(D.6)}$$

The subscript "*S*" denotes standard temperature and pressure (STP) conditions. The above equality allows us to conveniently choose air at STP conditions (e.g., 15°C, 1013.25 hPa referred to as STP15 thereafter, or 0°C, 1013.25 hPa referred to as STP0 thereafter) to determine (measure) accurate values of the refractive index, and then calculate the Rayleigh scattering coefficient using the following re-arranged **Eq.** (D.1**)**:

$$\alpha_{Ray}(\lambda) = \frac{24\pi^3 N_a}{\lambda^4} L^2(\lambda) \sum_i \nu_i \frac{1}{N_{iS}^2} \left(\frac{n_{iS}^2(\lambda) - 1}{n_{iS}^2(\lambda) + 2} \right)^2 F_i(\lambda) \quad \textbf{(D.7)}$$

Eq. (D.7) represents the most exact analytical expression of the Rayleigh volume scattering coefficient of a mixture of gases (Eberhard, 2010). Many published works however consider the parameters of this equation for the whole air mixture instead of considering each species separately. In addition, they often ignore the dispersion of the King factor. Differences further rise when different gas mixtures are considered. In earlier works, authors used a 3-compound standard air mixture (N_2, O_2, and Ar). Most authors later used a 4-compound standard dry air mixture taking into account typically 300-350 ppmv of CO_2 (e.g., Penndorf, 1957), and a 5-compound moist air mixture for cases of elevated water vapor mixing ratios (e.g., Eberhard, 2010). The approximations made by various authors led to the publication of many versions of **Eq. (D.7)**. The key parameters of this equation are compiled for a number of published works in

Table D.1. Eberhard (2010) provides an excellent review of the various approximations made in the literature, and their impact on the actual values of the Rayleigh scattering coefficient.

Table D.1: Key parameters used for published analytical formulation of extinction coefficient

Dataset	STP N_s	$n_s(\lambda)$	δ_n	F	Empirical $\lambda^{-4+x(\lambda)}$
Penndorf, 1957	STP15	Edlén, 1953	0.0350	1.061	/
Hoyt, 1976	unspecified	Edlén, 1953	0.0139	1.024	/
Fröhlich and Shaw, 1980	STP15	Peck, 1972	N_2: 2 references O_2: 2 references Ar: 2 reference	1.016	/
Bates, 1984	STP0	$N_2(\lambda)$: 2 references $O_2(\lambda)$: 3 references Ar(λ): 1 reference $CO_2(\lambda)$: 1 reference	/	$N_2(\lambda)$: 3 references $O_2(\lambda)$: 6 references Ar(λ): 1 reference $CO_2(\lambda)$: 1 reference	/
Nicolet, 1984	/	/	/	/	Applies to Bates, 1984 data
Bucholtz, 1995	STP15	Peck, 1972	/	Bates, 1984	/
Eberhard, 2010	STP0	Bates, 1984	/	Bates, 1984	/

D.2 The air "Mean" Rayleigh scattering cross-section

Once we ensure that the air number density N_a is consistent with the assumptions made on the composition of the air (i.e., proper derivation of the King factors F_i and refractive indices n_{is}), we can derive a convenient and self-consistent analytical expression for the air "mean" Rayleigh extinction cross-section:

$$\sigma_{Ray}(\lambda) = \frac{24\pi^3}{\lambda^4} L^2(\lambda) \sum_i \nu_i \frac{1}{N_{iS}^2} \left(\frac{n_{iS}^2(\lambda) - 1}{n_{iS}^2(\lambda) + 2} \right)^2 F_i(\lambda) \qquad \textbf{(D.8)}$$

The above expression is identical to equation (56) of Eberhard (2010). It differs from many other expressions found in literature due to the presence of the Lorentz factor, and to the joint summation of the polarizability and King factor. Referring to other authors, the closest expression to **Eq. (D.8)** would be that of Bates (1984):

$$\sigma_{Ray}(\lambda) = \frac{32\pi^3}{3\lambda^4} \frac{1}{N_S^2} \sum_i \nu_i \left(n_{iS}^2(\lambda) - 1 \right)^2 F_i(\lambda) \qquad \text{with } i = 1,4 \ (N_2, O_2, Ar, CO_2) \qquad \textbf{(D.9)}$$

Bates's results were fitted by Nicolet (1984) using:

$$\sigma_{Ray}(\lambda) = \frac{A}{\lambda^{4+B+C\lambda+D/\lambda}} \quad \textbf{(D.10)}$$

with λ in μm, and A=4.02.10^{-28}, B=-0.3228, C=0.389 and D=0.09426

Other expressions used are, for example, those of Penndorf (1957) and Hoyt (1976):

$$\sigma_{Ray}(\lambda) = \frac{8\pi^3}{3\lambda^4}\frac{1}{N_S^2}\left(n_S^2(\lambda)-1\right)^2 F_{air} \quad \textbf{(D.11)}$$

with F_{air} = constant (no dispersion)

Fröhlich and Shaw (1980):

$$\sigma_{Ray}(\lambda) = \frac{24\pi^3}{\lambda^4 N_a^2}\left(\frac{n_a^2(\lambda)-1}{n_a^2(\lambda)+1}\right)^2 \sum_i \nu_i F_i(\lambda) \quad \textbf{(D.12)}$$

with i = 1,5 (N_2, O_2, Ar, CO_2 ,H_2O), and Bucholtz (1995):

$$\sigma_{Ray}(\lambda) = \frac{24\pi^3}{\lambda^4 N_s^2}\left(\frac{n_S^2(\lambda)-1}{n_S^2(\lambda)+2}\right)^2 F'_{air}(\lambda) \quad \textbf{(D.13)}$$

with $F'_{air}(\lambda)$ = Effective air King factor from Bates (1984).

The air refractive index and King Factor used in the above formulations are plotted in **Figure D.1.** Frequently-used formulations of standard air refractive index (e.g., Edlén, 1953; Edlén, 1966; Peck and Reeder, 1972; Ciddor, 1996), and temperature and pressure-dependent air refractive index (Edlén, 1966) yield conservative absolute uncertainty estimates of 10^{-7}, which leads to 0.07% and 0.00001% relative uncertainty from the polarizability and Lorentz factor contributions respectively.

Figure D.2 illustrates the differences found between the Rayleigh scattering cross-sections formulation of Eberhard (2010) and that of authors listed above. The cross-section differences are plotted in the left panel (applies to lidar temperature retrieval). The impact on the cross-section differential (for the ozone DIAL retrieval) is shown in the right panel. The dominant source of uncertainty is the King factor as was already revealed by **Figure D.2** (right). The King factors of Hoyt (1976) and Fröhlich and Shaw (1980) are too small (-3% to -5%) because they do not include the Raman contributions to total scattering. Not taking into account the mixing ratios weights of the individual gases when computing the air King Factor or polarizability leads to errors in the cross-section of roughly +/-0.2% (Eberhard, 2010). Ignoring cases of elevated water vapor mixing ratio shifts positively by about the same amount (0.3%) any bias associated with dry air mixture approximations (Eberhard, 2010).

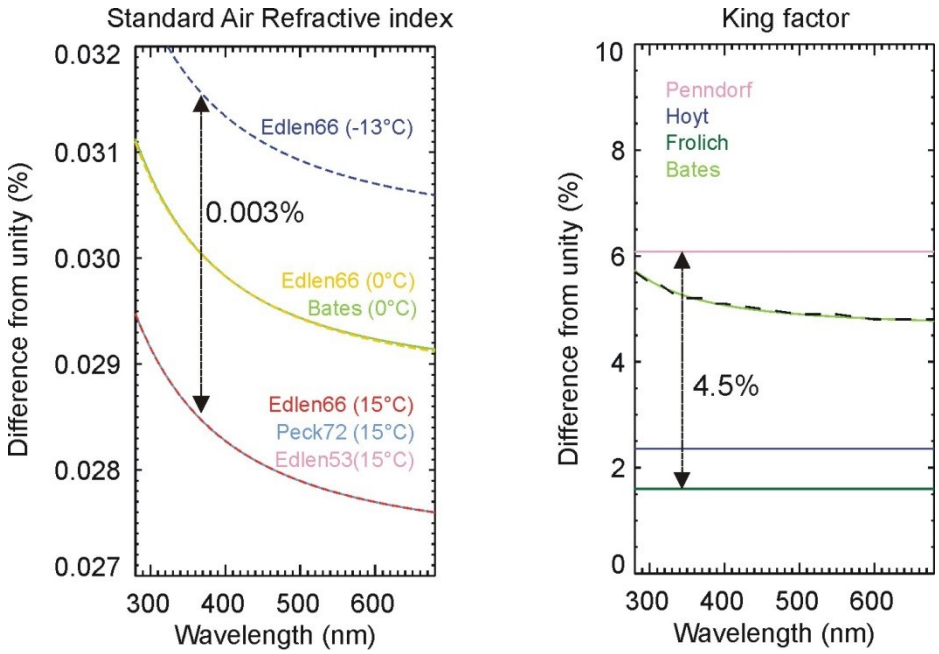


Figure D.1 Relative spread (%) of the various formulations of air refractive index (left) and King factor (right) found in literature

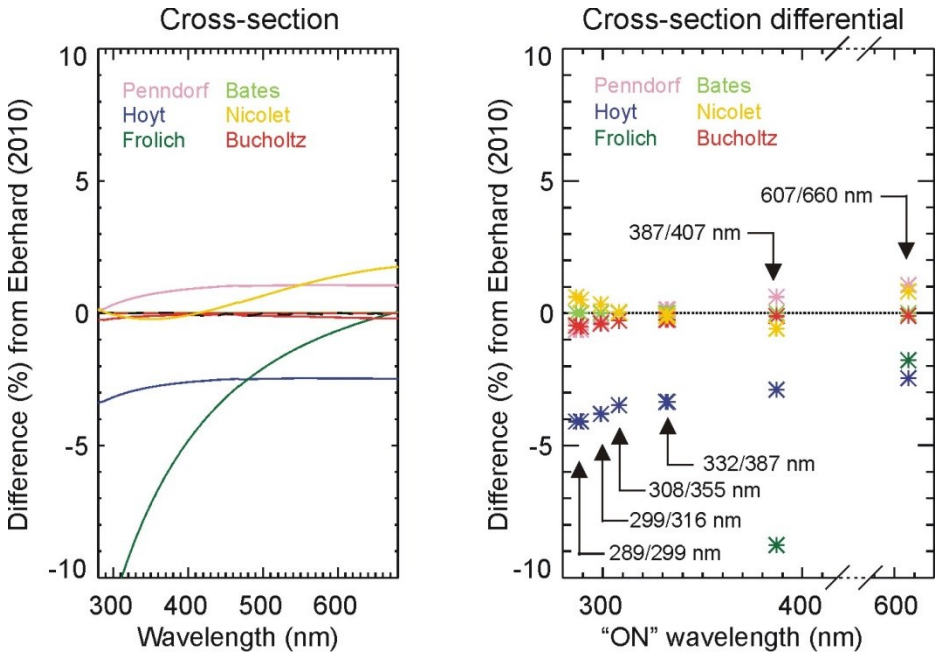


Figure D.2 Relative differences (%) of the Rayleigh scattering cross-section (left) and cross-section differential (right) computed for the various equations found in literature and with respect to Eberhard (2010)

E Assessment of eight contemporary ozone cross-section datasets

The cross-section differential term $\Delta\sigma_{O3}$ in **Eq. (1.4)** contributes non-negligibly to the ozone DIAL uncertainty budget. For this reason, it is important to quantify precisely the uncertainty associated with the cross-sections used in the retrieval, and their impact on the cross-section differential. Because of its complexity, the complete, line-line theoretical absorption spectrum of the ozone molecule is still not available today. The cross-sections used by the atmospheric research community originate in laboratory measurements using different experimental apparatus. The main issues for obtaining highly accurate absorption cross-sections of ozone in the UV are the determination of the vapor pressure of ozone, the photo-decomposition of ozone during the measurements and the wavelength calibration. Hanson and Mauersberger (1985) used freezing of undesired fragments plus mass-spectrometric control to obtain an accurate value for the vapor pressure of ozone. A modified set-up was then used to determine the absorption cross-section at 253.7 nm (air wavelength of the Hg line) with an uncertainty of just 0.7 % (Mauersberger et al., 1986). A simpler approach, based on a precision measurement of the total pressure of O_2 and O_3, taking into account the decomposition of ozone during the cross-section measurement, was used by Daumont et al. (1992), Malicet et al. (1995) and Brion et al. (1998), allowing the determination of cross-section values over a wider spectrum. The deviation between the values of Mauersberger et al. (1986) and Malicet et al. (1995) was found to around 0.6 %.

Today there are several publicly available ozone absorption cross-section datasets obtained from laboratory measurements (Serdyuchenko et al., 2014) and covering the range of wavelengths of interest to ozone DIAL, temperature and water vapor Raman lidars (250-700 nm). Even for the most recent measurements, relative differences of up to several percent have been found. The reasons for these discrepancies are as follows:

- Measurements of absolute cross-sections are not always available, and some datasets are indeed provided after normalization using values from previously published works
- The spectral resolution at which measurements are made or at which datasets are provided varies typically from 0.01 nm to 5 nm, therefore affecting the cross-section values eventually used, fitted, interpolated, or convolved with a specific spectral window
- In the 300-410 nm region (Huggins band and region of minimum absorption), the ozone absorption cross-sections have a pronounced temperature dependence. This temperature dependence is not fully described by the ozone molecule's quantum assignments (Qu et al., 2004), and therefore requires arbitrary assumptions for its empirical quantification.
- Some of the publicly available datasets can be found at various on-line or physical locations, which increases the chances for a loss of traceability and eventually leading to "duplicate" versions that are not strictly identical (see examples below)

In this report, eight datasets covering a wide range of temperatures and wavelengths are compared and reviewed. These datasets are available online on the UV-Visible Atlas of Gaseous Molecules website of the Max Planck Institute (Mainz, Germany) http://satellite.mpic.de/spectral_atlas/cross_sections/Ozone/O3.spc (thereafter referred to as "MPI-Mainz" or "Mainz" for brevity), and on the WMO Ad-hoc Working Group on Absorption Cross-sections of Ozone (ACSO) website http://igaco-o3.fmi.fi/ACSO/index.html. Three of the eight datasets considered comprise multiple data versions.

The reviewed datasets include, from most recent to oldest:

1) Serdyuchenko et al. (2014), obtained from ACSO
2) Chehade et al., (2013) and Bogumil et al. (2003)
    SCIAMACHY FM version 4 (2013), obtained from Univ. Bremen
    SCIAMACHY FM version 3 (2003), obtained from ACSO
    SCIAMACHY FM, unspecified version (v1 or v2), obtained from MPI-Mainz
3) Voigt et al. (2001), obtained from ACSO
4) Brion et al. (1998), Daumont et al. (1992), Malicet et al. (1995), referred thereafter to as "DMB"
    Partially smoothed data, obtained from ACSO
    Raw data, obtained directly from personal communication with the DMB group
5) Burrows et al. (1999) (GOME FM), obtained from ACSO
6) Burkholder et al. (1994), obtained from MPI-Mainz
7) Molina and Molina (1986), obtained from MPI-Mainz, referred thereafter to as "Molina"
8) Bass and Paur (1984), referred thereafter to as "BP"
    Data including updated quadratic coefficients, obtained from ACSO
    Data from personal communication with the BP group, obtained from MPI-Mainz

The measurements presented in references (1), (4), (5), (7) are absolute measurements obtained from the experimental measurement of total pressure (Gorshelev et al., 2014; Brion et al., 1998; Molina and Molina 1986) or using chemical titration (Burrows et al., 1997), while the measurements presented in references (2), (3), (6), (8), were normalized using values from Bass and Paur (1984), Burrows et al. (1999), Anderson and Mauersberger (1993), and Hearn (1961), respectively.

When duplicate datasets exist, the first dataset listed among those of the same source (i.e., datasets "a)") will be used for the comparisons shown thereafter because they are the most likely datasets currently used or expected to be used by the broader atmospheric remote sensing community. For reference, a short paragraph and two f**igure**s will be included at the beginning of this appendix in order to highlight differences observed between the duplicate datasets.

Each dataset has its own temperature coverage, spectral coverage, and spectral resolution. Because of its extensive wavelength and temperature coverage, and high spectral resolution, the dataset from Serdyuchenko et al. (2014) will often serve as reference. After reviewing the main features of the cross-section spectra and their differences as a function of wavelength and temperature, we will show the impact of these differences on the cross-section differentials used with the most common Rayleigh and Raman ozone DIAL pairs of wavelengths (289/299 nm, 299/316 nm, 308/355 nm, 308+331/355+387 nm, 308+332/355+387 nm) and Raman water vapor lidar pairs (387/407.5 nm, and 607/660 nm).

E.1 Sample Spectrum

Figure E.1 shows a typical measured ozone absorption cross-section spectrum (Serdyuchenko et al., 2013). The left panel shows the spectrum over the 200-750 nm region of interest to ozone, temperature, and water vapor lidars, and the right panel shows this spectrum zoomed in the Huggins band (300-350 nm).

The main features are the large absorption in the Hartley band (200-300 nm), the strong temperature dependence in the Huggins band (300-350 nm), a strong temperature dependence and increasing measurement relative uncertainty in the region of minimum absorption (350-410 nm), and a weak temperature dependence and moderate absorption in the Chapuis band (450-750 nm). Though the Hartley and Huggins bands have been attributed long time ago to the electronic and vibrational energy transitions of the ozone molecule respectively, there is still no exact theoretical model to accurately reproduce every line-by-line feature in the Huggins band and in the region of minimum absorption. As a result, only empirical models currently account for the temperature dependence of the cross-sections in these regions. This caveat together with low-absorption values, lead to large relative uncertainties and large relative differences between the compared datasets in the 350-410 nm region.

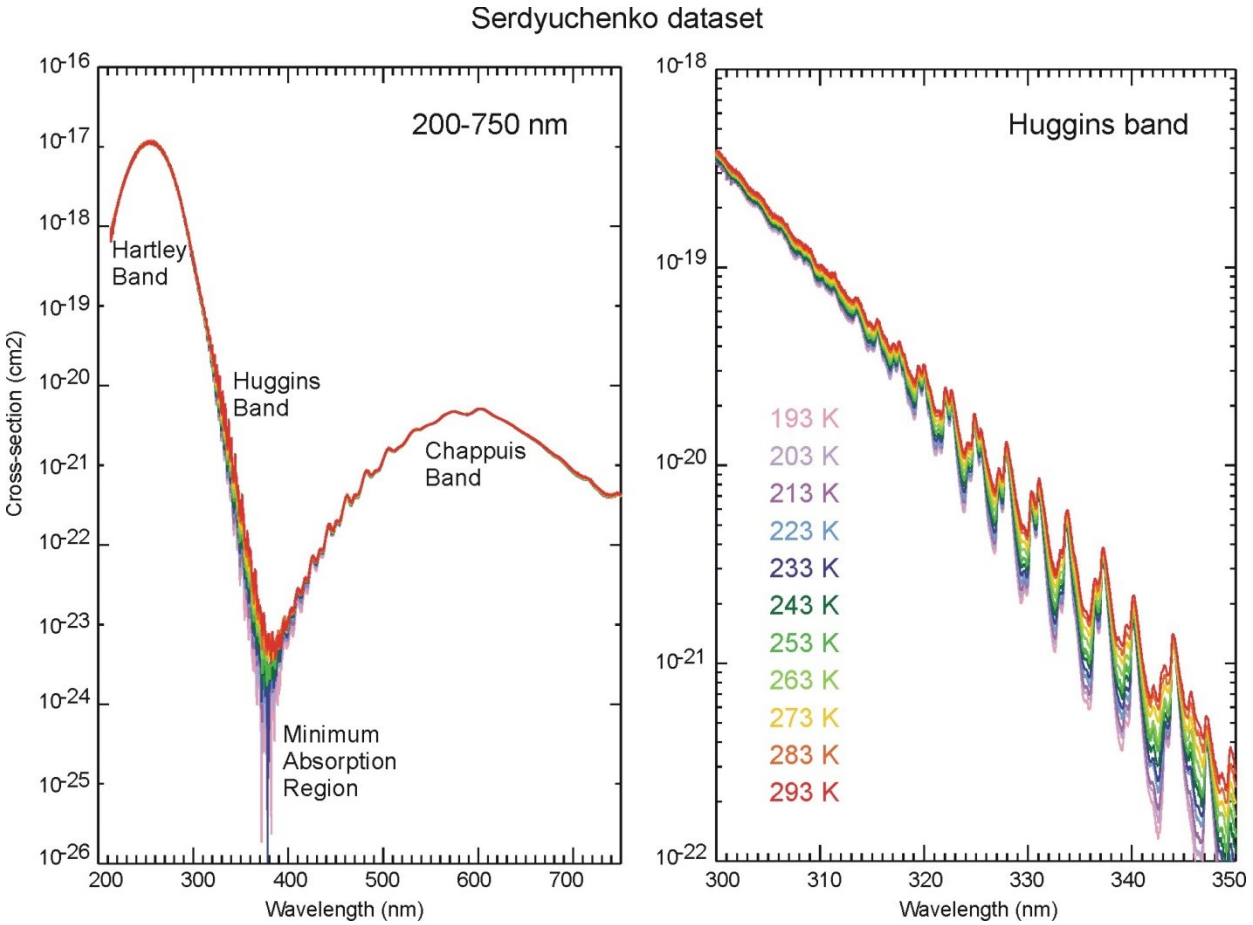


Figure E.1 Left: Example of measured ozone absorption cross-section spectrum in the UV and visible (Serdyuchenko dataset). Right: Zoomed-in to the Huggins band (300-350 nm)

E.2 Duplicate datasets

Figure E.2 shows the cross-section relative differences (%) between the duplicate datasets of Bogumil (two top rows), DMB (middle panels, bottom), and BP (bottom left panels) as a function of wavelength and for selected temperatures. For the BP datasets, differences exist over the entire 300-340 nm range for which cross-section values are available, but mainly affect the 330-340 nm region. For the Bogumil datasets, differences exist over the entire 200-750 nm region of interest. The UV region is shown in the left column of plots, the region of minimum absorption is shown in the center, and the Chappuis band is shown in the right column. For the DMB datasets, differences exist only in the region of minimum absorption (370-400 nm). For the DMB and BP datasets, the differences are significantly reduced (two bottom plots) when the spectra convoluted beforehand by a Gaussian window of full-width at half-max (FWHM) of 0.5 nm (equivalent to typical lidar interference filter widths or laser line-widths).

Figure E.3 shows the cross-section relative differences between the duplicate datasets as a function of temperature for selected wavelengths. The cross-section temperature dependence was either fitted or interpolated (see **section E.5** below for details). The relative differences between the original cross-sections (i.e., neither interpolated nor fitted) are represented by triangle symbols in the top-right panel (DMB) together with the differences between the interpolated (dotted curves) and fitted (solid curves) cross-sections. Symbols are not included in the three other panels to avoid overloading the plots. As will be explained in **section E.5**, quadratic fits usually provide the best (or at least smoothest) account of temperature dependence. We will therefore limit the present discussion to the "fitted" datasets.

Large relative differences (>20%) are found in the Huggins band for the Bogumil duplicate datasets (left panels). Smaller differences were found for the BP and DMB datasets (0-8% for BP and 0-2% for DMB). There will be no attempt here to interpret the calculated differences (beyond the scope of this paper). For brevity, we will only use one of the duplicate datasets in the rest of the comparisons. For Bogumil, we will use the latest available version, i.e., version 4 (referred to as "Bogumilv4" thereafter). For DMB we will use the smooth dataset ("DMB-smooth", referred to as "DMB" thereafter), and for BP, we will use the dataset obtained from the ACSO website (BP-ACSO, referred to as "BP" thereafter). Users interested in the behavior of the earlier Bogumil versions, the DMB raw dataset, and the BP dataset from MPI-Mainz, should first refer to the behavior of the datasets used by default (comparisons thereafter), then use **Figure E.2** and **Figure E.3**

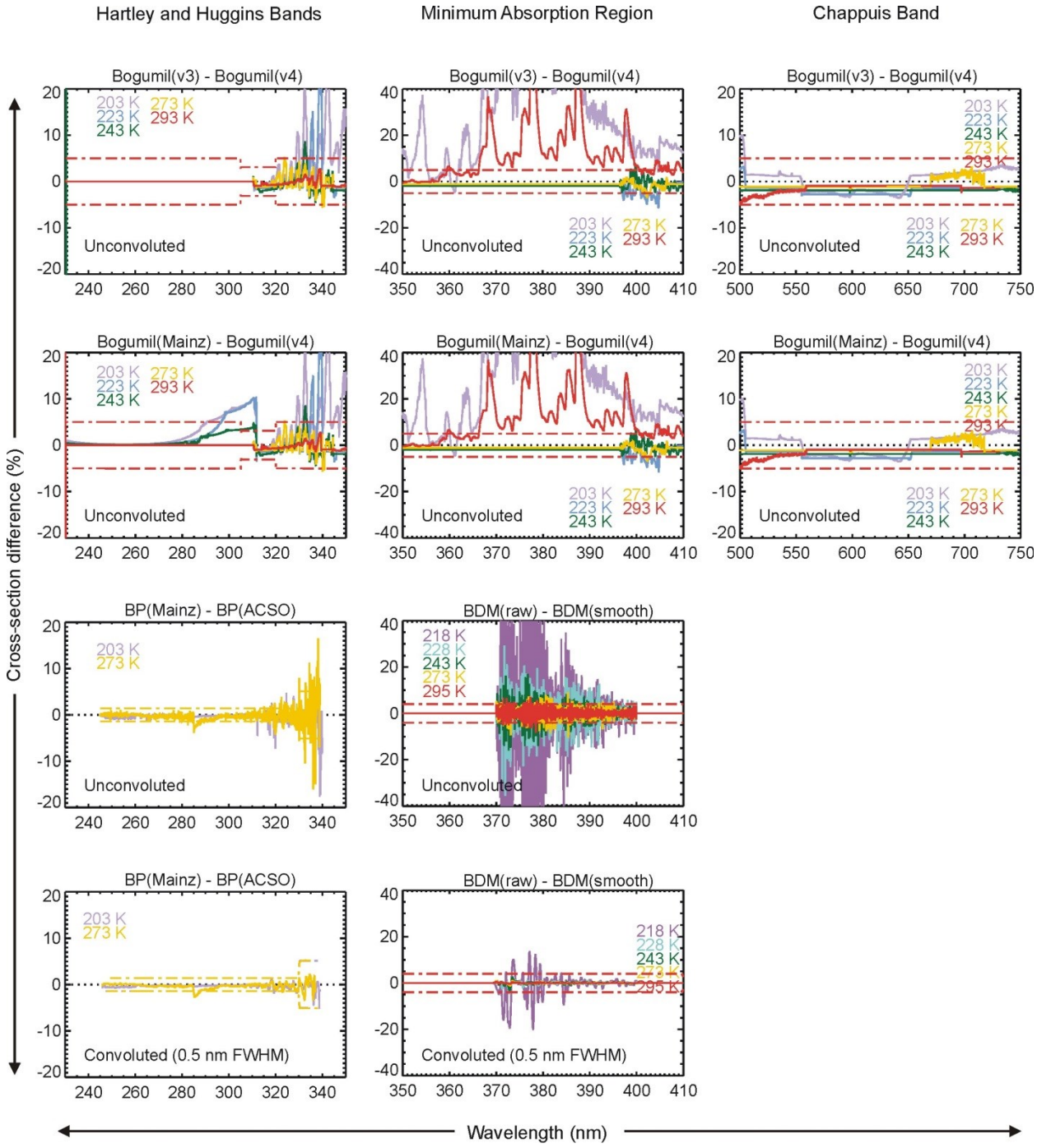


Figure E.2 Relative difference (%) in ozone cross-section between the "duplicate" datasets of Bogumil (top two rows), Bass and Paur (two bottom-left panels), and DMB (two bottom-center panels) for three different spectral regions (Hartley and Huggins bands on the left side, minimum absorption region in the center, and Chappuis band on the right side)

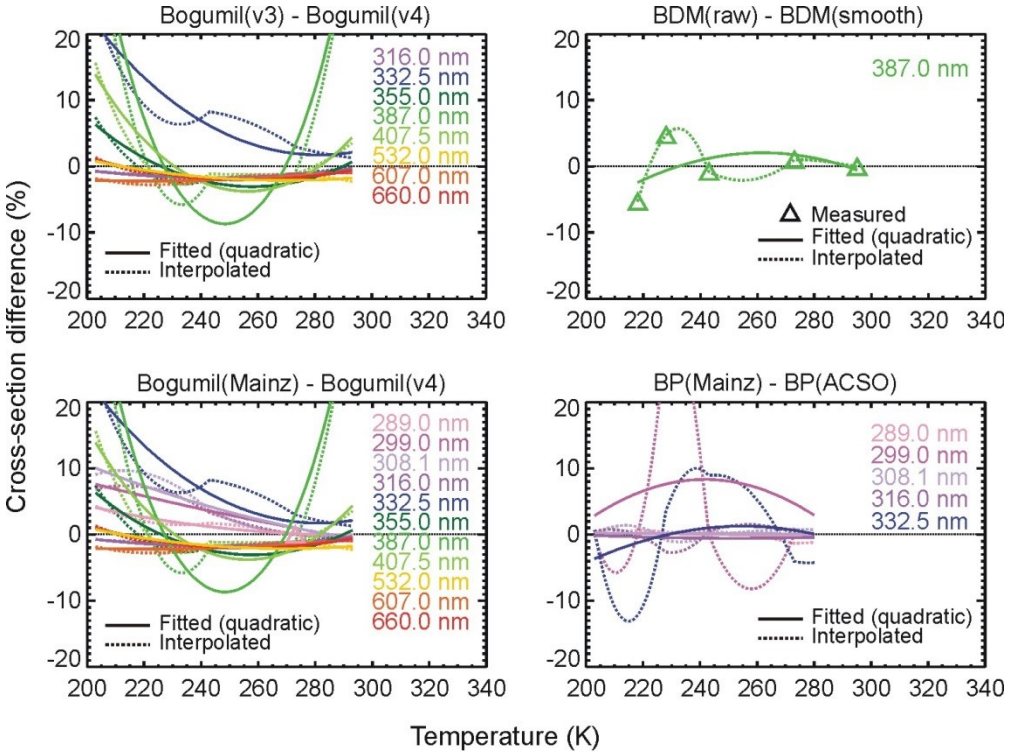


Figure E.3 Relative difference (%) in ozone cross-section between the "duplicate" datasets of Bogumil (left panels), DMB (top right panel), and Bass and Paur (bottom right panel), as a function of temperature and for selected lidar wavelengths

E.3 Spectral and temperature coverage

Figure E.4 shows the spectral and temperature coverage of all eight compared datasets. The most complete dataset is by Serdyuchenko with a full spectral coverage in the 200-750 nm region, and with the largest temperature coverage, ranging from 193 K to 293 K. The Bogumil, Voigt, and Burrows datasets also have a large spectral coverage, but with a smaller temperature coverage than that of Serdyuchenko. The DMB dataset has poor temperature coverage in the Chappuis band. The Molina and BP datasets cover only the UV region, and the Burkholder dataset covers only the Chappuis band.

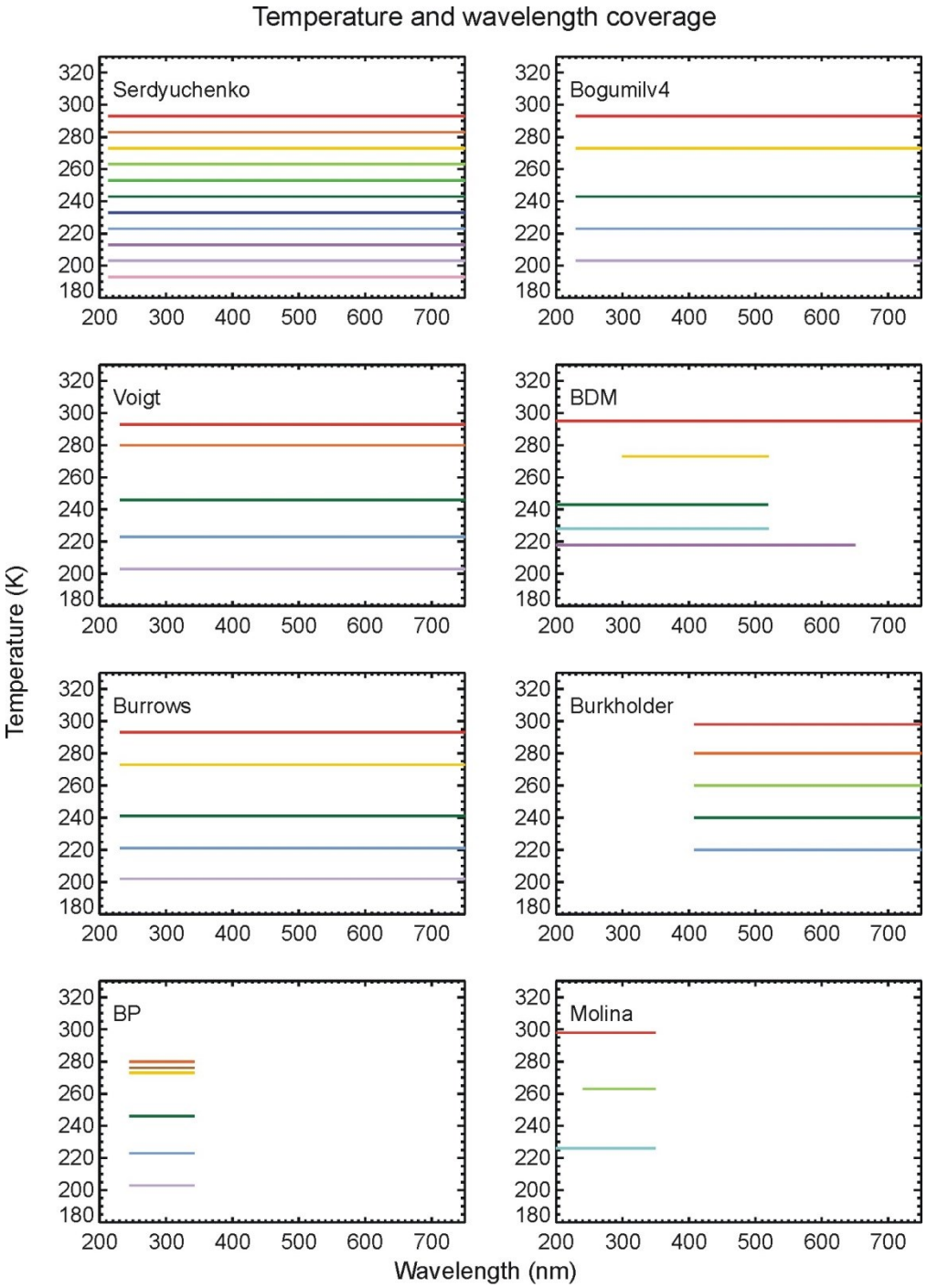


Figure E.4 Spectral and temperature coverage of the ozone absorption cross-sections for all eight available datasets (coverage is represented by horizontal lines where data is available)

E.4 Spectral resolution

Figure E.5 shows, for all eight compared datasets, the spectral resolution at which the cross-section data are provided, which is not necessarily the measurements sampling resolution. They range from a high resolution of 0.01 nm (Serdyuchenko, DMB) to very low resolutions of 0.5 nm (Molina) and 1 nm (Burkholder). For meaningful comparisons, all datasets except DMB were

interpolated onto a 0.01 nm resolution grid ranging from 200.00 nm to 750.00 nm (wavelengths in air) using a smooth, 3-point running cubic spline. For the DMB datasets, this interpolation is not necessary since the data are already provided on this grid. Before the cross-sections were interpolated, the values of wavelength in vacuum provided in the Serdyuchenko, Bogumil and Voigt datasets were converted to wavelength in air using the dispersion formula of Edlén (1966). Uncertainty associated with interpolation was estimated by comparing interpolated values and their closest measured neighbor. For most datasets, differences do not exceed 0.2% (not shown). Exceptions are the Voigt dataset for which differences reach 5% in the region of minimum absorption (350-420 nm) and in the Hartley band (200-250 nm), the Burrows dataset for which differences reach 1-3% in the region of minimum absorption (350-400 nm), and the Bogumil dataset for which the differences reach 1% in the region of minimum absorption (350-400 nm). All the differences can be explained by the combination of lower sampling resolution and larger statistical noise found in the original datasets in these spectral regions.

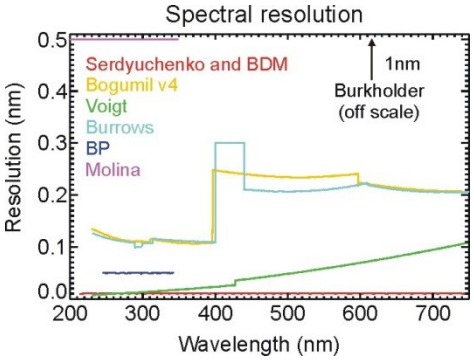


Figure E.5 Spectral resolution (in nm) of the ozone absorption cross-sections of all eight compared datasets

E.5 Temperature dependence

The electronic energy transitions of the ozone molecule responsible for the broad Hartley band is well described, but not all vibrational assignments in the Huggins Band are known to date (Qu et al, 2004; O'Keefe et al., 2001). Therefore, only empirical models have been used so far to quantify the temperature dependence of the ozone cross-section spectrum. Several fitting functions (polynomials, exponentials, and combinations of the two) have been used in the past. Consensus today favors the use of 2nd degree polynomials (Orphal, 2002). For some datasets (e.g., BP), the coefficients of the polynomials are publicly available at each wavelength. In the present work, we used our own coefficients, obtained by fitting the temperature-dependent, spectrally-interpolated (0.01 nm resolution) cross-sections with 2nd degree polynomials. For each of the eight datasets, the coefficients of the quadratic fits are used to re-grid the cross-section values between the lowest and highest temperatures available (no extrapolation) with a resolution of 0.1 K. The quadratic fitting procedure is replaced by a linear interpolation when only two measurement temperatures are available (e.g., DMB in the Chappuis band).

To complement the results obtained from the quadratic fits, the cross-sections were also interpolated onto a 0.1 K temperature grid. The interpolation scheme is a smooth, running cubic-spline. Interpolation and fitting yield almost identical results, except in the region of minimum absorption, where the sensitivity is the lowest and can lead to potentially large errors and a non-

monotonic temperature dependence. Unlike the interpolation procedure, the quadratic fitting procedure conserves monotonicity, with only rare exceptions at low temperatures in the region of minimum absorption.

The cross-section changes with temperature relative to their values at 273 K are shown in **Figure E**.6 (respectively **Figure E**.7) for selected wavelengths in the Huggins band (respectively the region of minimum absorption). The cross-section changes with temperature relative to their values at 293 K for selected wavelengths in the Chappuis band are shown in **Figure E**.8. The values of 273 K and 293 K were chosen because they allow the largest number of matching pairs in their respective spectral region.

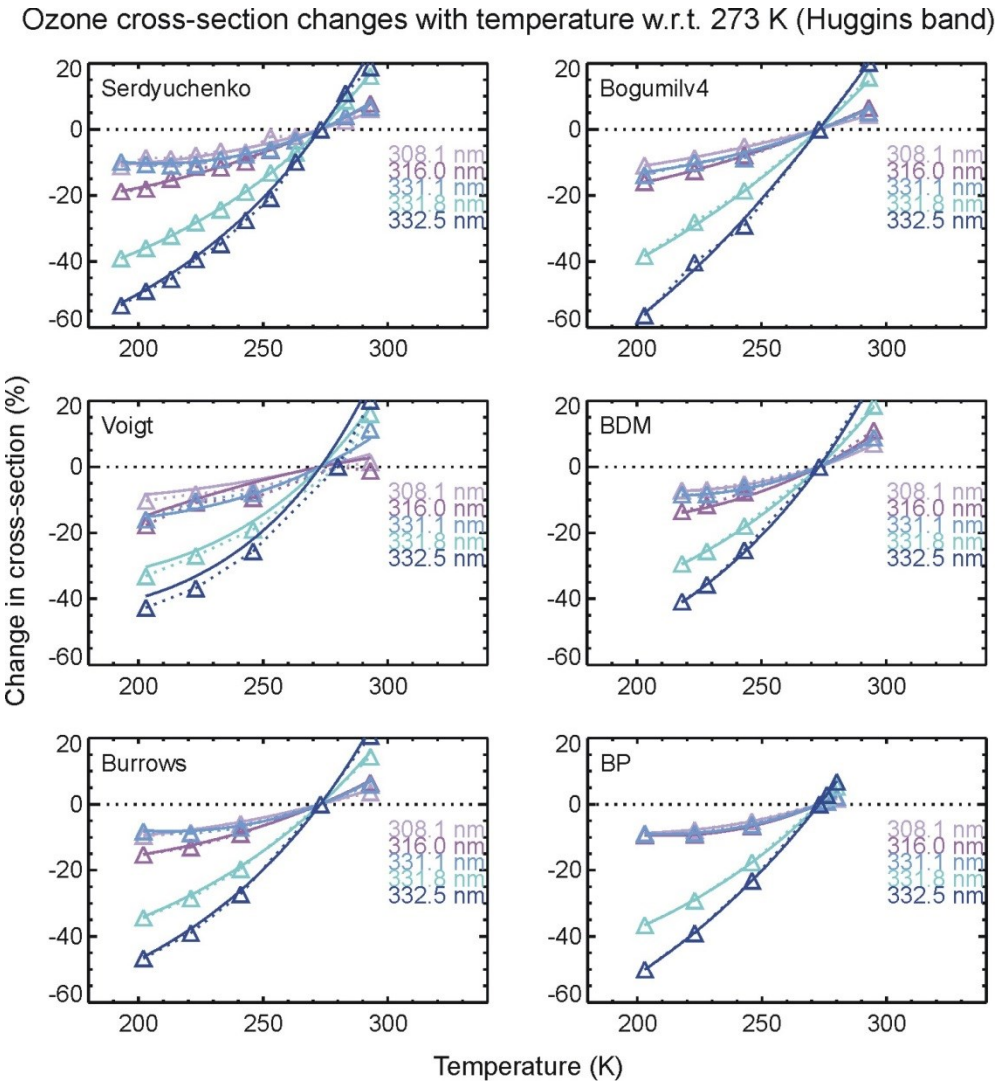


Figure E.6 Cross-section changes (%) with respect to their value at 273 K for selected wavelengths in the Huggins band. The triangle symbols indicate the spectrally-interpolated values (no temperature fitting or interpolation). The solid curves denote the spectrally-interpolated and temperature-fitted values. The dotted curves denote the spectrally-interpolated and temperature-interpolated values (see text for details)

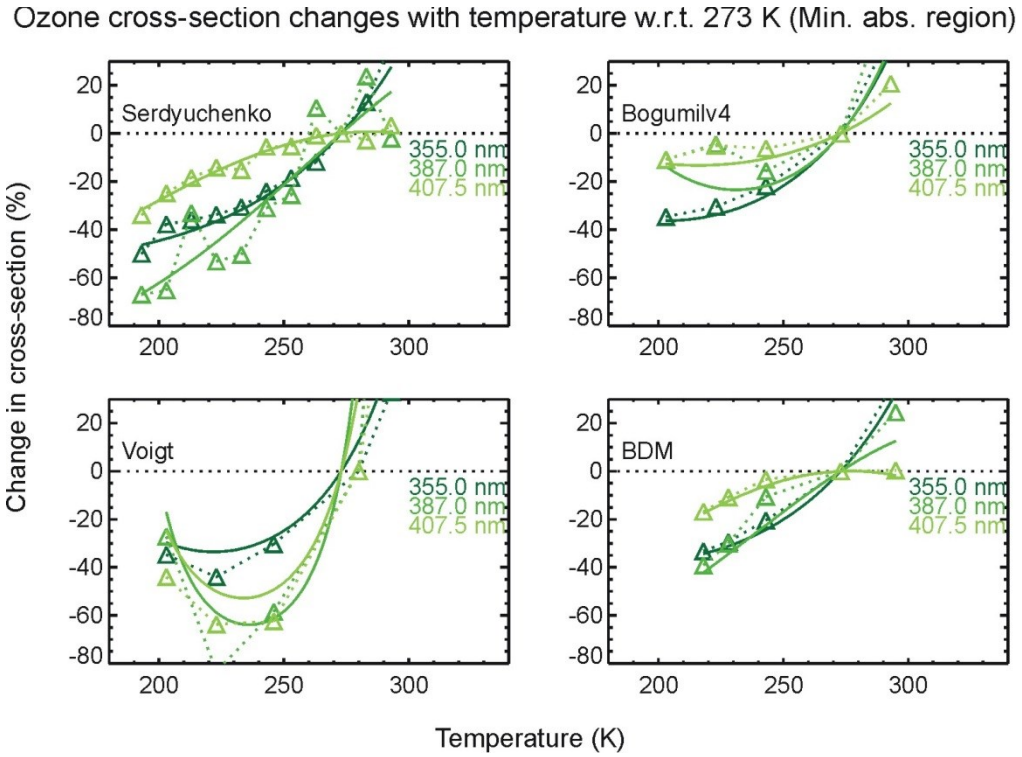


Figure E.7 Cross-section changes (%) with respect to their value at 273 K for selected wavelengths in the region of minimum absorption. The triangle symbols indicate the spectrally-interpolated values (no temperature fitting or interpolation). The solid curves denote the spectrally-interpolated and temperature-fitted values. The dotted curves denote the spectrally-interpolated and temperature-interpolated values (see text for details)

In the Huggins band, the temperature dependence near 332 nm is highly variable, as can be seen by the significantly different curve shapes in **Figure E.6** for 331.1 nm, 331.8 nm, and 332.5 nm curves). This highlights the importance of using high-resolution cross-section data, and of having a precise knowledge of the lidar receiver's filters' center line and bandwidth.

In the region of minimum absorption, the relative changes in cross-section are strongly affected by measurement precision, resulting in higher order structures (non-monotonic temperature dependence). Despite high spectral resolution and sampling over a higher number of temperatures, the Serdyuchenko dataset appears noisier than other datasets (e.g., DMB and BP). Finally, all datasets extending below 220 K except Serdyuchenko's show a reversal in the sign of the temperature dependence below 220-230 K, with cross-sections increasing with decreasing temperature. Considering the low signal-to-noise ratio at these very low cross-section values, this sign reversal must be interpreted with caution.

In the Chapuis band (**Figure E**.8), the temperature dependence is more pronounced in the red tail (660 nm) with changes up to 2.5% between room temperature and 200 K. The Serdyuchenko, DMB and Burkholder datasets agree best, while the Burrows dataset acts as an outlier by showing an increase in cross-section at very low temperatures at 660 nm.

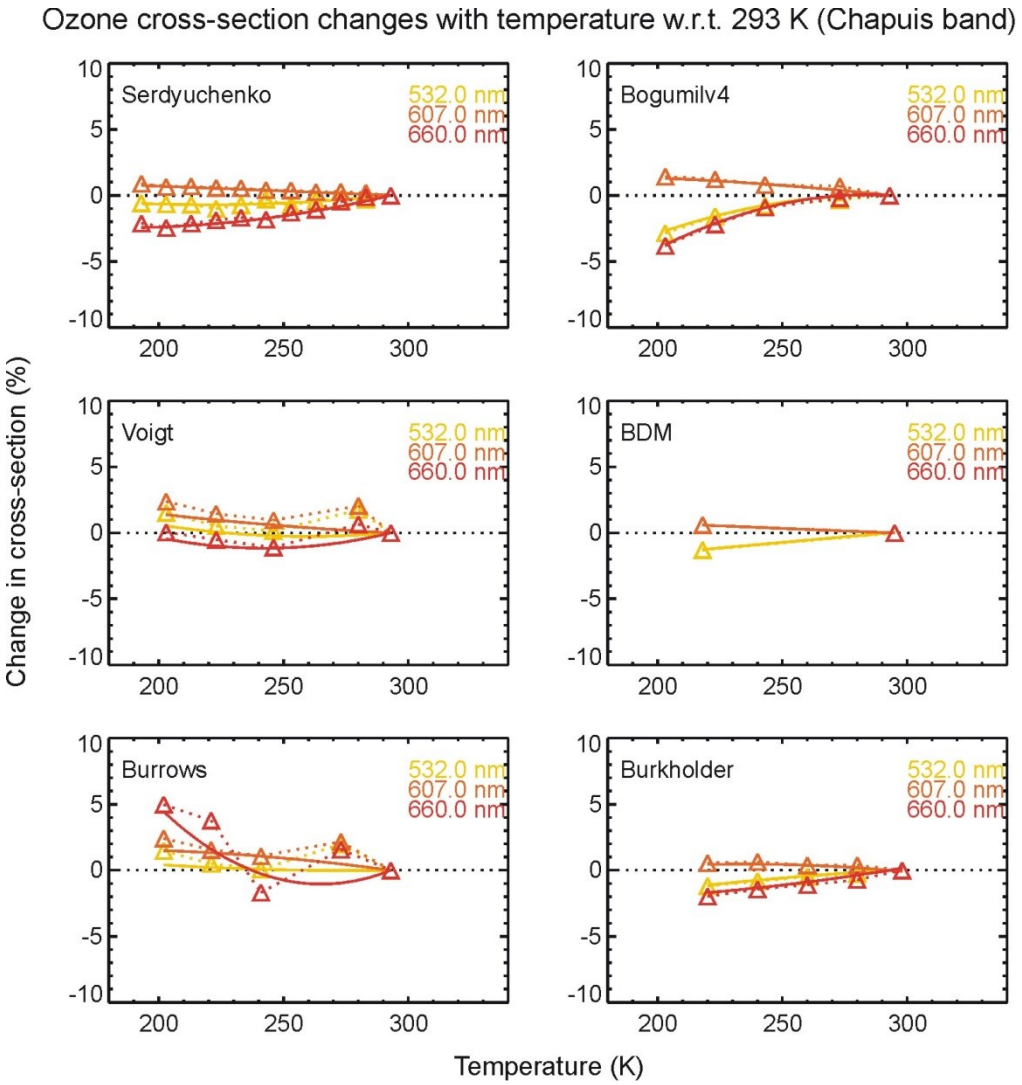


Figure E.8 Cross-section changes (%) with respect to their value at 293 K for selected wavelengths in the Chappuis band. Triangles indicate the spectrally-interpolated values connected by dotted lines. The solid curves denote the spectrally- and temperature-interpolated values

E.6 Combined effect of spectral resolution and temperature dependence

The ozone absorption cross-sections depend more or less on temperature over the entire 200-750 nm region of interest, but in relative magnitude, this dependence is much more pronounced in the 300-450 nm region. The magnitude of the changes is illustrated in the two panels of **Figure E**.9. Changes can reach up to 1%/K, which translates into 50%-100% change between the atmospheric temperature extremes of 190 K and 290 K. The right panels of **Figure E**.9 shows the magnitude of the changes in a narrow spectral window of the Huggins band for one high-resolution dataset (DMB) and one low-resolution dataset (Molina). The data is presented on the 0.01 nm-resolution spectral interpolation grid (i.e., oversampled for Molina). Not surprisingly, the low-resolution dataset of Molina misses the secondary structures seen by DMB at 330.8 nm

and 333.4 nm, which results in the underestimation of the temperature dependence of the cross-section by up to 30% (e.g., 0.3%/K instead of 0.4%/K at 333.5 nm).

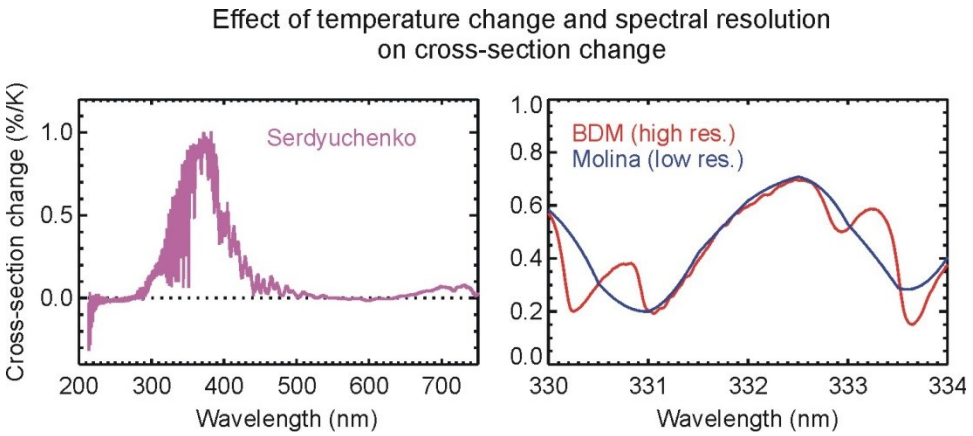


Figure E.9 Cross-section relative changes (%/K) caused by temperature changes (calculated from the difference between the cross-section values at the highest available temperature). Left: Serdyuchenko dataset (193-293 K). Right: DMB dataset (218-295 K) and Molina dataset (226-298K). See text for details

E.7 Comparison between interpolated datasets

Spectrally-interpolated data (0.01 nm resolution), fitted for temperature dependence and re-gridded every 0.1 K, are now compared. When interpreting the results, it is important to keep in mind the limitations of the original spectral resolution and the questionable temperature dependence of some datasets. Inspection of all possible combinations of matching pairs (8*7=56 comparison pairs) facilitates the identification of features that are unique and/or systematic to specific datasets. It is impossible however to show all of the matching pairs here. We will therefore only show key comparisons in view of our conclusions and recommendations on the use of specific datasets. The four plots of **Figure E**.10 show the cross-section relative differences (in %) between the Serdyuchenko, Bogumil v4, DMB and BP datasets in the UV region. The dash-dotted curves on these plots (as well as all future plots where they appear) indicate the best uncertainty estimates, as reported in the literature, for the dataset listed second in the difference (i.e., the "reference" dataset). The calculated differences remain within 3% for wavelengths shorter than 310 nm, but cannot be characterized by constant biases. Therefore, differences often calculated and reported in the literature at the reference wavelength of 253.7 nm (Hg line) do not necessarily reflect the differences found at other wavelengths, even within the same region (e.g., Hartley band). In the 320-420 nm region, the differences observed in **Figure E**.10 are highly structured and can reach 30%, maximizing in the region of minimum absorption (low measurement sensitivity). The structure and magnitude of these large differences are discussed in the next subsection.

Figure E.11 shows the cross-section relative differences between the Serdyuchenko, Bogumil, DMB and Burkholder datasets in the Chappuis band. There are no more highly structured differences like in the 320-410 nm region, but instead clear apparent biases are visible, especially between Serdyuchenko and the other datasets (1% with Burkholder, 2% with DMB, and 5% with Bogumil). The differences observed at wavelengths longer than 675 nm are not discussed here because they do not impact the ozone DIAL, temperature and water vapor Raman lidar retrievals.

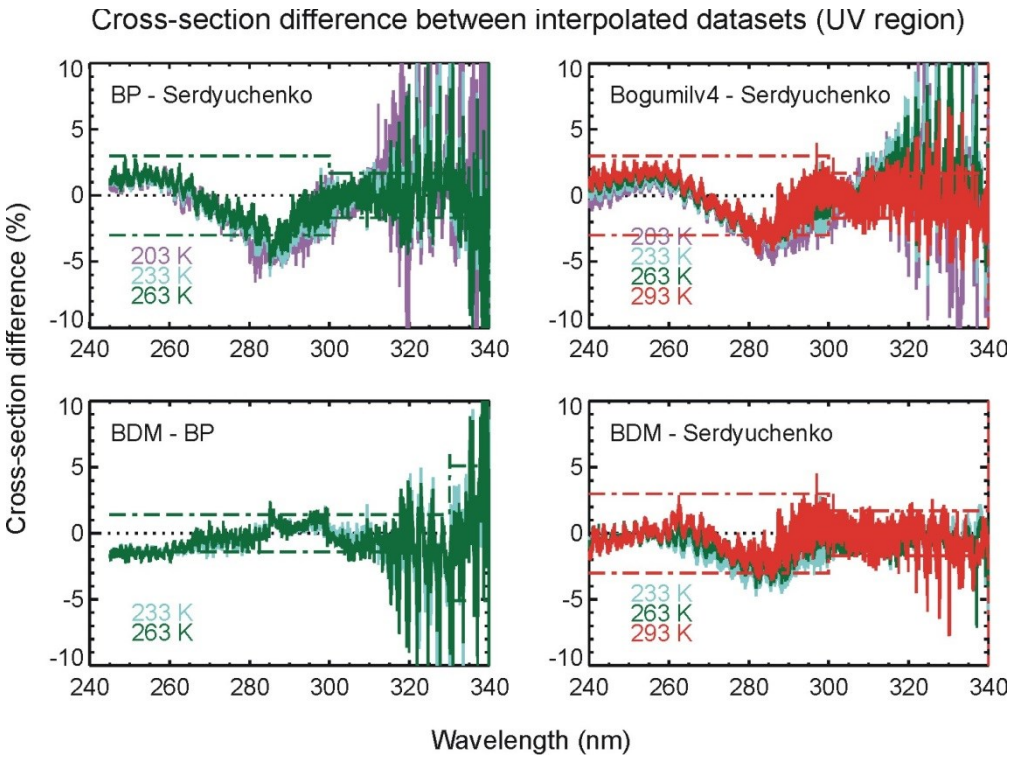


Figure E.10 Cross-section relative differences (%) between selected datasets (Serdyuchenko, Bogumilv4, DMB and BP) in the UV region for selected temperatures, when available

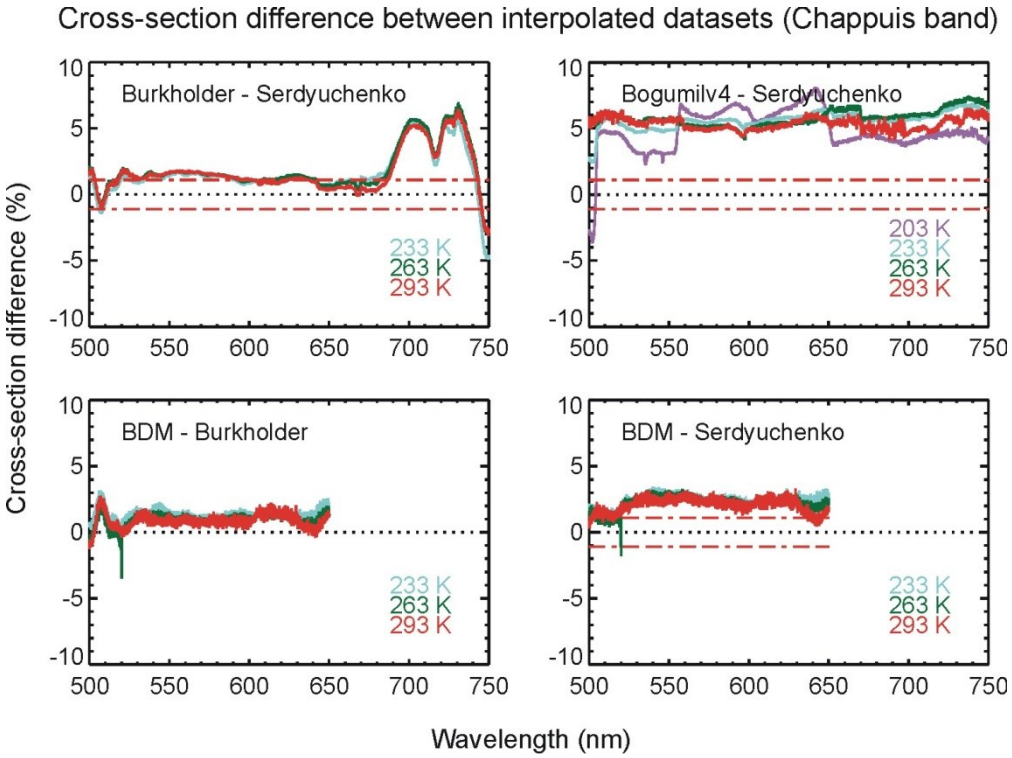


Figure E.11 Cross-section relative differences (%) between selected datasets (Serdyuchenko, Bogumilv4, DMB and Burkholder) in the Chappuis band for selected temperatures, when available

E.8 Effect of wavelength shifts, baseline and cross-section calibration

The large and highly-structured differences between the datasets seen in the Huggins band and in the region of minimum absorption (**Figure E**.10) are mostly related to the loss of measurement sensitivity, and to slight differences in sampling wavelength calibration between the datasets (Orphal, 2003). On the other hand, the broadly-structured differences observed in **Figure E**.10 (Hartley and Huggins bands) and the biases observed in **Figure E**.11 (Chappuis band) are mostly related to differences in cross-section calibration and baseline determination. In order to identify and quantify such differences, each dataset was spectrally-shifted and their cross-section values re-scaled until the relative difference between two compared datasets was found minimal. The minimization method was applied to the sum of the squared differences of the logarithm of the cross-section values spectrally shifted and renormalized using finite spectral resolution steps of 0.01 nm and finite rescaling steps of 0.01%. There is no need to provide higher scaling and shifting resolutions since the resulting discretization errors are already well below the measurement standard deviations (typically of the order of 2%).

Figure E.12 shows the shifts and scaling factors between selected datasets, as calculated in the UV region. The computation was performed using three different spectral windows: 260-310 nm, 300-340 nm, and 323-340 nm. The spectral resolution of the original datasets impacts significantly the consistency of the results. The results are much more consistent for the high resolution datasets (DMB and Serdyuchenko). The results from the two computations performed over the Huggins band are also consistent. Larger wavelength shifts but with reduced consistency were computed in the Hartley band. Consistent shifts of 0.01(+/-0.005) nm and scaling factors of 0.5-1.5% are found between the DMB and Serdyuchenko datasets. Consistent shifts of 0.02-0.03(+/-0.005) nm are observed between DMB and BP. Shifts of 0.02-0.03 nm between DMB and Molina with 0.5-3.5% scaling factors, and shifts of 0.03-0.04 nm between Molina and BP with 1-3% scaling factors were found consistently at 233 K and 263 K.

Figure E.13 shows the shifts and scaling factors between selected datasets calculated in the Chappuis band. The computation was performed for two different spectral windows. The wavelength shifts were found to strongly depend on temperature while the scaling factors remain consistent for multiple temperatures. Because of the slow variation of cross-section with wavelength in the Chappuis band, the wavelength shifts obtained from the minimization technique have larger uncertainties in this spectral region. In addition, calculated wavelength shifts smaller than 0.5 nm are meaningless for any comparison involving the Burkholder dataset due to the latter's low spectral resolution. Therefore only the scaling factors, which here reflect the magnitude of the biases seen in **Figure E**.11, can really be trusted. They are characterized by a 1% low bias for Serdyuchenko relative to Burkholder, a 1.0-1.5% low bias for Burkholder relative to DMB, and logically a 2.0-2.5% bias between Serdyuchenko and DMB.

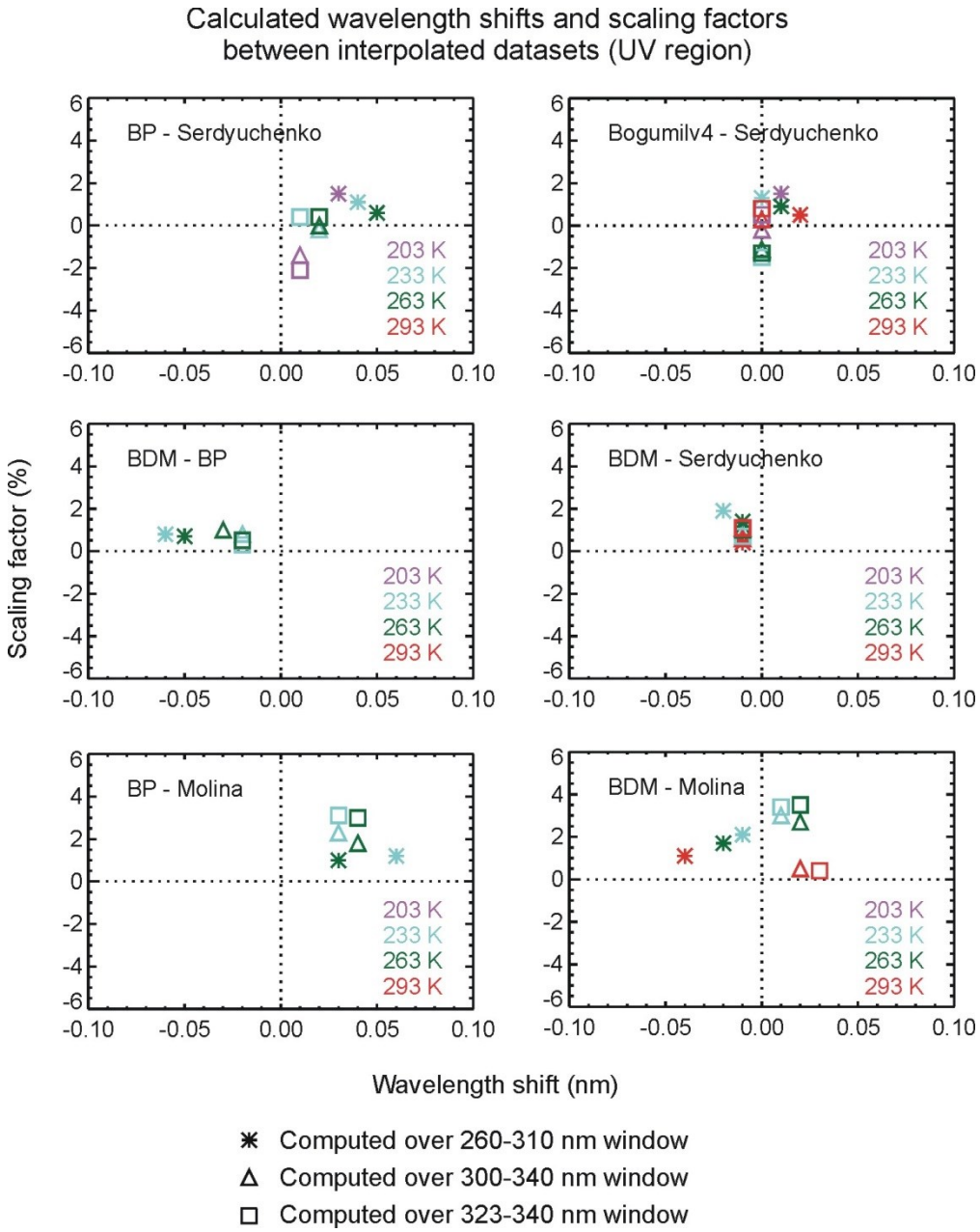


Figure E.12 Wavelength shifts and cross-section scaling factors between selected datasets (Serdyuchenko, Bogumil v4, DMB, BP, and Molina) calculated in the UV region with a minimization technique of 0.01 nm by 0.01% finite step resolution (see text for details)

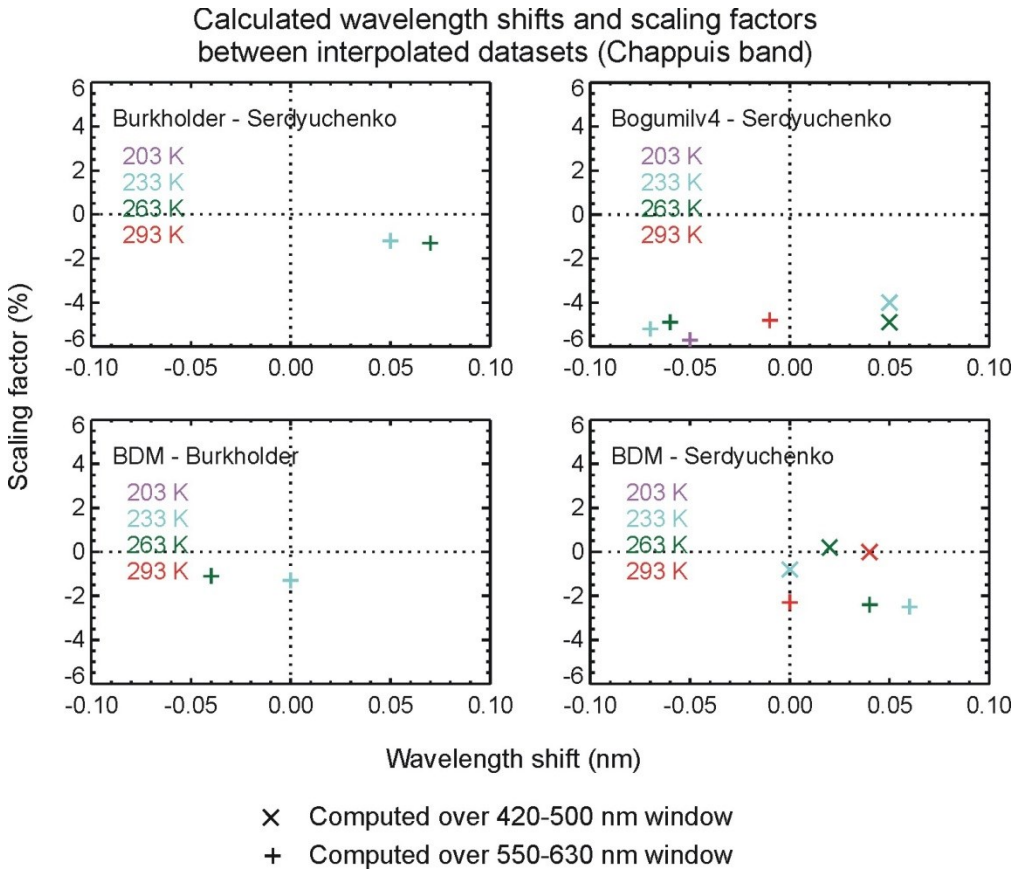


Figure E.13 Wavelength shifts and cross-section scaling factors between selected datasets (Serdyuchenko, Bogumil v4, DMB and Burkholder) calculated in the Chappuis band with a minimization technique of 0.01 nm by 0.01% finite step resolution (see text for details)

Two examples of the differences found between the DMB, BP and Serdyuchenko datasets after the calculated wavelength shifts and scaling factors were applied are plotted in **Figure E**.14. The left (respectively right) panel should be compared to the bottom-left panel of **Figure E**.10 (respectively the bottom-right panel of **Figure E**.11) showing the differences between the uncorrected datasets. Differences reaching typically 10% have now been reduced to 5% in the 315-340 nm region (left plot). Correcting the datasets produces smaller differences within a limited spectral window. However, the values of wavelength shifts and scaling calculated for different spectral windows are often inconsistent with each other. This is due to the different sensitivities of the multiple instruments used to produce the various datasets and to inconsistent baseline corrections throughout the spectrum. It is therefore not recommended to use a "corrected" dataset due to the lack of understanding of all the underlying effects producing shifts and calibration differences. However this is not a limitation for ozone DIAL, temperature, and water vapor Raman lidar applications because, as we will see in the next sub-section, the laser line-widths and filter bandwidths are typically much larger (e.g., x10) than the calculated wavelength shifts, and therefore act as a smoothing device reducing the cross-section differences by an order of magnitude similar to that of the effect of correcting for shifts.

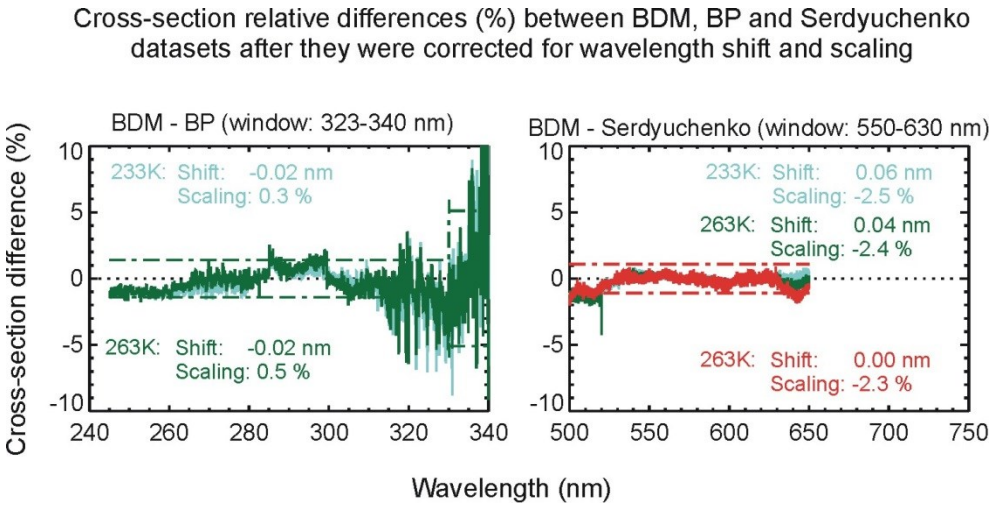


Figure E.14 Cross-section relative differences (%) between the Serdyuchenko, DMB and BP datasets in the UV region (left) and Chappuis band (right) for selected temperatures and after the datasets were corrected for wavelength shifts and cross-section scaling. Compare the left (respectively right) plot with the bottom left (respectively bottom right) plot of figure 4.4.7a (respectively 4.4.7b) showing the same differences, calculated between the uncorrected datasets

E.9 Effect of spectral filtering

For ozone and temperature lidar applications of interest, the laser and/or receiver spectral widths can be wider than 0.01 nm. Our review of the datasets must therefore not only cover the differences between spectra taken at 0.01 nm resolution, but also after they are convolved with realistic spectral windows, typically anything between 0.05 nm and 1 nm. **Figure E**.15 is identical to **Figure E**.10, but each dataset was convoluted beforehand with a Gaussian window of Full Width at Half Maximum (FWHM) of 0.5 nm. In the Huggins band where differences between datasets are highly-structured, the main effect of the convolution is to reduce the magnitude of these differences. Depending on the FWHM, the magnitude can be reduced by up to a factor of two (40% in the DMB-BP case shown in **Figure E**.10 and **Figure E**.15). Outside the Huggins band, the cross-section differences do not show highly-variable structures but rather biases, and the convolution process does not significantly reduce the calculated differences. **Figure E**.15 summarizes well the relative differences found at wavelengths shorter than 320 nm between the Serdyuchenko, Bogumil, DMB and BP datasets: 2% between DMB and Serdyuchenko, 2% between Bogumil v4 and Serdyuchenko, 1.5% between DMB and BP, and 2.0% between DMB and Serdyuchenko. A systematic behavior is observed for the Serdyuchenko dataset with low values (positive differences) near 255 nm, and high values (negative differences) near 285 nm.

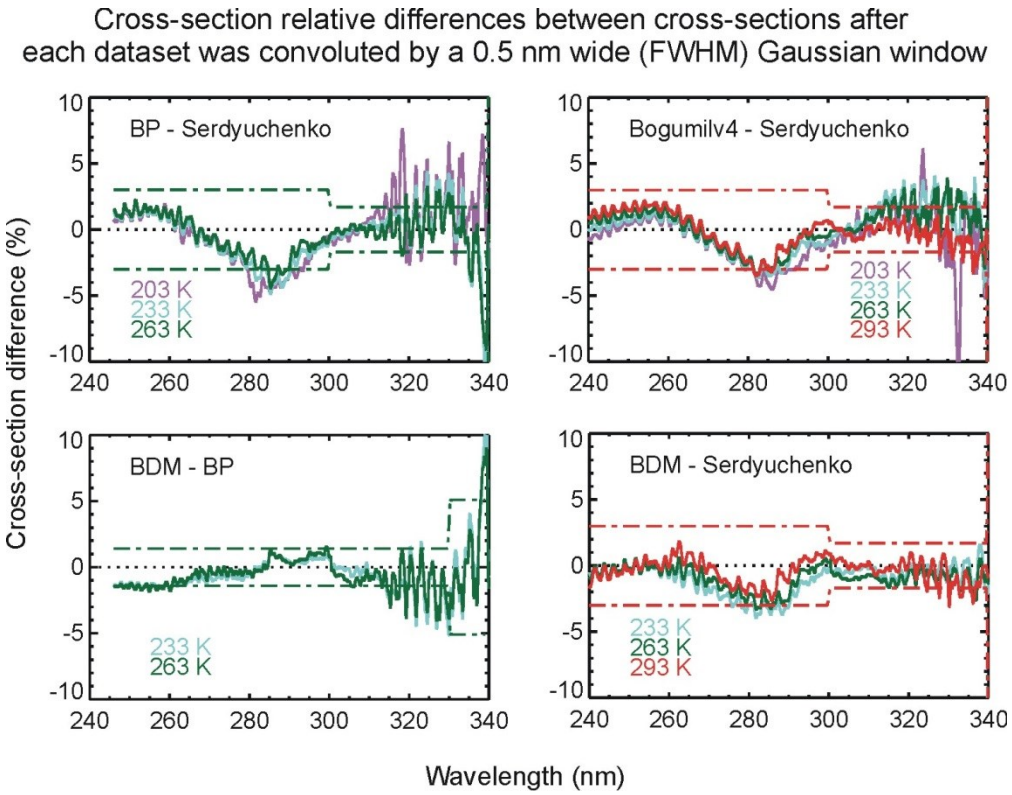


Figure E.15 Same as Figure E.10, but with each dataset first convoluted with a 0.5 nm wide Gaussian window

E.10 Effect on DIAL and Raman water vapor wavelength pairs

For ozone DIAL, it is ultimately the impact of the differences between the absorption cross-section differentials as it appears in **Eq. (1.4)**, uncertainty term #5 that must be quantified. We therefore calculated the differences in the differentials of several ozone Rayleigh and Raman DIAL pairs in the UV (from 289/299 nm to 308/355 nm), and the differences in the cross-section ratio for two water vapor Raman lidar pairs (387/407 nm and 607/660 nm). The calculated differences are plotted in **Figure E**.16 as a function of temperature. The differences calculated using the original cross-sections (no interpolation or fitting) are represented by triangle symbols wherever available (requires matching wavelength and temperature). The dotted curves represent the differences of the differentials of the cross-sections obtained after interpolating the temperature dependence, while the solid curves represent the differences of the differentials of the the cross-sections obtained after fitting (quadratic) the temperature dependence. Six ozone DIAL pairs and two water vapor Raman lidar pairs are shown. Two neighboring Raman ozone DIAL pairs (308+331.8 nm/355+387 nm and 308+332.5 nm/355+387 nm) were purposely shown in **Figure E**.16 (curves and symbols with two different shades of green) to emphasize the minor role of the backscattered wavelength 332 nm (α_{DOWN}) with respect to the emitted wavelength 308 nm (α_{UP}). **Figure E**.6 showed a significant difference (>20%) in the temperature dependence at 331.8 nm and 332.5 nm, but this difference does not appear on **Figure E**.15 because the cross-section value at 308 nm is about 10 times larger than that at 332 nm. For DIAL pairs in the Hartley and Huggins bands, differences of up to 5% are found for the 289/299 nm (tropospheric ozone) pair. Though the largest relative differences (>50%, off-scale) were found

for the pair 387/407 nm, the large magnitudes are compensated by the low absolute absorption cross-section values in this spectral region, and therefore do not impact significantly the water vapor retrieval (see our Monte-Carlo simulations in chapter 6).

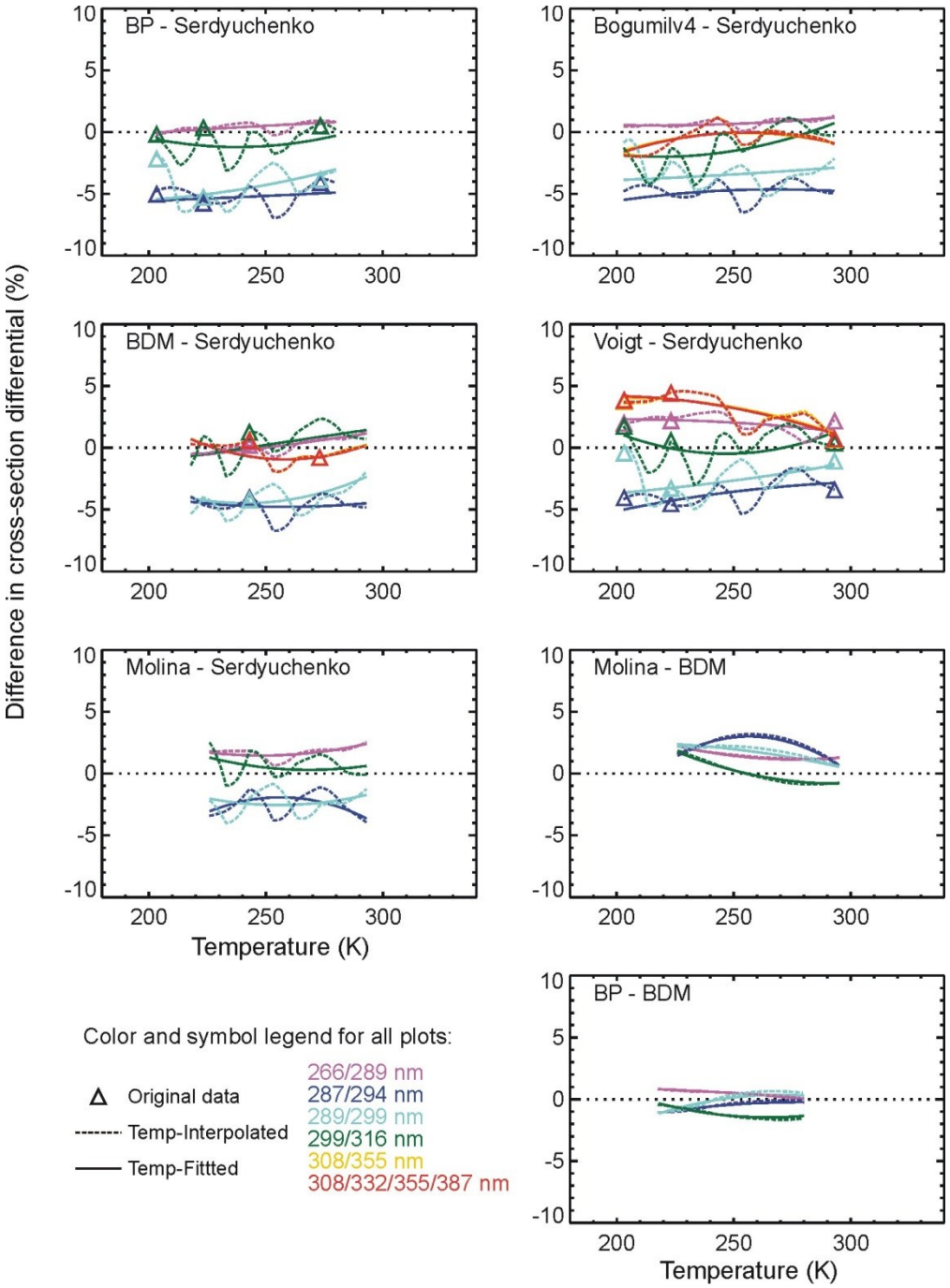


Figure E.16 Relative differences (%) between the cross-section differentials or ratios of several datasets. (six DIAL wavelength pairs and two water vapor Raman lidar wavelength pairs

E.11 Reported Uncertainties

Uncertainty estimates for the datasets reviewed in this appendix are compiled in **Table E**.1 below. For most of these datasets, the uncertainty information reported in the literature consists of lower bound and upper bound "systematic" uncertainty estimates, and a value or a range of values for random uncertainties. In most cases, both uncertainty types are reported for a broad spectral region (e.g., the Huggins band) rather than as a function of wavelength and temperature.

Table E.1: Ozone absorption cross-section uncertainties as reported in current literature

Reference	Spectral uncertainty (nm)	Random uncertainty	Systematic uncertainty (or accuracy*)	Total uncertainty	Remarks
Serdyuchenko et al., 2014	0.005	1.1-30% Hartley 1.2-14% Huggins 1.7-28% Abs. Min. 1% Chapuis	<0.4%	1.3-3% Hartley 1.7% Huggins 14-30% Abs. Min. 1.1% Chapuis	Total is reported for OD=1
Chehade et al., 2013 (Bogumil v4)					No estimates
Bogumil et al., 2004 (Bogumil v3)	0.01			>3.1% 305-320 nm >3.1% Abs. Min. <3.1% elsewhere	
Voigt et al., 2001	0.0005 230 nm 0.0072 850 nm			>7% <270 nm 4-7% 270-351 nm 3-6% >476 nm	Provided as fct. of T and λ
Daumont et al., 1992	0.005	0.9-2.2%	*1-1.5% Hartley *1-3% Huggins		
Malicet et al., 1995	0.005-0.015	0.3-2%	*1.3-1.5% Hartley *1.3-3.5% Huggins		
Brion et al., 1998	(0.005)	0.9-2%		1.5% Chapuis 4% Abs. Min.	
Burrows et al., 1997	0.0002-0.03	<1%		2.6% Hartley 2.6% Huggins >2.6% elsewhere	
Burkholder and Talukdar, 1994					Only experim. estimates
Molina and Molina, 1986	<0.05		*1% Hartley *2% Huggins		
Bass and Paur, 1984	<0.025	1%	1% 245-330 nm 5% 335-337.5 nm		

F Computing uncertainty based on geophysical standard deviation of ancillary datasets

In the absence of proper documentation on uncertainty for a given input quantity X (e.g., ancillary air temperature $X=T_a$ or ancillary NO_2 number density $X=N_{NO2}$), the uncertainty at an altitude z should be computed from the input quantity's horizontal and temporal standard deviations $\sigma_{X(\Delta x)}$, $\sigma_{X(\Delta y)}$ and $\sigma_{X(\Delta t)}$ calculated over appropriate intervals Δx, Δy, and Δt characterizing this quantity's variability. In this **appendix**, we provide simple recommendations on a standardized way to compute this uncertainty based on the type of ancillary dataset used. In most cases, the recommendations herein can be used in conjunction with a more sophisticated collocation uncertainty model, for example the statistical model presented in Fassò et al. (2014).

F.1 Actual measurements at a single geo-location $X(z,t)$

Examples of such correlative measurements are balloon-borne measurements (radiosonde, ozonesonde) or measurements from other ground-based instruments.

F.1.1 Simultaneous and co-located

No horizontal information is needed. Uncertainty can be computed based on the input quantity's standard deviation $\sigma_{X(\delta t)}$ calculated over the smallest scale of temporal variability δt:

$$u_X(z) = \sigma_{X(\delta t)}(z) \tag{F.1}$$

F.1.2 Co-located, with ancillary profile found within the time interval Δt

No horizontal information is needed. Uncertainty can be computed based on the input quantity's standard deviation $\sigma_{X(\Delta t)}$ calculated over a scale of temporal variability similar to that of the time difference Δt between the lidar and ancillary measurement:

$$u_X(z) = \sigma_{X(\Delta t)}(z) \tag{F.2}$$

F.1.3 Simultaneous, with ancillary profile found at distance Δd

Uncertainty should be computed based on the input quantity's horizontal standard deviation $\sigma_{X(\Delta d)}$. This standard deviation must be calculated over the horizontal distance Δd, and must be either provided as part of the dataset, or calculated using similar (repeatable) measurements at neighboring locations.

$$u_X(z) = \sigma_{X(\Delta d)} = \sqrt{\sigma^2_{X(\Delta x)}(z) + \sigma^2_{X(\Delta y)}(z)} \quad \text{with} \quad \Delta d = \sqrt{\Delta x^2 + \Delta y^2} \tag{F.3}$$

F.1.4 Ancillary profile found within time interval Δt and at distance Δd

Uncertainty should be computed based on the input quantity's temporal standard deviation $\sigma_{X(\Delta t)}$ and horizontal standard deviation $\sigma_{X(\Delta d)}$. The standard deviation values must be either calculated over the time interval Δt and horizontal distance Δd, and must be either provided as part of the dataset, or calculated using similar (repeatable) measurements at neighboring locations.

$$u_X(z) = \sigma_{X(\Delta d)} = \sqrt{\sigma^2_{X(\Delta t)}(z) + \sigma^2_{X(\Delta x)}(z) + \sigma^2_{X(\Delta y)}(z)} \quad \text{with} \quad \Delta d = \sqrt{\Delta x^2 + \Delta y^2} \tag{F.4}$$

F.2 Four-dimensional tabulated datasets *X*(*x*,*y*,*z*,*t*)

Examples of such datasets are state-of-the-art assimilation models (NCEP, ECMWF), chemistry-climate models (WACCM), or gridded satellite data (level 3). The recommended expressions are identical to the case of an actual measurement (**section F.1**).

F.2.1 Simultaneous and co-located

No horizontal information is needed. Uncertainty can be computed based on the input quantity's standard deviation $\sigma_{X(\delta t)}$ calculated over the smallest scale of temporal variability δt:

$$u_X(z) = \sigma_{X(\delta t)}(z) \tag{F.5}$$

F.2.2 Co-located, with ancillary profile found within the time interval Δt

No horizontal information is needed. Uncertainty can be computed based on the input quantity's standard deviation $\sigma_{X(\Delta t)}$ calculated over a scale of temporal variability similar to that of the time difference Δt between the lidar and ancillary dataset:

$$u_X(z) = \sigma_{X(\Delta t)}(z) \tag{F.6}$$

F.2.3 Simultaneous, with ancillary profile found at distance Δd

Uncertainty should be computed based on the input quantity's horizontal standard deviation $\sigma_{X(\Delta d)}$. This standard deviation must be calculated over the horizontal distance Δd, and must be either provided as part of the dataset, or calculated using the dataset values at neighboring locations.

$$u_X(z) = \sigma_{X(\Delta d)} = \sqrt{\sigma^2_{X(\Delta x)}(z) + \sigma^2_{X(\Delta y)}(z)} \quad \text{with} \quad \Delta d = \sqrt{\Delta x^2 + \Delta y^2} \tag{F.7}$$

F.2.4 Ancillary profile found within time interval Δt and at distance Δd

Uncertainty should be computed based on the input quantity's temporal standard deviation $\sigma_{X(\Delta t)}$ and horizontal standard deviation $\sigma_{X(\Delta d)}$. The standard deviation values must be either calculated over the time interval Δt and horizontal distance Δd, and must be either provided as part of the dataset, or calculated using the dataset values at neighboring locations.

$$u_X(z) = \sigma_{X(\Delta d)} = \sqrt{\sigma^2_{X(\Delta t)}(z) + \sigma^2_{X(\Delta x)}(z) + \sigma^2_{X(\Delta y)}(z)} \quad \text{with} \quad \Delta d = \sqrt{\Delta x^2 + \Delta y^2} \tag{F.8}$$

F.3 Three-dimensional tabulated climatological or empirical model *X*(*x*,*y*,*z*)

Examples of such datasets are .CIRA-86 or MSISE-00. The recommendations are similar to those provided before, except that the standard deviation of X along the un-sampled time dimension must be accounted for in the standard uncertainty of X.

F.3.1 Co-located

No horizontal information is needed. Uncertainty can be computed based on the input quantity's standard deviation calculated over all scales of temporal variability $\sigma_{X(\Delta t)}$, which must be provided as part of the dataset.

$$u_X(z) = \sigma_{X(\Delta t)}(z) \tag{F.9}$$

F.3.2 Ancillary profile found at distance Δd

Uncertainty should be computed based on the input quantity's temporal standard deviation $\sigma_{X(\Delta t)}$ and horizontal standard deviation $\sigma_{X(\Delta d)}$. The temporal standard deviation values must be provided as part of the dataset. The horizontal standard deviation values must be calculated over the horizontal distance Δd, and must be either provided as part of the dataset, or calculated using the dataset values at neighboring locations.

$$u_X(z) = \sigma_{X(\Delta d)} = \sqrt{\sigma^2_{X(\Delta t)}(z) + \sigma^2_{X(\Delta x)}(z) + \sigma^2_{X(\Delta y)}(z)} \quad \text{with} \quad \Delta d = \sqrt{\Delta x^2 + \Delta y^2} \tag{F.10}$$

F.4 Three-dimensional zonal-mean tabulated datasets *X*(*y*,*z*,*t*)

An example of such datasets is .GOZCARDS. The recommendations are similar to those provided before, except that the standard deviation of X along the un-sampled longitudinal dimension must be accounted for in the standard uncertainty of X.

F.4.1 Simultaneous and same latitude

No temporal variability information is needed. Uncertainty can be computed based on the input quantity's standard deviation $\sigma_{X(\Delta x)}$ calculated over all scale of longitudinal variability Δx:

$$u_X(z) = \sigma_{X(\Delta x)}(z) \tag{F.11}$$

F.4.2 Simultaneous and within latitude difference Δy

No temporal variability information is needed. Uncertainty should be computed based on the input quantity's horizontal standard deviation $\sigma_{X(\Delta d)}$. The longitudinal standard deviation values must be provided as part of the dataset. The latitudinal standard deviation values must be calculated over the distance Δy, and must be either provided as part of the dataset, or calculated using the dataset values at neighboring latitudes:

$$u_X(z) = \sigma_{X(\Delta d)} = \sqrt{\sigma^2_{X(\Delta x)}(z) + \sigma^2_{X(\Delta y)}(z)} \tag{F.12}$$

F.4.3 Same latitude and within time interval Δt

No latitudinal variability information is needed. Uncertainty should be computed based on the input quantity's longitudinal standard deviation $\sigma_{X(\Delta x)}$ and temporal standard deviation $\sigma_{X(\Delta t)}$. The longitudinal standard deviation values must be provided as part of the dataset. The temporal standard deviation values must be calculated over the time interval Δt, and must be either provided as part of the dataset, or calculated using the dataset values at neighboring time stamps.

$$u_X(z) = \sqrt{\sigma^2_{X(\Delta x)}(z) + \sigma^2_{X(\Delta t)}(z)} \tag{F.13}$$

F.4.4 Ancillary profile found within time interval Δt and latitude difference Δy

Uncertainty should be computed based on the input quantity's horizontal standard deviation $\sigma_{X(\Delta d)}$ and temporal standard deviation $\sigma_{X(\Delta t)}$. The longitudinal standard deviation values must be provided as part of the dataset. The latitudinal standard deviation values must be calculated over the time interval Δy, and must be either provided as part of the dataset, or calculated using the dataset values at neighboring latitudes. The temporal standard deviation values must be calculated over the time interval Δt, and must be either provided as part of the dataset, or calculated using the dataset values at neighboring time stamps.

$$u_X(z) = \sqrt{\sigma^2_{X(\Delta x)}(z) + \sigma^2_{X(\Delta y)}(z) + \sigma^2_{X(\Delta t)}(z)} \tag{F.14}$$

F.5 Two-dimensional zonal mean climatological or empirical models *X*(*y*,*z*)

The recommendations are similar to those provided before, except that the standard deviation of X along the un-sampled longitudinal and time dimensions must be accounted for in the standard uncertainty of X.

F.5.1 Same latitude

No latitudinal variability information is needed. Uncertainty should be computed based on the input quantity's standard deviation $\sigma_{X(\Delta x)}$ calculated over all scales of longitudinal variability Δx and standard deviation $\sigma_{X(\Delta t)}$ calculated over all time scales Δt. Both the longitudinal and temporal standard deviation values must be provided as part of the dataset:

$$u_X(z) = \sqrt{\sigma^2_{X(\Delta x)}(z) + \sigma^2_{X(\Delta t)}(z)} \quad \textbf{(F.15)}$$

F.5.2 Ancillary profile found at latitude difference Δy

Uncertainty should be computed based on the input quantity's horizontal standard deviation $\sigma_{X(\Delta d)}$ and temporal standard deviation $\sigma_{X(\Delta t)}$. The longitudinal and temporal standard deviation values must be provided as part of the dataset. The latitudinal standard deviation values must be calculated over the distance Δy, and must be either provided as part of the dataset, or calculated using the dataset values at neighboring latitudes:

$$u_X(z) = \sqrt{\sigma^2_{X(\Delta t)}(z) + \sigma^2_{X(\Delta x)}(z) + \sigma^2_{X(\Delta y)}(z)} \quad \textbf{(F.16)}$$

F.6 Constant profile *X*(*z*) (e.g., global average or empirical profile)

The standard deviation of X along all dimensions must be accounted for. Uncertainty should be computed based on the input quantity's horizontal standard deviation $\sigma_{X(\Delta d)}$ and temporal standard deviation $\sigma_{X(\Delta t)}$ over all scales of variability. The longitudinal, latitudinal and temporal standard deviation values must be provided as part of the dataset.

$$u_X(z) = \sqrt{\sigma^2_{X(\Delta x)}(z) + \sigma^2_{X(\Delta y)}(z) + \sigma^2_{X(\Delta t)}(z)} \quad \textbf{(F.17)}$$

G Alternate expression of uncertainty propagation for simple uncertainty components arising from systematic and random effects

The present appendix illustrates the complexity of integrating uncertainty components arising from systematic effects, and provides a simple, practical solution to this integration. Starting from the input quantities X_n introduced in **section 2.2**, we can take the example of an uncertainty budget comprising two independent randomized components δ_n and ε_n arising from systematic effects, and one random component, the quantity's experimental standard deviation for a given sample σ_n.(n=1,N). The three components are independent of each other (uncorrelated) for a given sample, but each of the two components arising from systematic effects implies dependencies within part, or the totality of the sampling population, in other words:

Each of the δ_n (n=1,N) implies correlation with at least one other point:

$$-1 \le r_{\delta nm} \le 1 \quad (n,m=1,N)) \tag{G.1}$$

Each of the ε_n (n=1,N) implies correlation with at least one other point:

$$-1 \le r_{\varepsilon nm} \le 1 \quad (n,m=1,N)) \tag{G.2}$$

None of the σ_n (n=1,N) implies correlation with any other points:

$$r_{\sigma nm} = 0 \ \text{ for all } \ n \ne m \ \text{ and } \ r_{\sigma nm} = 1 \ \text{ for all } \ n = m) \tag{G.3}$$

Because each component is independent of each other, the combined uncertainty of an individual sample x_n is easily computed as:

$$u_n = \sqrt{\delta_n^2 + \varepsilon_n^2 + \sigma_n^2} \tag{G.4}$$

Considering the output model Y defined in **section 2.2**, computing the combined standard uncertainty u_y from the individual components u_n, requires the calculation of every covariance term (or correlation coefficient r_{nm}) in **Eq. (2.4)** or **Eq. (2.7)** owing to the introduction of the components δ_n and ε_n. Unless the complete covariance model is known, this task is impossible to achieve. However, an alternate procedure more likely to succeed consists of first propagating the uncertainty for each independent component δ_n, ε_n, σ_n, then calculating the combined uncertainty for the output model Y. The overall equation describing the process is:

$$u_y = \sqrt{\sum_{m=1}^{N}\left(\sum_{n=1}^{N}\frac{\partial y}{\partial x_n}\frac{\partial y}{\partial x_m} r_{\delta nm}\delta_n\delta_m\right) + \sum_{m=1}^{N}\left(\sum_{n=1}^{N}\frac{\partial y}{\partial x_n}\frac{\partial y}{\partial x_m} r_{\varepsilon nm}\varepsilon_n\varepsilon_m\right) + \sum_{n=1}^{N}\left(\frac{\partial y}{\partial x_n}\right)^2\sigma_n^2} \tag{G.5}$$

The first two terms under the square root provide expressions of the propagated uncertainties arising from systematic effects, and the third term provides an expression of the propagated uncertainties arising from all random effects. The advantage of this procedure is that the

correlation coefficients $r_{\delta nm}$ and $r_{\varepsilon nm}$ are more likely to be known because they characterize the actual sources of the systematic effects.

To illustrate the impact of components arising from systematic effects, two simple examples of the application of **Eq.** (2.3.1) are given below. First let's consider the input quantity X_n (n=1,2,3) with the following simplistic uncertainty budget:

$\delta_n = 1$ for all n, with $r_{\delta nm}$ =1 for all n and all m

$\varepsilon_n = 2$ for all n, with $r_{\varepsilon nm}$ =1 for all n and all m

$\sigma_n = 4$ for all n (random, uncorrelated)

For the output model Y, we first consider the simple case of a 3-point average:

$$y = \frac{1}{3}\sum_{n=1}^{3} x_n \qquad \textbf{(G.6)}$$

When all uncertainty components are treated following their correct classification (correlated and uncorrelated), the total uncertainty computed using (**Eq.** 2.3.1) is:

$$u_y = \frac{\sqrt{93}}{3}. \qquad \textbf{(G.7)}$$

If the components characterizing the systematic effects were treated as uncorrelated components, then the application of either (**Eq.** 2.2.3) or (**Eq.** 2.3.1) would yield:

$$u_y = \frac{\sqrt{63}}{3}. \qquad \textbf{(G.8)}$$

Finally if the components characterizing the systematic effects were neglected, then the application of either **Eq.** (2.2.3) or **Eq.** (2.3.1) would yield the answer:

$$u_y = \frac{\sqrt{48}}{3}. \qquad \textbf{(G.9)}$$

In this example, neglecting uncertainty components arising from systematic effects, or treating them as "uncorrelated" results in the underestimation of the combined uncertainty.

Another simple example of output model is the subtraction of two samples x_1 and x_2 (same uncertainty budget as before for the input quantities). When all uncertainty components are treated following their correct classification, the total uncertainty computed using (**Eq.** 2.3.1) is:

$$u_y = \sqrt{32}.$$

If the components characterizing the systematic effects were treated as uncorrelated components, then the application of (**Eq.** 2.2.3) or (**Eq.** 2.3.1) would yield the value:

$$u_y = \sqrt{42}.$$

In this example, treating uncertainty components arising from systematic effects as uncorrelated results in the overestimation of the combined uncertainty. If the components characterizing the systematic effects were simply neglected, then the application of either **Eq.** (2.2.3) or **Eq.** (2.3.1) would yield:

$$u_y = \sqrt{32}\,, \tag{G.10}$$

This is the same value as if all components are treated following their correct classification. This is not surprising since the subtraction of two measured quantities known to have the same value of uncertainty arising from the same systematic effect leads to the cancelation of this effect in the combined uncertainty budget.

H Estimation of uncertainty associated with paralyzable saturation correction

Following **Eq. (4.20)**, the function f for which the root x must be found is written:

$$f(x) = x\exp(c_1 x) + c_2 \tag{H.1}$$

With $x = S_1(i,k)$ $\quad c_1 = -\tau(i)\frac{2\delta z}{cL}$ $\quad c_2 = S_0(i,k)$

For the Newton-Raphson method, an estimated value of the root at iteration j+1 is:

$$x(j+1) = x(j) - g(j) \tag{H.2}$$

Where $g(j) = \frac{f(j)}{f'(j)}$, $\quad f(j) = x(j)\exp(c_1 x(j)) + c_2$, and $f'(j) = \frac{\partial f(j)}{\partial x(j)} = (1 + c_1 x(j))\exp(c_1 x(j))$

The method is initialized using a first guess value $x(0)$. At the final n^{th} iteration, the corrected signal $S_1(i,k)$ is set to the estimated value $x(n)$. If x_T is the "true" (unknown) value of the root, the residual error at iteration $j+1$ can be written:

$$\varepsilon_{S1}(j+1) = x(j+1) - x_T \tag{H.3}$$

The residual error ε_{S1} after the final n^{th} iteration is:

$$\varepsilon_{S1}(n) = -\varepsilon_{S1}^2(n-1)\frac{f''(x_T)}{f'(x_T)} = (-1)^n\left(x(0) - x_T\right)^{2n}\left(\frac{f''(x_T)}{f'(x_T)}\right)^n \tag{H.4}$$

With $f''(j) = \frac{\partial^2 f(j)}{\partial x^2(j)} = (2c_1 + c_1^2 x(j))\exp(c_1 x(j))$

If we assume that the corrected signal estimated at the last iteration $x(n)=S_1(i,k)$ is very close to the true root x_T, the standard uncertainty introduced by the root-finding method can then be taken as the absolute value of the approximated residual error at the last iteration:

$$u_{S1(ITER)}(i,k) \approx \left(x(0) - x(n)\right)^{2n}\left(\frac{f''(n)}{f'(n)}\right)^n = \left(x(0) - x(n)\right)^{2n}\left(\frac{\left(2c_1 + c_1^2 x(n)\right)}{\left(1 + c_1 x(n)\right)}\right)^n \tag{H.5}$$

The uncertainty owed to the dead-time u_τ and to the detection noise u_{S0} can be analytically propagated through the iteration process. When expanded, the function g can be written as a function of the independent variables c_1 which relates to the dead-time, and c_2 which relates to the non-corrected signal S_0. The dead-time uncertainty propagated at iteration j+1 can therefore be written:

$$u_{x(SAT)}(j+1) = \sqrt{u_{x(SAT)}^2(j) + \left(\frac{\partial g(j)}{\partial c_1}\right)^2 u_{c1}^2 - 2\,\mathrm{cov}_{SAT}\left(x(j), g(j)\right)}$$

4748 **(H.6)**

4749 with $u_{c1} = \frac{2\delta z}{cL} u_\tau$

4750 The term "cov$_{SAT}$" describes the covariance of $x(j)$ and $g(j)$ in absence of any other uncertainty
4751 sources (i.e., by setting u_{S0}=0). It can be estimated using a Monte-Carlo experiment similar to
4752 what is described in **appendix A** for the target parameter $\tau(i)$.

4753 Similarly, the uncertainty due to the detection noise propagated at the iteration j+1 can be
4754 written:

4755 $$u_{x(DET)}(j+1) = \sqrt{u_{x(DET)}^2(j) + \left(\frac{\partial g(j)}{\partial c_2}\right)^2 u_{c2}^2 - 2\,\mathrm{cov}_{DET}\left(x(j), g(j)\right)}$$

4756 **(H.7)**

4757 with $u_{c2} = u_{S0(DET)}$

4758 The term "cov$_{DET}$" describes the covariance of $x(j)$ and $g(j)$ in absence of any other uncertainty
4759 sources (i.e., by setting u_τ=0). Again it can be estimated using a Monte-Carlo experiment similar
4760 to what is described in **appendix A** for the target parameter $S_0(i,k)$. The resulting dead-time
4761 uncertainty component propagated to the corrected signal S_1 after the final n^{th} iteration can then
4762 be written:

4763 $$u_{S1(SAT)}(i,k) = u_{x(SAT)}(n)$$

4764 **(H.8)**

4765 Similarly the resulting detection noise uncertainty component propagated to the corrected signal
4766 S_1 after the final n^{th} iteration can be written:

4767 $$u_{S1(DET)}(i,k) = u_{x(DET)}(n)$$

4768 **(H.9)**

4769
4770

I Derivation of fitting coefficients uncertainty and their co-variance for general least-squares and singular value decomposition fitting methods

I.1 Least squares case

We start from **Eq. (4.30)** re-written in a more general form and using the range r along which the background noise $b(r)$ needs to be estimated:

$$S_2(r) = S_1(r) - b(r) \tag{I.1}$$

For Least-squares (LS) fitting methods, all the uncertainty and co-variance terms can be analytically calculated. These analytical solutions take a simpler form for polynomials of degree 1 or 2 (Press et al., 1986). We first define the fitting function B in the general form:

$$B(r) = \sum_{j=1}^{M} b_j X_j(r) \tag{I.2}$$

Here the sum of the M Basis Functions X_j is, for example, a combination of polynomial or exponential functions of altitude range r. We then minimize the chi-square merit function defined by:

$$\chi^2(b_0, b_1, ..., b_M) = \sum_{i=1}^{N} \left[\frac{1}{\delta S_1(r_i)} \left(S_1(r_i) - \sum_{j=1}^{M} b_j X_j(r_i) \right) \right]^2 \tag{I.3}$$

The chi-square is minimized by resolving the following M Normal Equations:

$$\sum_{i=1}^{N} \frac{1}{\delta S_1^2(r_i)} \left[S_1(r_i) - \sum_{j=1}^{M} b_j X_j(r_i) \right] X_k(r_i) = 0 \quad \text{with } k=1,2,\ldots,M \tag{I.4}$$

The coefficients of the fitting function and their uncertainty can be found analytically by transforming the Normal Equations into a matrix form:

$$\sum_{j=1}^{M} b_j \alpha_{kj} = \beta_k \tag{I.5}$$

where the (k,j) element of an M x M matrix is:

$$\alpha_{kj} = \sum_{j=1}^{N} \frac{X_j(r_i) X_k(r_i)}{\delta S_1^2(r_i)} \tag{I.6}$$

and the k^{th} element of a vector of length M is:

$$\beta_k = \sum_{i=1}^{N} \frac{S_1(r_i) X_k(r_i)}{\delta S_1^2(r_i)}$$

4802 **(I.7)**

4803 Using the inverse matrix $B_{kj} = [\alpha_{kj}]^{-1}$, the M coefficients b_j of the fitting function can be written:

4804 $$b_j = \sum_{k=1}^{M} B_{jk} \left(\sum_{i=1}^{N} \frac{S_1(r_i) X_k(r_i)}{\delta S_1^2(r_i)} \right)$$

4805 **(I.8)**

4806 The uncertainties associated with the M coefficients b_j correspond to the diagonal elements of the
4807 matrix **B**:

4808 $$u_{bj} = \sum_{i=1}^{N} \delta S_1^2(r_i) \left(\frac{\partial b_j}{\partial S_1(r_i)} \right)^2 = \sqrt{B_{jj}}$$

4809 **(I.9)**

4810 and the covariance between b_j and b_k is contained in the non-diagonal elements:

4811 $$\mathrm{cov}(b_j, b_k) = B_{jk}$$

4812 **(I.10)**

4813 More details on the general LS fitting method as well as specific solutions for the coefficients
4814 and their uncertainty for polynomials of degree 1 and 2 can be found in Press et al. (1986) as
4815 well as many other tutorials.

4816

4817 ## I.2 Singular value Decomposition case

4818 Though a LS method may work in most cases, the Normal Equations may occasionally turn out
4819 singular or close to singular. In this case, using the Singular Value Decomposition (SVD)
4820 method is more appropriate. We start by constructing the N x M elements of the so-called
4821 "design matrix" **A** of the fitting process, and the vector space b of length N as follows:

4822 $$A_{ij} = \frac{X_j(r_i)}{u_1(r_i)} \qquad c_i = \frac{S_1(r_i)}{u_1(r_i)}$$

4823 **(I.11)**

4824 Writing the coefficients *bj* of the fitting function into the vector form **b**, the minimization of χ^2
4825 can then be written in the following SVD form:

4826 $\chi^2 = |\mathbf{A} \cdot \mathbf{c} - \mathbf{b}|$ minimum $\Rightarrow \mathbf{A} \cdot \mathbf{c} = \mathbf{b}$, which can be written: $\mathbf{c} = \mathbf{V} \cdot \mathbf{W} \cdot \mathbf{U}^T \cdot \mathbf{b}$ with $\mathbf{A} = \mathbf{U} \cdot \mathbf{W} \cdot \mathbf{V}^T$

4827 **(I.12)**

4828 **U**, **W**, and **V** represent a unique set of diagonal and orthogonal matrices. **U** is an M x N column-
4829 orthogonal matrix, **W** is an N x N diagonal matrix (the singular values), and **V** is an M x M
4830 orthogonal matrix. Once the SVD system is solved (i.e., **U**, **V**, and **W** are found), the coefficients
4831 b_j of the fitting function, their uncertainty u_{bj}, and the covariance between b_j and b_k can be jointly
4832 calculated:

4833 $$b_j = \sum_{i=1}^{M} \left(\frac{V_{ji}}{w_i} \sum_{k=1}^{N} U_{ik} c_k \right) \qquad u_{bj} = \sqrt{\sum_{i=1}^{M} \left(\frac{V_{ji}}{w_i} \right)^2} \qquad \mathrm{cov}(b_j, b_k) = \sum_{i=1}^{M} \frac{V_{ji} V_{ki}}{w_i^2}$$

4834 **(I.13)**

As mentioned in the main text, there is generally no need for full access to the analytical derivation of b_j , u_{bj} and covar(b_j, b_k) as most scientific programming languages (e.g., FORTRAN, IDL, MATLAB) include bundled fitting routines that provide both the fitting coefficients and their uncertainty and covariances.

J Example of partial overlap correction and its associated uncertainty

In the region of partial overlap, the lidar equation cannot be applied properly without correcting the backscattered signals for an altitude-dependent overlap factor, meant to compensate the missing fraction of the laser beam image collected on the surface of the detector. This factor, comprised between 0 and 1, is altitude-dependent and strongly dependent on the instrumental setup geometry and hardware used for a particular channel. The overlap factor can be estimated by fitting the lidar signal to a reference signal of two possible types: 1) the actual lidar signal from another channel, backscattered from a laser beam that is assumed to be fully encompassed within the receiver field-of-view of this channel, or 2) use an extinction-corrected, a priori atmospheric density profile assumed to be proportional to the lidar signal after this latter is corrected. In both cases, the signal to be corrected must be corrected for background noise and saturation first (see **chapter 4**). The uncertainty associated with this correction must be treated similarly to that of the saturation correction using a reference channel (**section 4.2.3**), i.e., use the uncertainties and covariances of the coefficients of the fitting function used to make the correction.

Starting from the end of **section 4.3**, and introducing the signals S_2 and S_3 to be the signal before and after overlap correction respectively, if i_R is the reference channel, we have:

$$S_3(i_R,k) = S_2(i_R,k) \tag{J.1}$$

The overlap correction is done by fitting the ratio:

$$\frac{S_2(i_R,k)}{S_2(i,k)} \approx f_{OVER}(k,c_1,c_2,...,c_m) \tag{J.2}$$

with the assumption that, after overlap correction, the ratio of the signals in the corrected and reference channels are proportional:

$$\frac{S_3(i,k)}{S_3(i_R,k)} = cst = c_R \tag{J.3}$$

The overlap-corrected signal S_3 is then calculated from the best fit, and from the uncorrected signal S_2:

$$S_3(i,k) = c_R f_{OVER}(k,c_1,c_2,...,c_m) S_2(i,k) \tag{J.4}$$

A new uncertainty component $u_{S3(OVER)}$ must be introduced to account for the fitting procedure associated with the saturation correction:

$$u_{S3(OVER)}(i,k) = c_R S_2(i,k) u_{fOVER}(k,c_1,c_2,...,c_m) \tag{J.5}$$

The uncertainty u_{fOVER} associated with the fitting procedure can be written in generic form:

$$u^2_{fOVER}(i,k) = \sqrt{\sum_{n=1}^{N}\left(\frac{\partial f_{OVER}(k,c_1,...,c_m)}{\partial c_n}\right)^2 u^2_{cn} + 2\sum_{m=1}^{N-1}\sum_{n=m+1}^{N}\frac{\partial f_{OVER}(k,c_1,...,c_m)}{\partial c_n}\frac{\partial f_{OVER}(k,c_1,...,c_m)}{\partial c_m}\chi_{c_n,c_m}}$$

(J.6)

The fitting coefficients' uncertainty u_{cn} and covariance $\chi_{cn,cm}$ terms are calculated and returned by the fitting routine.

The uncertainty components previously introduced propagated to the corrected signal S_3 can be written:

$$u_{S3(DET)}(i,k) = c_R f_{OVER}(k,c_1,c_2,...,c_m) u_{S2(DET)}(i,k)$$

(J.7)

$$u_{S3(SAT)}(i,k) = c_R f_{OVER}(k,c_1,c_2,...,c_m) u_{S2(SAT)}(i,k)$$

(J.8)

$$u_{S3(BKG)}(i,k) = c_R f_{OVER}(k,c_1,c_2,...,c_m) u_{S2(BKG)}(i,k)$$

(J.9)

The combined standard uncertainty for the overlap-corrected signal S_3 is:

$$u_{S3}(i,k) = \sqrt{u^2_{S3(DET)}(i,k) + u^2_{S3(SAT)}(i,k) + u^2_{S3(BKG)}(i,k) + u^2_{S3(OVER)}(i,k)}$$

(J.10)

It is discouraged to use an overlap correction if the correction factor (i.e., the coefficients of the fitting function) is found to change significantly from one correction occurrence to the next one. The correction would be equivalent to a simple adjustment of the lidar measurement to an *a priori* state, and would not reflect the actual lidar measurement in this particular channel.

Acknowledgements

This work was initiated in response to the 2010 Call for International Teams of Experts in Earth and Space Science by the International Space Science Institute (ISSI) in Bern, Switzerland. It could not have been performed without the travel and logistical support of ISSI. Part of the work described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under agreements with the National Aeronautics and Space Administration. Part of this work was carried out in support of the VALID Project. RJS would like to acknowledge the support of the Canadian National Sciences and Engineering Research Council for support of The University of Western Ontario lidar work.

List of acronyms and abbreviations

ACE-FTS	Atmospheric Chemistry Experiment-Fourier Transform Spectrometer
ACSO	WMO ad-hoc working group on Absorption Cross-sections of Ozone
AD	Analog-to-Digital
AIM	Aeronomy of Ice in the Mesosphere
AIRS	Atmospheric InfraRed Sounder
AMSU	Advanced Microwave Sounding Unit
BIPM	Bureau International des Poids et Mesures
BKG	in subscript: related to background noise extraction
BP	Bass and Paur
CIRA	COSPAR International Reference Atmosphere
CODATA	COmmittee on DATA for science and technology
COSPAR	COmmittee on SPAce Research
cov	covariance
DET	In subscript: due to detection noise
DIAL	DIfferential Absorption Lidar
DOAS	Differential Optical Absorption Spectroscopy
DMB	Daumont, Malicet and Brion
ECC	Electro-Chemical Cell
ECMWF	European Centre for Medium-range Weather Forecast
EGM	Earth Gravitational Model
ENVISAT	ENVIronmental Satellite

4935	EOS-Aura	Earth Observing System
4936	ERBS	Earth Radiation Budget Satellite
4937	ESA	European Space Agency
4938	FM	Flight Model
4939	FTIR	Fourier Transform InfraRed spectroscopy
4940	FWHM	Full Width Half Maximum
4941	GCOS	Global Climate Observing System
4942	GEOS	Goddard Earth Observing System Model
4943	GGM03	GRACE Gravity Model 03
4944	GOMOS	Global Ozone Monitoring by Occultation of Stars
4945	GOZCARDS	Global OZone Chemistry And Related trace gas Data records for the Stratosphere
4946	GPS	Global Positioning System
4947	GUM	Guide to the expression of Uncertainty in Measurement
4948	GRACE	Gravity Recovery And Climate Experiment
4949	GRS	Geodetic Reference System
4950	GRUAN	GCOS Reference Upper Air Network
4951	HALOE	Halogen Occultation Experiment
4952	IASB	Belgian Institute for Space Aeronomy
4953	ICSU	International Council for Science
4954	INTEX-B	INtercontinental chemical Transport EXperiment - phase B
4955	ISSI	International Space Science Institute
4956	JPL	Jet Propulsion Laboratory
4957	LS	Least-squares
4958	LSB	Least Significant Bit
4959	MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
4960	MLS	Microwave Limb Sounder
4961	MPI	Max Planck Institute
4962	MSIS	Mass-Spectrometer-Incoherent-Scatter
4963	NASA	US National Aeronautics and Space Administration
4964	NCEP	National Centers for Environmental Prediction
4965	NDACC	Network for the Detection of Atmospheric Composition Change
4966	NIMA	US National Imagery and Mapping Agency
4967	O3S	Stratospheric Ozone lidar (JPL)

4968	O3T	Tropospheric Ozone lidar (JPL)
4969	OA	Operational Amplifier
4970	OMI	Ozone Monitoring Instrument
4971	OSIRIS	Optical Spectrograph and InfraRed Imager System
4972	OVER	in subscript: related to overlap correction
4973	PI	Principal Investigator
4974	PC	Photon Counting
4975	SABER	Sounding of the Atmosphere using Broadband Emission Radiometry
4976	SAGE	Stratospheric Aerosol and Gas Experiments
4977	SAT	in subscript: saturation
4978	SCIAMACHY	SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY
4979	SOFIE	Solar Occultation For Ice Experiment
4980	SSU	Stratospheric Sounding Unit
4981	STP	Standard Temperature and Pressure (at 0°C or 15°C, 1013.15 hPa)
4982	SVD	Singular Value Decomposition
4983	TES	Tropospheric Emission Spectrometer
4984	TIMED	Thermosphere Ionosphere Mesosphere Energetics Dynamics
4985	TGFC	Task Group on Fundamental Constants
4986	TMP	Temperature lidar
4987	TOLNet	Tropospheric Ozone Lidar Network
4988	UARS	Upper Atmosphere Research Satellite
4989	UV	Ultraviolet
4990	VIM	International Vocabulary of basic and general terms in Metrology
4991	WGS	World Geodetic System
4992	WMO	World Meteorological Organization

4993
4994

References

Ahmad, Z., McClain, C. R., Herman, J. R., Franz, B. A., Kwiatkowska, E. J., Robinson, W. D., Bucsela, E. J., and Tzortziou, M.: Atmospheric correction for NO2 absorption in retrieving water-leaving reflectances from the SeaWiFS and MODIS measurements, Appl. Opt., 46, 6504-6512, 2007.

Anderson, S. M., and Mauersberger, K.: Laser measurements of ozone absorption cross-sections in the Chappuis band, Geophys. Res. Lett., 19, 933-936, 10.1029/92gl00780, 1992.

Argall, P. S.: Upper altitude limit for Rayleigh lidar, Ann. Geophys., 25, 19-25, 10.5194/angeo-25-19-2007, 2007.

Arshinov, Y. F., Bobrovnikov, S. M., Zuev, V. E., and Mitev, V. M.: Atmospheric-temperature measurements using a pure rotational Raman lidar, Appl. Opt., 22, 2984-2990, 1983.

Bass, A. M., and Paur, R. J.: The Ultraviolet Cross-sections of Ozone: I. The Measurements, Proc. Quadriennial Ozone Symp., Halkidiki, Greece, 606-616, 1984.

Bates, D. R.: Rayleigh-scattering by air, Planet Space Sci., 32, 785-790, 10.1016/0032-0633(84)90102-8, 1984.

Bauer, R., Rozanov, A., McLinden, C. A., Gordley, L. L., Lotz, W., Russell, J. M., Walker, K. A., Zawodny, J. M., Ladstatter-Weissenmayer, A., Bovensmann, H., and Burrows, J. P.: Validation of SCIAMACHY limb NO2 profiles using solar occultation measurements, Atmos. Meas. Tech., 5, 1059-1084, 10.5194/amt-5-1059-2012, 2012.

Beyerle, G., and McDermid, I. S.: Altitude Range Resolution of Differential Absorption Lidar Ozone Profiles, Appl. Opt., 38, 924-927, 1999.

Boersma, K. F., Jacob, D. J., Trainic, M., Rudich, Y., DeSmedt, I., Dirksen, R., and Eskes, H. J.: Validation of urban NO2 concentrations and their diurnal and seasonal variations observed from the SCIAMACHY and OMI sensors using in situ surface measurements in Israeli cities, Atmos. Chem. Phys., 9, 3867-3879, 2009.

Bogumil, K., Orphal, J., Homann, T., Voigt, S., Spietz, P., Fleischmann, O. C., Vogel, A., Hartmann, M., Kromminga, H., Bovensmann, H., Frerick, J., and Burrows, J. P.: Measurements of molecular absorption spectra with the SCIAMACHY pre-flight model: instrument characterization and reference data for atmospheric remote-sensing in the 230-2380 nm region, J. Photochem. Photobiol. A, 157, 167-184, 10.1016/s1010-6030(03)00062-5, 2003.

Bracher, A., Sinnhuber, M., Rozanov, A., and Burrows, J. P.: Using a photochemical model for the validation of NO2 satellite measurements at different solar zenith angles, Atmos. Chem. Phys., 5, 393-408, 10.5194/acp-5-393-2005, 2005.

Brion, J., Chakir, A., Charbonnier, J., Daumont, D., Parisse, C., and Malicet, J.: Absorption Spectra Measurements for the Ozone Molecule in the 350–830 nm Region, J. Atmos. Chem., 30, 291-299, 10.1023/a:1006036924364, 1998.

5033 Brohede, S., McLinden, C. A., Berthet, G., Haley, C. S., Murtagh, D., and Sioris, C. E.: A
5034 stratospheric NO2 climatology from Odin/OSIRIS limb-scatter measurements, Can. J. Phys.,
5035 85, 1253-1274, 10.1139/p07-141, 2007.

5036 Brühl, C., Lelieveld, J., Höpfner, M., and Tost, H.: Stratospheric SO2 and sulphate aerosol,
5037 model simulations and satellite observations, Atmos. Chem. Phys. Discuss., 13, 11395-
5038 11425, 10.5194/acpd-13-11395-2013, 2013.

5039 Bucholtz, A.: Rayleigh-scattering calculations for the terrestrial atmosphere, Appl. Opt., 34,
5040 2765-2773, 1995.

5041 Burkholder, J. B., and Talukdar, R. K.: Temperature dependence of the ozone absorption
5042 spectrum over the wavelength range 410 to 760 nm, Geophysical Research Letters, 21, 581-
5043 584, 10.1029/93gl02311, 1994.

5044 Burrows, J. P., Richter, A., Dehn, A., Deters, B., Himmelmann, S., and Orphal, J.: Atmospheric
5045 remote-sensing reference data from GOME - 2. Temperature-dependent absorption cross-
5046 sections of O-3 in the 231-794 nm range, J. Quant. Spectr. Rad. Trans., 61, 509-517,
5047 10.1016/s0022-4073(98)00037-5, 1999.

5048 Butz, A., Bösch, H., Camy-Peyret, C., Chipperfield, M., Dorf, M., Dufour, G., Grunow, K.,
5049 Jeseck, P., Kühl, S., Payan, S., Pepin, I., Pukite, J., Rozanov, A., von Savigny, C., Sioris, C.,
5050 Wagner, T., Weidner, F., and Pfeilsticker, K.: Inter-comparison of stratospheric O3 and NO2
5051 abundances retrieved from balloon borne direct sun observations and Envisat/SCIAMACHY
5052 limb measurements, Atmos. Chem. Phys., 6, 1293-1314, 10.5194/acp-6-1293-2006, 2006.

5053 Cao, N., Fukuchi, T., Fujii, T., Collins, R. L., Li, S., Wang, Z., and Chen, Z.: Error analysis for
5054 NO2 DIAL measurement in the troposphere, Appl. Phys. B: Lasers and Optics, 82, 141-148,
5055 10.1007/s00340-005-2050-8, 2006.

5056 Cede, A., Herman, J., Richter, A., Krotkov, N., and Burrows, J.: Measurements of nitrogen
5057 dioxide total column amounts using a Brewer double spectrophotometer in direct Sun mode,
5058 J. Geophys. Res., 111, D05304, 10.1029/2005jd006585, 2006.

5059 Chehade, W., Gorshelev, V., Serdyuchenko, A., Burrows, J. P., and Weber, M.: Revised
5060 temperature-dependent ozone absorption cross-section spectra (Bogumil et al.) measured
5061 with the SCIAMACHY satellite spectrometer, Atmos. Meas. Tech., 6, 3055-3065,
5062 doi:10.5194/amt-6-3055-2013, 2013.

5063 Ciddor, P. E.: Refractive index of air: New equations for the visible and near infrared, Appl.
5064 Opt., 35, 1566-1573, 10.1364/ao.35.001566, 1996.

5065 Committee on Extension to the Standard Atmosphere, "U.S. standard atmosphere, 1976," Tech.
5066 Rep. NOAA-S/T 76-1562, National Oceanic and Atmospheric Administration, Washington
5067 D.C. 20402, 1976.

5068 Cracknell, A. P., and Varotsos, C. A.: Satellite systems for atmospheric ozone observations, Int.
5069 J. Remote Sens., 35, 5566-5597, 10.1080/01431161.2014.945013, 2014.

5070 Daumont, D., Brion, J., Charbonnier, J., and Malicet, J.: Ozone UV Spectroscopy I: Absorption
5071 Cross-Sections at Room Temperature, J. Atmos Chem., 15, 145-155, 10.1007/bf00053756,
5072 1992.

5073 Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U.,
5074 Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L.,
5075 Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L.,
5076 Healy, S. B., Hersbach, H., Holm, E. V., Isaksen, L., Kallberg, P., Kohler, M., Matricardi,
5077 M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay,
5078 P., Tavolato, C., Thepaut, J. N., and Vitart, F.: The ERA-Interim reanalysis: configuration
5079 and performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553-597,
5080 10.1002/qj.828, 2011.

5081 Dirksen, R. J., Boersma, K. F., Eskes, H. J., Ionov, D. V., Bucsela, E. J., Levelt, P. F., and
5082 Kelder, H. M.: Evaluation of stratospheric NO2 retrieved from the Ozone Monitoring
5083 Instrument: Intercomparison, diurnal cycle, and trending, J. Geophys. Res., 116, D08305,
5084 10.1029/2010jd014943, 2011.

5085 Donovan, D. P., Whiteway, J. A., and Carswell, A. I.: Correction for nonlinear photon-counting
5086 effects in lidar systems, Appl. Opt., 32, 6742-6753, 1993.

5087 Dorf, M., Bösch, H., Butz, A., Camy-Peyret, C., Chipperfield, M. P., Gerilowski, K., Grunow,
5088 K., Gurlit, W., Harder, H., Kühl, S., Kritten, L., Lindner, A., Payan, S., Rozanov, A.,
5089 Savigny, C. v., Simmes, B., Sioris, C., Weidner, F., and Pfeilsticker, K.: Validation of
5090 SCIAMACHY Level-1 and Level-2 Products by Balloon-Borne Differential Optical
5091 Absorption Spectroscopy (DOAS), Proc. of the First Atmospheric Science Conference, 2006.

5092 Eberhard, W. L.: Correct equations and common approximations for calculating Rayleigh scatter
5093 in pure gases and mixtures and evaluation of differences, Appl. Opt., 49, 1116-1130, 2010.

5094 Edlen, B.: The dispersion of standard air, J. Opt. Soc. Am., 43, 339-344,
5095 10.1364/josa.43.000339, 1953.

5096 Edlén, B.: The Refractive Index of Air, Metrologia, 2, 71, 1966.

5097 Fally, S., Vandaele, A. C., Carleer, M., Hermans, C., Jenouvrier, A., Merienne, M. F., Coquart,
5098 B., and Colin, R.: Fourier transform spectroscopy of the O-2 Herzberg bands. III. Absorption
5099 cross-sections of the collision-induced bands and of the Herzberg continuum, J. Mol.
5100 Spectrosc., 204, 10-20, 10.1006/jmsp.2000.8204, 2000.

5101 Fasso, A., Ignaccolo, R., Madonna, F., Demoz, B. B., and Franco-Villoria, M.: Statistical
5102 modelling of collocation uncertainty in atmospheric thermodynamic profiles, Atmos. Meas.
5103 Tech., 7, 1803-1816, 10.5194/amt-7-1803-2014, 2014

5104 Fröhlich, C., and Shaw, G. E.: New determination of Rayleigh scattering in the terrestrial
5105 atmosphere, Appl. Opt., 19, 1773-1775, 1980.

5106 Garcia, R. R., Marsh, D. R., Kinnison, D. E., Boville, B. A., and Sassi, F.: Simulation of secular
5107 trends in the middle atmosphere, 1950-2003, J. Geophys. Res.-Atmos., 112, D09301,
5108 10.1029/2006jd007485, 2007.

5109 Godin-Beekmann, S., and Nair, P. J.: Sensitivity of stratospheric ozone lidar measurements to a
5110 change in ozone absorption cross-sections, J. Quant. Spectr. Rad. Trans., 113, 1317-1321,
5111 10.1016/j.jqsrt.2012.03.002, 2012.

5112 Godin-Beekmann, S., Porteneuve, J., and Garnier, A.: Systematic DIAL lidar monitoring of the
5113 stratospheric ozone vertical distribution at Observatoire de Haute-Provence (43.92 degrees N,
5114 5.71 degrees E), J. Environ. Monit., 5, 57-67, 10.1039/b205880d, 2003.

5115 Godin, S., Carswell, A. I., Donovan, D. P., Claude, H., Steinbrecht, W., McDermid, I. S.,
5116 McGee, T. J., Gross, M. R., Nakane, H., Swart, D. P. J., Bergwerff, H. B., Uchino, O., von
5117 der Gathen, P., and Neuber, R.: Ozone Differential Absorption Lidar Algorithm
5118 Intercomparison, Appl. Opt., 38, 6225-6236, 1999.

5119 Gorshelev, V., Serdyuchenko, A., Weber, M., Chehade, W., and Burrows, J. P.: High spectral
5120 resolution ozone absorption cross-sections - Part 1: Measurements, data analysis and
5121 comparison with previous measurements around 293 K, Atmos. Meas. Tech., 7, 609-624,
5122 10.5194/amt-7-609-2014, 2014.

5123 Gross, M. R., McGee, T. J., Ferrare, R. A., Singh, U. N., and Kimvilakani, P.: Temperature
5124 measurements made with a combined Rayleigh-Mie and Raman lidar, Appl. Opt., 36, 5987-
5125 5995, 1997.

5126 Hanson, D., and Mauersberger, K.: Precision ozone vapor-pressure measurements, J. Chem.
5127 Phys., 83, 326-328, 10.1063/1.449773, 1985.

5128 Hauchecorne, A., and Chanin, M. L.: Density and temperature profiles obtained by lidar between
5129 35-km and 70-km, Geophys. Res. Lett., 7, 565-568, 1980.

5130 He, H., Loughner, C. P., Stehr, J. W., Arkinson, H. L., Brent, L. C., Follette-Cook, M. B.,
5131 Tzortziou, M. A., Pickering, K. E., Thompson, A. M., Martins, D. K., Diskin, G. S.,
5132 Anderson, B. E., Crawford, J. H., Weinheimer, A. J., Lee, P., Hains, J. C., and Dickerson, R.
5133 R.: An elevated reservoir of air pollutants over the Mid-Atlantic States during the 2011
5134 DISCOVER-AQ campaign: Airborne measurements and numerical simulations, Atmos.
5135 Env., 85, 18-30, 10.1016/j.atmosenv.2013.11.039, 2014.

5136 Hearn, A. G.: Absorption of ozone in ultra-violet and visible regions of spectrum, Proc. Phys.
5137 Soc., 78, 932-&, 10.1088/0370-1328/78/5/340, 1961.

5138 Hedin, A. E.: Extension of the MSIS thermosphere model into the middle and lower atmosphere,
5139 J. Geophys. Res., 96, 1159-1172, 10.1029/90ja02125, 1991.

5140 Hendrick, F., Granville, J., Lambert, J.-C., and Roozendael, M. V.: Validation of Schiamachy
5141 OL 3.0 NO2 Profiles and Columns using Ground-based DOAS Profiling, Proc. Third
5142 Workshop on the Atmospheric Chemistry and Validation of ENVISAT (ACVE-3), 2003.

5143 Hinkley, E. D.: Laser monitoring of the atmosphere, Topics in applied physics, 14, Springer-
5144 Verlag, 380 pp., 1976.

5145 Hopfner, M., Glatthor, N., Grabowski, U., Kellmann, S., Kiefer, M., Linden, A., Orphal, J.,
5146 Stiller, G., von Clarmann, T., Funke, B., and Boone, C. D.: Sulfur dioxide (SO2) as observed
5147 by MIPAS/Envisat: temporal development and spatial distribution at 15-45 km altitude,
5148 Atmos. Chem. Phys., 13, 10405-10423, 10.5194/acp-13-10405-2013, 2013.

5149 Hoyt, D. V.: Redetermination of rayleigh optical depth and its application to selected solar-
5150 radiation problems, J. Appl. Meteorol., 16, 432-436, 10.1175/1520-
5151 0450(1977)016<0432:arotro>2.0.co;2, 1977.

5152 Hurst, D. F., Hall, E. G., Jordan, A. F., Miloshevich, L. M., Whiteman, D. N., Leblanc, T.,
5153 Walsh, D., Vomel, H., and Oltmans, S. J.: Comparisons of temperature, pressure and
5154 humidity measurements by balloon-borne radiosondes and frost point hygrometers during
5155 MOHAVE-2009, Atmos. Meas. Tech., 4, 2777-2793, 10.5194/amt-4-2777-2011, 2011.

5156 Immler, F. J., Dykema, J., Gardiner, T., Whiteman, D. N., Thorne, P. W., and Vömel, H.:
5157 Reference Quality Upper-Air Measurements: guidance for developing GRUAN data
5158 products, Atmos. Meas. Tech., 3, 1217-1231, 10.5194/amt-3-1217-2010, 2010.

5159 JCGM: International vocabulary of basic and general terms in metrology (VIM), Tech. Rep.
5160 JCGM 200:2008, International Bureau of Weights and Measures (BIPM), 2008.

5161 JCGM: Evaluation of measurement data – Guide to the expression of uncertainty in
5162 measurement (GUM), Tech. Rep. JCGM 100: 2008, International Bureau of Weights and
5163 Measures (BIPM), 2008.

5164 JCGM: Evaluation of measurement data – An introduction to the "Guide to the expression of
5165 uncertainty in measurement" and related documents, Tech. Rep. JCGM 104: 2009,
5166 International Bureau of Weights and Measures (BIPM), 2009.

5167 Jenouvrier, A., Merienne, M. F., Coquart, B., Carleer, M., Fally, S., Vandaele, A. C., Hermans,
5168 C., and Colin, R.: Fourier transform spectroscopy of the O-2 Herzberg bands - I. Rotational
5169 analysis, J. Mol. Spectrosc., 198, 136-162, 10.1006/jmsp.1999.7950, 1999.

5170 Jones, A., Walker, K. A., Jin, J. J., Taylor, J. R., Boone, C. D., Bernath, P. F., Brohede, S.,
5171 Manney, G. L., McLeod, S., Hughes, R., and Daffer, W. H.: Technical Note: A trace gas
5172 climatology derived from the Atmospheric Chemistry Experiment Fourier Transform
5173 Spectrometer (ACE-FTS) data set, Atmos. Chem. Phys., 12, 5207-5220, 10.5194/acp-12-
5174 5207-2012, 2012.

5175 Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha,
5176 S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J.,
5177 Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.:
5178 The NCEP/NCAR 40-year reanalysis project, Bull. Amer. Meteorol. Soc., 77, 437-471,
5179 10.1175/1520-0477(1996)077<0437:tnyrp>2.0.co;2, 1996.

5180 Keckhut, P., Hauchecorne, A., and Chanin, M. L.: A critical-review of the database acquired for
5181 the long-term surveillance of the middle atmosphere by the French Rayleigh lidars, Journal
5182 of Atmospheric and Oceanic Technology, 10, 850-867, 10.1175/1520-
5183 0426(1993)010<0850:acrotd>2.0.co;2, 1993.

5184 Keckhut, P., Randel, W. J., Claud, C., Leblanc, T., Steinbrecht, W., Funatsu, B. M., Bencherif,
5185 H., McDermid, I. S., Hauchecorne, A., Long, C., Lin, R., and Baumgarten, G.: An evaluation
5186 of uncertainties in monitoring middle atmosphere temperatures with the ground-based lidar
5187 network in support of space observations, J. Atmos. Solar-Terr. Phys., 73, 627-642,
5188 10.1016/j.jastp.2011.01.003, 2011.

5189 King, L. V.: On the complex anisotropic molecule in relation to the dispersion and scattering of
5190 light, Proc. R. Soc. Lond. A, 104, 333-357, 10.1098/rspa.1923.0113, 1923.

5191 Kuhl, S., Pukite, J., Deutschmann, T., Platt, U., and Wagner, T.: SCIAMACHY limb
5192 measurements of NO2, BrO and OClO. Retrieval of vertical profiles: Algorithm, first results,

5193 sensitivity and comparison studies, Adv. Space Res., 42, 1747-1764,
5194 10.1016/j.asr.2007.10.022, 2008.

5195 Leblanc, T., McDermid, I. S., Hauchecorne, A., and Keckhut, P.: Evaluation of optimization of
5196 lidar temperature analysis algorithms using simulated data, J. Geophys. Res., 103, 6177-
5197 6187, 1998.

5198 Lemoine, F. C., Kenyon, S. C., Factor, J. K., Trimmer, R. G., Pavlis, N. K., Chinn, D. S., Cox, C.
5199 M., Klosko, S. M., Luthcke, S. B., Torrence, M. H., Wang, Y. M., Williamson, R. G., Pavlis,
5200 E. C., Rapp, R. H., and Olson, T. R.: The Development of the Joint NASA GSFC and the
5201 National Imagery and Mapping Agency (NIMA) Geopotential Model EGM96, Tech. Rep.,
5202 1998.

5203 Liu, G., Liu, J., Tarasick, D. W., Fioletov, V. E., Jin, J. J., Moeini, O., Liu, X., Sioris, C. E., and
5204 Osman, M.: A global tropospheric ozone climatology from trajectory-mapped ozone
5205 soundings, Atmos. Chem. Phys., 13, 10659-10675, 10.5194/acp-13-10659-2013, 2013.

5206 Malicet, J., Daumont, D., Charbonnier, J., Parisse, C., Chakir, A., and Brion, J.: Ozone uv
5207 spectroscopy .2. Absorption cross-sections and temperature-dependence, J. Atmos Chem.,
5208 21, 263-273, 10.1007/bf00696758, 1995.

5209 Mauersberger, K., Barnes, J., Hanson, D., and Morton, J.: Measurement of the ozone absorption
5210 cross-section at the 253.7 nm mercury line, Geophys. Res. Lett., 13, 671-673,
5211 10.1029/GL013i007p00671, 1986.

5212 McGee, T. J., Gross, M., Ferrare, R., Heaps, W., and Singh, U.: Raman dial measurements of
5213 stratospheric ozone in the presence of volcanic aerosols, Geophys. Res. Lett., 20, 955-958,
5214 10.1029/93gl00751, 1993.

5215 Megie, G., and Menzies, R. T.: Complementarity of UV and IR differential absorption lidar for
5216 global measurements of atmospheric species, Appl. Opt., 19, 1173-1183, 1980

5217 McGee, T. J., Whiteman, D., Ferrare, R., Butler, J. J., and Burris, J. F.: STROZ LITE -
5218 STRatospheric OZone Lidar Trailer Experiment, Opt. Engin., 30, 31-39, 1991

5219 McLinden, C. A., Fioletov, V., Boersma, K. F., Kharol, S. K., Krotkov, N., Lamsal, L., Makar,
5220 P. A., Martin, R. V., Veefkind, J. P., and Yang, K.: Improved satellite retrievals of NO2 and
5221 SO2 over the Canadian oil sands and comparisons with surface measurements, Atmos.
5222 Chem. Phys., 14, 3637-3656, 10.5194/acp-14-3637-2014, 2014.

5223 Megie, G., Allain, J. Y., Chanin, M. L., and Blamont, J. E.: Vertical Profile of Stratospheric
5224 Ozone by Lidar Sounding from Ground, Nature, 270, 329-331, 1977.

5225 Meng, Z. Y., Ding, G. A., Xu, X. B., Xu, X. D., Yu, H. Q., and Wang, S. F.: Vertical
5226 distributions of SO2 and NO2 in the lower atmosphere in Beijing urban areas, China, Sci.
5227 Total Env., 390, 456-465, 10.1016/j.scitotenv.2007.10.012, 2008.

5228 Merienne, M. F., Jenouvrier, A., Coquart, B., Carleer, M., Fally, S., Colin, R., Vandaele, A. C.,
5229 and Hermans, C.: Fourier transform spectroscopy of the O-2 Herzberg bands - II. Band
5230 oscillator strengths and transition moments, J. Mol. Spectrosc., 202, 171-193,
5231 10.1006/jmsp.2000.8126, 2000.

5232 Miyazaki, K., Eskes, H. J., and Sudo, K.: Global NOx emission estimates derived from an
5233 assimilation of OMI tropospheric NO2 columns, Atmos. Chem. Phys., 12, 2263-2288,
5234 10.5194/acp-12-2263-2012, 2012a.

5235 Miyazaki, K., Eskes, H. J., Sudo, K., Takigawa, M., van Weele, M., and Boersma, K. F.:
5236 Simultaneous assimilation of satellite NO2, O-3, CO, and HNO3 data for the analysis of
5237 tropospheric chemical composition and emissions, Atmos. Chem. Phys., 12, 9545-9579,
5238 10.5194/acp-12-9545-2012, 2012b.

5239 Mohr, P. J., Taylor, B. N., and Newell, D. B.: CODATA recommended values of the
5240 fundamental physical constants: 2006, Rev. Mod. Phys., 80, 633-730,
5241 10.1103/RevModPhys.80.633, 2008.

5242 Molina, L. T., and Molina, M. J.: Absolute absorption cross-sections of ozone in the 185-nm to
5243 350-nm wavelength range, J. Geophys. Res., 91, 14501-14508, 10.1029/JD091iD13p14501,
5244 1986.

5245 Muller, J. W.: Dead-time problems, Nucl. Instr. and Meth., 112, 47-57, 10.1016/0029-
5246 554x(73)90773-8, 1973.

5247 Muller, J. W.: Some formulas for a dead-time-Distorted Poisson process, Nucl. Instr. and Meth.,
5248 117, 401-404, 1974.

5249 Newsom, R. K., Turner, D. D., Mielke, B., Clayton, M., Ferrare, R., and Sivaraman, C.:
5250 Simultaneous analog and photon counting detection for Raman lidar, Appl. Opt., 48, 3903-
5251 3914, 2009.

5252 Nicolet, M.: On the molecular-scattering in the terrestrial atmosphere - an empirical-formula for
5253 its calculation in the homosphere, Planet. Space Sci., 32, 1467-1468, 10.1016/0032-
5254 0633(84)90089-8, 1984.

5255 NIMA: Department of Defense World Geodetic System 1984, Tech. Rep., 1984.

5256 O'Keeffe, P., Ridley, T., Lawley, K. P., and Donovan, R. J.: Re-analysis of the ultraviolet
5257 absorption spectrum of ozone, J. Chem. Phys., 115, 9311-9319, 10.1063/1.1412254, 2001.

5258 Omote, K.: Dead-time effects in photon-counting distributions, Nucl. Instrum. Methods, 293,
5259 582-588, 10.1016/0168-9002(90)90327-3, 1990.

5260 Orphal, J.: A Critical Review of the Absorption Cross-Sections of O3 and NO2 in the 240–790
5261 nm Region, ESA Technical Report, 1-126, 2002.

5262 Orphal, J.: A critical review of the absorption cross-sections of O3 and NO2 in the ultraviolet
5263 and visible, J. Photochem. Photobiol., 157, 185-209, 10.1016/s1010-6030(03)00061-3, 2003.

5264 Papayannis, A., Ancellet, G., Pelon, J., and Megie, G.: Multiwavelength lidar for ozone
5265 measurements in the troposphere and the lower stratosphere, Appl. Opt., 29, 467-476, 1990

5266

5267 Pavlis, N. K., Holmes, S. A., Kenyon, S. C., and Factor, J. K.: The development and evaluation
5268 of the Earth Gravitational Model 2008 (EGM2008), J. Geophys. Res., 117, B04406,
5269 10.1029/2011jb008916, 2012.

5270 Peck, E. R., and Reeder, K.: Dispersion of air, J. Opt. Soc. Am., 62, 958,
5271 10.1364/josa.62.000958, 1972.

5272 Pelon, J., and Megie, G.: Ozone monitoring in the troposphere and lower stratosphere -
5273 evaluation and operation of a ground-based lidar station, J. Geophys. Res., 87, 4947-4955,
5274 10.1029/JC087iC07p04947, 1982

5275 Penndorf, R.: Tables of the refractive index for standard air and the rayleigh scattering
5276 coefficient for the spectral region between 0.2 and 20.0-mu and their application to
5277 atmospheric optics, J. Opt. Soc. Am., 47, 176-182, 10.1364/josa.47.000176, 1957.

5278 Picone, J. M., Hedin, A. E., Drob, D. P., and Aikin, A. C.: NRLMSISE-00 empirical model of
5279 the atmosphere: Statistical comparisons and scientific issues, J. Geophys. Res-Space Phys.,
5280 107, 1468,10.1029/2002ja009430, 2002.

5281 Press, W. H.; Flannery, B. P., Teukolsky, S. A., and Vetterling W. T.: Numerical Recipes: The
5282 Art of Scientific Computing (1st ed.), New York, Cambridge University Press. ISBN 978-0-
5283 521-88068-8, 1986.

5284 Qu, Z. W., Zhu, H., Grebenshchikov, S. Y., Schinke, R., and Farantos, S. C.: The Huggins band
5285 of ozone: A theoretical analysis, J. Chem. Phys., 121, 11731-11745, 10.1063/1.1814098,
5286 2004.

5287 Reale, A., Tilley, F., Ferguson, M., and Allegrino, A.: NOAA operational sounding products for
5288 advanced TOVS, Int. J. Remote Sensing, 29, 4615-4651, 2008.

5289 Remsberg, E. E., Marshall, B. T., Garcia-Comas, M., Krueger, D., Lingenfelser, G. S., Martin-
5290 Torres, J., Mlynczak, M. G., Russell, J. M., Smith, A. K., Zhao, Y., Brown, C., Gordley, L.
5291 L., Lopez-Gonzalez, M. J., Lopez-Puertas, M., She, C. Y., Taylor, M. J., and Thompson, R.
5292 E.: Assessment of the quality of the Version 1.07 temperature-versus-pressure profiles of the
5293 middle atmosphere from TIMED/SABER, J. Geophys. Res., 113, D17101.,
5294 10.1029/2008jd010013, 2008.

5295 Rozanov, A., Bovensmann, H., Bracher, A., Hrechanyy, S., Rozanov, V., Sinnhuber, M., Stroh,
5296 F., and Burrows, J. P.: NO2 and BrO vertical profile retrieval from SCIAMACHY limb
5297 measurements: Sensitivity studies, in: Atmospheric Remote Sensing: Earth's Surface,
5298 Troposphere, Stratosphere and Mesosphere - I, edited by: Burrows, J. P., and Eichmann, K.
5299 U., Adv. Space Res., 5, Elsevier Science Ltd, Oxford, 846-854, 2005.

5300 Savitzky, A., and Golay, M. J. E.: Smoothing and differentiation of data by simplified least
5301 squares procedures, Anal. Chem., 36, 1627-&, 1964.

5302 Schwartz, M. J., Lambert, A., Manney, G. L., Read, W. G., Livesey, N. J., Froidevaux, L., Ao,
5303 C. O., Bernath, P. F., Boone, C. D., Cofield, R. E., Daffer, W. H., Drouin, B. J., Fetzer, E. J.,
5304 Fuller, R. A., Jarnot, R. F., Jiang, J. H., Jiang, Y. B., Knosp, B. W., Kruger, K., Li, J. L. F.,
5305 Mlynczak, M. G., Pawson, S., Russell, J. M., Santee, M. L., Snyder, W. V., Stek, P. C.,
5306 Thurstans, R. P., Tompkins, A. M., Wagner, P. A., Walker, K. A., Waters, J. W., and Wu, D.
5307 L.: Validation of the aura microwave limb sounder temperature and geopotential height
5308 measurements, J. Geophys. Res., 113, D15s11,10.1029/2007jd008783, 2008.

5309 Serdyuchenko, A., Gorshelev, V., Weber, M., Chehade, W., and Burrows, J. P.: High spectral
5310 resolution ozone absorption cross-sections - Part 2: Temperature dependence, Atmos. Meas.
5311 Tech., 7, 625-636, 10.5194/amt-7-625-2014, 2014.

5312 Sheese, P. E., Strong, K., Llewellyn, E. J., Gattinger, R. L., Russell, J. M., Boone, C. D., Hervig,
5313 M. E., Sica, R. J., and Bandoro, J.: Assessment of the quality of OSIRIS mesospheric

5314 temperatures using satellite and ground-based measurements, Atmos. Meas. Tech., 5, 2993-
5315 3006, 10.5194/amt-5-2993-2012, 2012.

5316 Sica, R. J. and A. Heafelee, Retrieval of temperature from a multiple channel Rayleigh-Scatter
5317 lidar using an Optimal Estimation Method, in press, Appl. Optics, 2014.

5318 Sica, R. J., Izawa, M. R. M., Walker, K. A., Boone, C., Petelina, S. V., Argall, P. S., Bernath, P.,
5319 Burns, G. B., Catoire, V., Collins, R. L., Daffer, W. H., De Clercq, C., Fan, Z. Y., Firanski,
5320 B. J., French, W. J. R., Gerard, P., Gerding, M., Granville, J., Innis, J. L., Keckhut, P.,
5321 Kerzenmacher, T., Klekociuk, A. R., Kyro, E., Lambert, J. C., Llewellyn, E. J., Manney, G.
5322 L., McDermid, I. S., Mizutani, K., Murayama, Y., Piccolo, C., Raspollini, P., Ridolfi, M.,
5323 Robert, C., Steinbrecht, W., Strawbridge, K. B., Strong, K., Stuebi, R., and Thurairajah, B.:
5324 Validation of the Atmospheric Chemistry Experiment (ACE) version 2.2 temperature using
5325 ground-based and space-borne measurements, Atmos. Chem. Phys., 8, 35-62, 2008.

5326 Sica, R. J., Zylawy, Z. A., and Argall, P. S.: Ozone Corrections for Rayleigh-Scatter
5327 Temperature Determinations in the Middle Atmosphere, J. Atmos. Ocean. Tech., 18, 1223-
5328 1228, doi:10.1175/1520-0426(2001)018<1223:OCFRST>2.0.CO;2, 2001.

5329 Sioris, C. E., McLinden, C. A., Martin, R. V., Sauvage, B., Haley, C. S., Lloyd, N. D.,
5330 Llewellyn, E. J., Bernath, P. F., Boone, C. D., Brohede, S., and McElroy, C. T.: Vertical
5331 profiles of lightning-produced NO2 enhancements in the upper troposphere observed by
5332 OSIRIS, Atmos. Chem. Phys., 7, 4281-4294, 10.5194/acp-7-4281-2007, 2007.

5333 Smit, H. G. J, Oltmans, S., Deshler, T., Tarasick, D., Johnson, B., Schmidlin, F., Stuebi, R., and
5334 Davies, J.: SI2N/O3SDQA, O3S-DQA Activity: Guide Lines for Homogenization of Ozone
5335 Sonde Data (Version 2.0: 12.October 2012), 2012.

5336 Smit, H. G. J., Straeter, W., Johnson, B. J., Oltmans, S. J., Davies, J., Tarasick, D. W., Hoegger,
5337 B., Stubi, R., Schmidlin, F. J., Northam, T., Thompson, A. M., Witte, J. C., Boyd, I., and
5338 Posny, F.: Assessment of the performance of ECC-ozonesondes under quasi-flight conditions
5339 in the environmental simulation chamber: Insights from the Juelich Ozone Sonde
5340 Intercomparison Experiment (JOSIE), J. Geophys. Res., 112, D19306,
5341 10.1029/2006jd007308, 2007.

5342 Sofieva, V. F., Dalaudier, F., Kivi, R., and Kyro, E.: On the variability of temperature profiles in
5343 the stratosphere: Implications for validation, Geophys. Res. Lett., 35, L23808,
5344 10.1029/2008gl035539, 2008.

5345 Stauffer, R. M., Morris, G. A., Thompson, A. M., Joseph, E., Coetzee, G. J. R., and Nalli, N. R.:
5346 Propagation of radiosonde pressure sensor errors to ozonesonde measurements, Atmos.
5347 Meas. Tech., 7, 65-79, 10.5194/amt-7-65-2014, 2014.

5348 Stevens, M. H., Deaver, L. E., Hervig, M. E., Russell, J. M., Siskind, D. E., Sheese, P. E.,
5349 Llewellyn, E. J., Gattinger, R. L., Hoffner, J., and Marshall, B. T.: Validation of upper
5350 mesospheric and lower thermospheric temperatures measured by the Solar Occultation for
5351 Ice Experiment, J. Geophys. Res., 117, D16304, 10.1029/2012jd017689, 2012.

5352 Strauch, R. G., Derr, V. E., and Cupp, R. E.: Atmospheric temperature measurement using raman
5353 backscatter, Appl. Opt., 10, 2665-&, 10.1364/ao.10.002665, 1971.

5354 Strutt, J. W. (Lord Rayleigh): XXXIV. On the transmission of light through an atmosphere
5355 containing small particles in suspension, and on the origin of the blue of the sky, Philos.
5356 Mag., 47, 375-384, 10.1080/14786449908621276, 1899.

5357 Sussmann, R., Stremme, W., Burrows, J. P., Richter, A., Seiler, W., and Rettinger, M.:
5358 Stratospheric and tropospheric NO2 variability on the diurnal and annual scale: a combined
5359 retrieval from ENVISAT/SCIAMACHY and solar FTIR at the Permanent Ground-Truthing
5360 Facility Zugspitze/Garmisch, Atmos. Chem. Phys., 5, 2657-2677, 10.5194/acp-5-2657-2005,
5361 2005.

5362 Tapley, B., Ries, J., Bettadpur, S., Chambers, D., Cheng, M., Condi, F., Gunter, B., Kang, Z.,
5363 Nagel, P., Pastor, R., Pekker, T., Poole, S., and Wang, F.: GGM02 - An improved Earth
5364 gravity field model from GRACE, J. Geodesy, 79, 467-478, 10.1007/s00190-005-0480-z,
5365 2005.

5366 Thompson, A. M., Witte, J. C., Smit, H. G. J., Oltmans, S. J., Johnson, B. J., Kirchhoff, V., and
5367 Schmidlin, F. J.: Southern Hemisphere Additional Ozonesondes (SHADOZ) 1998-2004
5368 tropical ozone climatology: 3. Instrumentation, station-to-station variability, and evaluation
5369 with simulated flight profiles, J. Geophys. Res., 112, D03304, 10.1029/2005jd007042, 2007.

5370 Uchino, O., and Tabata, I.: Mobile lidar for simultaneous measurements of ozone, aerosols, and
5371 temperature in the stratosphere, Appl. Opt., 30, 2005-2012, 1991

5372 Uppala, S. M., Kallberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V. D., Fiorino, M.,
5373 Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S.,
5374 Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., Van
5375 De Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M.,
5376 Fisher, M., Fuentes, M., Hagemann, S., Holm, E., Hoskins, B. J., Isaksen, L., Janssen, P.,
5377 Jenne, R., McNally, A. P., Mahfouf, J. F., Morcrette, J. J., Rayner, N. A., Saunders, R. W.,
5378 Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.:
5379 The ERA-40 re-analysis, Q. J. Roy. Meteor. Soc., 131, 2961-3012, 10.1256/qj.04.176, 2005.

5380 Vandaele, A. C., Hermans, C., Simon, P. C., Carleer, M., Colin, R., Fally, S., Merienne, M. F.,
5381 Jenouvrier, A., and Coquart, B.: Measurements of the NO2 absorption cross-section from 42
5382 000 cm(-1) to 10 000 cm(-1) (238-1000 nm) at 220 K and 294 K, J. Quant. Spectr. Rad.
5383 Trans., 59, 171-184, 10.1016/s0022-4073(97)00168-4, 1998.

5384 Vandaele, A. C., Hermans, C., and Fally, S.: Fourier transform measurements of SO2 absorption
5385 cross-sections: II. Temperature dependence in the 29 000-44 000 cm(-1) (227-345 nm)
5386 region, J. Quant. Spectr. Rad. Trans., 110, 2115-2126, 10.1016/j.jqsrt.2009.05.006, 2009.

5387 Voigt, S., Orphal, J., Bogumil, K., and Burrows, J. P.: The temperature dependence (203-293 K)
5388 of the absorption cross-sections of O-3 in the 230-850 nm region measured by Fourier-
5389 transform spectroscopy, J. Photochemistry and Photobiology a-Chemistry, 143, 1-9,
5390 10.1016/s1010-6030(01)00480-4, 2001.

5391 Voigt, S., Orphal, J., and Burrows, J. P.: The temperature and pressure dependence of the
5392 absorption cross-sections of NO2 in the 250-800 nm region measured by Fourier-transform
5393 spectroscopy, J. Photochem. Photobiol., 149, 1-7, 10.1016/s1010-6030(01)00650-5, 2002.

5394 Volten, H., Brinksma, E. J., Berkhout, A. J. C., Hains, J., Bergwerff, J. B., Van der Hoff, G. R.,
5395 Apituley, A., Dirksen, R. J., Calabretta-Jongen, S., and Swart, D. P. J.: NO2 lidar profile

measurements for satellite interpretation and validation, J. Geophys. Res., 114, D24301, 10.1029/2009jd012441, 2009.

Wang, W. H., and Zou, C. Z.: AMSU-A-Only Atmospheric Temperature Data Records from the Lower Troposphere to the Top of the Stratosphere, J. Atmos. Ocean Tech., 31, 808-825, 10.1175/jtech-d-13-00134.1, 2014.

Weitkamp, C.: Lidar: Range-Resolved Optical Remote Sensing of the Atmosphere, Springer Series in Optical Sciences, 102, Springer, 460 pp., 2005.

Welton, E. J., and Campbell, J. R.: Micropulse lidar signals: Uncertainty analysis, J. Atmos. Ocean Tech., 19, 2089-2094, 10.1175/1520-0426(2002)019<2089:mlsua>2.0.co;2, 2002.

Wetzel, G., Bracher, A., Funke, B., Goutail, F., Hendrick, F., Lambert, J. C., Mikuteit, S., Piccolo, C., Pirre, M., Bazureau, A., Belotti, C., Blumenstock, T., De Mazière, M., Fischer, H., Huret, N., Ionov, D., López-Puertas, M., Maucher, G., Oelhaf, H., Pommereau, J. P., Ruhnke, R., Sinnhuber, M., Stiller, G., Van Roozendael, M., and Zhang, G.: Validation of MIPAS-ENVISAT NO2 operational data, Atmos. Chem. Phys., 7, 3261-3284, 10.5194/acp-7-3261-2007, 2007.

Whiteman, D. N., Demoz, B., Rush, K., Schwemmer, G., Gentry, B., Di Girolamo, P., Comer, J., Veselovskii, I., Evans, K., Melfi, S. H., Wang, Z., Cadirola, M., Mielke, B., Venable, D., and Van Hove, T.: Raman Lidar Measurements during the International H2O Project. Part I: Instrumentation and Analysis Techniques, J. Atmos. Ocean. Tech., 23, 157-169, doi:10.1175/JTECH1838.1, 2006.

Young, A. T.: Revised depolarization corrections for atmospheric extinction, Appl. Opt., 19, 3427-3428, 10.1364/ao.19.003427, 1980.

Young, A. T.: Rayleigh-scattering, Appl. Opt., 20, 533-535, 10.1364/ao.20.000533, 1981.