

Article

Coupled Stratospheric Chemistry–Meteorology Data Assimilation. Part II: Weak and Strong Coupling

Richard Ménard ^{1,*} , Pierre Gauthier ² , Yves Rochon ¹ , Alain Robichaud ¹,
Jean de Grandpré ¹, Yan Yang ¹, Cécilien Charrette ³ and Simon Chabrilat ⁴ 

¹ Air Quality Research Division, Environment and Climate Change Canada, 2121 Transcanada Highway, Dorval, QC H9P 1J3, Canada; yves.rochon@canada.ca (Y.R.); alain.robichaud@canada.ca (A.R.); jean.degrandpre@canada.ca (J.d.G.); yan.yang@canada.ca (Y.Y.)

² Département des Sciences de la Terre et de L’atmosphère, Université du Québec à Montréal, Montréal, QC H2L 2C4, Canada; gauthier.pierre@uqam.ca

³ Meteorological Research Division, Environment and Climate Change Canada; Dorval, QC H9P 1J3, Canada; cecilien.charrette@ec.gc.ca

⁴ Royal Belgian Institute for Space Aeronomy, BIRA-IASB, 1180 Brussels, Belgium; simon.chabrilat@aeronomie.be

* Correspondence: richard.menard@canada.ca; Tel.: +1-514-421-4613

Received: 19 October 2019; Accepted: 5 December 2019; Published: 9 December 2019



Abstract: We examine data assimilation coupling between meteorology and chemistry in the stratosphere from both weak and strong coupling strategies. The study was performed with the Canadian operational weather prediction Global Environmental Multiscale (GEM) model coupled online with the photochemical stratospheric chemistry model developed at the Belgian Institute for Space Aeronomy, described in Part I. Here, the Canadian Meteorological Centre’s operational variational assimilation system was extended to include errors of chemical variables and cross-covariances between meteorological and chemical variables in a 3D-Var configuration, and we added the adjoint of tracer advection in the 4D-Var configuration. Our results show that the assimilation of limb sounding observations from the MIPAS instrument on board Envisat can be used to anchor the AMSU-A radiance bias correction scheme. Additionally, the added value of limb sounding temperature observations on meteorology and transport is shown to be significant. Weak coupling data assimilation with ozone–radiation interaction is shown to give comparable results on meteorology whether a simplified linearized or comprehensive ozone chemistry scheme is used. Strong coupling data assimilation, using static error cross-covariances between ozone and temperature in a 3D-Var context, produced inconclusive results with the approximations we used. We have also conducted the assimilation of long-lived species observations using 4D-Var to infer winds. Our results showed the added value of assimilating several long-lived species, and an improvement in the zonal wind in the Tropics within the troposphere and lower stratosphere. 4D-Var assimilation also induced a correction of zonal wind in the surf zone and a temperature bias in the lower tropical stratosphere.

Keywords: coupled chemistry-meteorology data assimilation; weak and strong data assimilation coupling; Canadian Quick Covariance method (CQC); assimilation of MIPAS temperature observations; ozone-temperature cross-covariance; tracer-wind 4D-Var assimilation

1. Introduction

Data assimilation is a process by which observations are integrated into a model of the atmosphere thereby changing the model state and its associated forecast. Tropospheric observations related to dynamical variables such as temperature, wind, and humidity are continuously collected and

routinely assimilated in weather prediction models. In the stratosphere, there are fewer meteorological observations available and these are mostly related to temperature, but there are, however, several research satellites measuring chemical composition in this region [1]. Important missions began in the early 1990s with the Upper Atmosphere Research Satellite (UARS) [2–4] followed by the Environmental Satellite Envisat [5–7] and NASA’s Earth Observing System (EOS) Aura [8,9]. Instruments on board these satellites typically perform measurements which are tangent to the atmosphere (also called limb soundings) and provide height-resolved retrievals of a number of chemical species as well as temperature. Chemical transformations, especially those related to stratospheric ozone, have an impact on the temperature, while winds change the distribution of chemical tracers (i.e., long-lived species). A natural question which then arises is *“To what extent does the assimilation of chemical observations, and in particular those provided by limb measurements, impact the meteorology on time scales relevant to numerical weather prediction?”* This is the main objective of this study. In Part I we focused on modelling aspects and introduced the coupled model GEM-BACH (GEM with the Belgian Atmospheric Chemistry model). Here we will discuss how these research satellite observations can provide useful information, and in particular, using weak and strong data assimilation coupling.

Coupled data assimilation is a relatively new area of research and development, where assimilation systems can broadly be classified as either weakly or strongly coupled [10–12]. In weakly coupled data assimilation system, each geophysical component (e.g., chemistry, meteorology) has its own independent analysis. The analyses are then used to initialize a coupled model, which produces a coupled model forecast (i.e., the coupling arises through the model forecast and not through the analysis). In a strongly coupled data assimilation system, the analysis is carried out on all variables together. Thus, observations of one geophysical component can have a direct impact on the analysis of the other geophysical component. Weak and strong data assimilation coupling strategies were developed for atmosphere-ocean [12–25] and atmosphere-land-surface coupled systems [26–30].

Coupled meteorology–chemistry data assimilation has primarily been examined in the context tropospheric aerosol-radiation interaction, on short time-scales [31–33], on subseasonal time scales [34] and decadal time scales [35] (also see [36] for a review of chemical data assimilation). It was also used to estimate parameters in the activation of aerosols into cloud droplets [37], and in determining cross-covariance between temperature and constituents (O_3 , NO_2 , and SO_2) using the coupled tropospheric model WRF-CHEM with an ensemble based approach [38]. Coupling can also occur through coupled observation operators. For example, infrared channels of operational meteorological satellites are sensitive to ozone and CO_2 and can benefit from using an ozone assimilation [39] and a CO_2 assimilation [40].

Data assimilation coupling in the stratosphere was also investigated in perspective of weak coupling through ozone–radiation interaction, and as strong coupling using the tracer-wind relation. Weak coupling was investigated at numerical weather prediction centers, such as ECMWF, by considering the ozone–radiation interaction [41] and at the Canadian Meteorological Center (CMC) with the experimental model GEM-BACH [42]. The experiments conducted at ECMWF (European Center for Medium Range Forecasting) were performed with a linearized ozone chemistry and using nadir-sounding stratospheric measurements, whereas those at CMC used a relatively low resolution model but with the full stratospheric chemistry and using limb sounding observations. The CMC study showed an impact on the meteorological forecast skill in the lower stratosphere.

Strong data assimilation coupling has been considered in the context of using chemical tracer observations to infer winds. Some earlier studies using an extended Kalman filter with a simplified two-dimensional transport model showed that wind recovery is very sensitive to the accuracy of chemical observations, and to the concentration fields having sufficient horizontal gradients and with small data voids [43]. It was also shown that constituents in zones of convergence could only determine the winds nearby. Experiments conducted with a one-dimensional model also showed that wind information can still be obtained in the case of a flat concentration field if there are gradients in the concentration error covariances [44]. Using the barotropic vorticity equation with a 4D-Var assimilation

system, Riishøjgaard [45] examined the issues of data density and length of the assimilation window, and arrived at similar conclusions. Using column measurements of ozone with a NWP (Numerical Weather Prediction) model and 4D-Var method, a small improvement in the winds was obtained using simulated observations, but a deterioration using real observations [46]. The negative impact was suspected to result from observational bias. In another study using an operational NWP model with a 4D-Var assimilation system, a small impact (about 0.1 m s^{-1}) was found on zonal wind with no reduction of error standard deviation [47]. These rather unsuccessful results conducted in an operational context suggested that additional studies were necessary. Using an ensemble Kalman filter and an intermediate-complexity model, Milewski and Bourqui [48] demonstrated that information about the ozone-wind cross-covariance is essential in constraining dynamical fields when ozone only is assimilated. Moreover they showed that a further reduction in error can be obtained with an ensemble Kalman smoother [49]. In a series of studies using 4D-Var and the ensemble Kalman filter, Allen et al. [50–52] showed that poorly-specified observation error could lead to an increase in RMS wind error, that observational coverage is important such that wind extraction could be improved if several chemical tracers were used, and that the balance between wind and temperature could be offset by the wind recovery from tracer measurements. We should note that the wind extraction from tracer observation is part of a more general class of joint state-parameter estimation problems (e.g., [53,54]).

The present study took place in the period 2005–2009 with funding, in part, from ESA/ESTEC [55]. This article, henceforth referred to as Part II, summarizes the data assimilation aspects. First, in Section 2, we present the extension of the CMC variational assimilation system to include chemical variables where we discuss in particular the analysis splitting and preconditioning, the extension of balance operators with chemical variables, and the validity of an incremental formulation of adjoint tracer operators for 4D-Var assimilation of long-lived species. In Section 3 we describe the error statistics of chemical variables using the Hollingsworth–Lönnberg method to estimate the error variances, and using the Canadian Quick Covariance (CQC) method to obtain non-separable error correlations. We also discuss the method and issues related to the cross-covariances between temperature and ozone. In Sections 4 and 5 we illustrate the benefits of using limb sounding temperatures from the MIPAS instrument on board Envisat to improve the AMSU-A bias correction and better simulation of temperature and transport in the stratosphere. We then discuss a weak coupling data assimilation experiment involving ozone and its impact on meteorological forecasts and show that a simplified linearized ozone chemistry gives comparable results to the full chemistry in terms of impact on thermodynamics. Then in Sections 7 and 8 we discuss results from strong coupling experiments, first in a 3D-Var context using a balance operator between ozone and temperature, and then in 4D-Var assimilation of several long-lived species, i.e., O_3 , CH_4 , and N_2O to correct the winds.

2. Extension of 3D-Var and 4D-Var for Chemical—Meteorological Coupling

The assimilation system scheme used here consists of a model integration step to obtain a six-hour forecast (called the background state), and an assimilation step in which observations are used to correct the background state and obtain an analysis. This analysis is then used to initialize the next six-hour forecast, and the cycle is repeated. In this study, the assimilation step employs a variational analysis solver that can be run in one of three modes:

- 3D-Var: In this case, all observations collected over the six-hour assimilation window are assumed to be valid at the central time. Observation departures from the model state (called innovations) are computed with respect to the background state valid at the central time of the window [56].
- 3D-FGAT (First Guess at Analysis Time): This scheme is a variant of 3D-Var in which the innovations are evaluated by comparing each observation with the model output valid at the observation time [57] (actually closest to a 1-hour bin). This is the default configuration used in all 3D-Var experiments in this study.

- 4D-Var: Extending 3D-Var to 4D-Var can be achieved by including the forward model integration as part of the observation operator (the observation operator computes the model equivalent of the observation) [58,59].

It is generally assumed in variational analysis that observation errors and background errors are uncorrelated, both unbiased, and Gaussian distributed. Producing a minimum variance estimate (or maximum a posteriori estimate for nonlinear observation operator), called the *analysis*, leading numerically to a large-scale minimization problem of a quadratic function that can be solved by unconstrained minimization techniques. This requires suitable preconditioning, and an adjoint observation operator (that is equivalent of a matrix transpose of the Jacobian of the non-linear observation operator).

4D-Var mode also requires the adjoint of the linearized (prognostic) model, commonly called the adjoint model. The linearization is made about a nonlinear model solution, but in the incremental form of 4D-Var, the linearized model is not required to be at the same resolution nor contain the same physics. In this study, however, the linearized model on which the adjoint model is based is at the same resolution (following the discussion in Section 6 Part I) but has no adjoint of ozone–radiation interaction and no adjoint of chemistry. All 4D-Var chemical data assimilation experiments were conducted using only long-lived species treated as passive tracer (over a period of six hours; the time length of the 4D-Var time window).

Background error covariances and observation error covariances are needed to compute a minimum variance estimate. The background error correlation model used in this study is, for each variable, homogeneous and isotropic (i.e., invariant under rotation) on a sphere, and non-separable, meaning that the vertical and horizontal correlation structures are interconnected. The cross-variable error correlations are obtained by a transformation of variables, involving what are called balance operators obtained from a regression analysis following a Gram–Schmidt orthogonalization procedure (Section 2.3). For the dynamical model variables, there are balance operators to represent the geostrophic and hydrostatic balance, and also the Ekman balance in the planetary boundary layer. In this study we introduced a balance operator between ozone and temperature that was obtained from either a fully coupled meteorological-chemical model or from a LINearized OZone photochemical model such as LINOZ [60] (see [61] for the implementation of LINOZ with semi-Lagrangian transport).

Applying balance operators between variables (when needed) and representing the error correlations in spherical harmonics, it is possible to completely diagonalize the error correlation matrix [62]. The covariances in physical space and between all variables can then be obtained through a series of transformations on a vector. With this formulation the background error covariance matrix can, in principle, be easily expanded to include other non-meteorological variables. This approach was taken to extend our meteorological data assimilation system to a coupled chemistry-dynamics data assimilation solver. The numerical coding effort began in another Canadian study [63] and was completed here with cross dynamics-chemistry balance operators and the 4D-Var chemical extension (passive tracer).

The last step of the development concerns the preconditioning, which will be discussed in Section 2.1. In principle, the control vector should contain all the meteorological and chemical variables and which, in our case, consists of 57 (constituent) + 4 (dynamical) = 61 three-dimensional fields. In developing the preconditioning, it was realized, however, that only the observed variables had to be added to the control variable. In our case, this amounts to 4 + 6 = 10 three-dimensional fields (horizontal winds, temperature, water vapor, O₃, CH₄, NO₂, N₂O, HNO₃, and ClONO₂). As far as unobserved constituents that have correlated errors with observed constituents, their minimum variance estimate can be obtained off-line after the minimization.

2.1. Analysis Splitting between Observed and Unobserved Variables

The state of a chemical-meteorological model comprises of all prognostic meteorological and chemical model variables, which, in our case, represents 61 three-dimensional fields (57 advected

chemical species + 4 dynamical variables). In principle, a state estimate should be conducted on all these variables. Yet, only a small fraction of these variables are observed. For example, MIPAS/ESA [5–7] chemical observations are mostly limited to O_3 , N_2O , NO_2 , CH_4 , and HNO_3 . We will derive in this section a computational simplification that allows splitting the analysis into observed and unobserved variables parts.

Let $\mathbf{Z}$ be the complete chemical-meteorological state vector be decomposed into *observed* variables $\mathbf{X}$ and *unobserved* variables $\mathbf{U}$, i.e.,

$$\mathbf{Z} = \begin{pmatrix} \mathbf{X} \\ \mathbf{U} \end{pmatrix} \quad (1)$$

The analysis of all state variables using a 3D-Var algorithm consists of minimizing the following cost function:

$$J(\mathbf{Z}) = \frac{1}{2}(\mathbf{Z} - \mathbf{Z}^f)^T \mathbf{B}^{-1}(\mathbf{Z} - \mathbf{Z}^f) + \frac{1}{2}(\mathbf{y} - H(\mathbf{Z}))^T \mathbf{R}^{-1}(\mathbf{y} - H(\mathbf{Z})) \quad (2)$$

where $\mathbf{y}$ denotes the observation vector (i.e., all observations of all observed variables at a given time), H is the observation operator, $\mathbf{R}$ the observation error covariance matrix, and $\mathbf{B}$ is the full state background error covariance matrix that can be decomposed into:

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{\text{XX}} & \mathbf{B}_{\text{XU}} \\ \mathbf{B}_{\text{UX}} & \mathbf{B}_{\text{UU}} \end{pmatrix} \quad (3)$$

which includes covariances and cross-covariances between observed and unobserved variables.

Developing a preconditioning for the cost function in Equation (2) with the full state vector $\mathbf{Z}$ we found that the minimization of $J(\mathbf{Z})$ can be split into two parts: A minimization of the cost function involving only the observed variables and observations, which takes the form:

$$J(\mathbf{X}) = \frac{1}{2}(\mathbf{X} - \mathbf{X}^f)^T \mathbf{B}_{\text{XX}}^{-1}(\mathbf{X} - \mathbf{X}^f) + \frac{1}{2}(\mathbf{y} - H(\mathbf{X}))^T \mathbf{R}^{-1}(\mathbf{y} - H(\mathbf{X})) = J_B + J_O \quad (4)$$

and a regression between the analysis increments of the unobserved variables with the increments of the observed variables, of the form:

$$\mathbf{U}^a - \mathbf{U}^f = \mathbf{B}_{\text{UX}} \mathbf{B}_{\text{XX}}^{-1}(\mathbf{X}^a - \mathbf{X}^f) \quad (5)$$

This property is called *analysis splitting*. Note that the cost function in Equation (4) is composed of two parts, the background cost function J_B and the observation cost function J_O . We should also note that analysis splitting is quite general, and holds, in particular, when the observation operator is nonlinear (the derivation is presented in Appendix A), but does not apply to 4D-Var as the J_O term is not separable in that case.

The concept of analysis splitting is interesting in terms of assimilation coupling and for practical reasons. Consider the behavior of unobserved variables $\mathbf{U}$ in either strongly-coupled or weakly-coupled data assimilation system. The analysis increment in a strongly-coupled data assimilation system would use $\mathbf{U}^a$ (Equation (5)) as part of the initial condition $\mathbf{Z}^{\text{initial}} = (\mathbf{X}^a, \mathbf{U}^a)^T$ for a coupled model. In a weakly-coupled data assimilation system, we would use $\mathbf{U}^f$ (instead of $\mathbf{U}^a$) to initialize the unobserved space, and furthermore $\mathbf{X}^a$ would be obtained from an uncoupled analysis. That means that, for example, in weakly-coupled chemistry–meteorology data assimilation, $\mathbf{X}^a = (\boldsymbol{\mu}^a, \boldsymbol{\chi}^a)^T$, where $\boldsymbol{\mu}$ are meteorological and $\boldsymbol{\chi}$ chemical variables so that in the coupled model the initial condition is given by $\mathbf{Z}^{\text{initial}} = (\boldsymbol{\mu}^a, \boldsymbol{\chi}^a, \mathbf{U}^f)^T$.

Analysis splitting is also practical as it reduces the optimization state space dimension for the 3D-Var. It also offers the possibility to examine the impact of the analysis on unobserved variables independently of the core variational optimization. In the absence of adequate information about cross-covariances between observed and unobserved variables, the increments of unobserved variables can be selectively removed from of the analysis in a simple manner.

2.2. General Description of the 3D-Var-CHEM

The CMC 3D-Var scheme developed for meteorology [56] and extended to include chemical variables [63] was further extended in this study to include cross-covariances between observed variables and between observed and unobserved variables using a balance operator. The general framework of balance operators will be discussed in Section 2.3 and the associated error statistics in Section 3.3. Cross-covariances involving chemical variables were estimated point-wise, while the meteorological variable error covariances (and cross-covariances) were computed in spectral space as in by Derber and Bouttier [64].

The coupled chemical-meteorological model state used in the 3DVar-Chem in Equation (4) consists of $\mathbf{X} = (\psi, \chi, T, \ln(q), c_1, \dots, c_N, p_s)^T$, where ψ is the streamfunction, χ the velocity potential (not the confuse with χ used in the previous sub-section), T the temperature, q the (tropospheric) water vapor mixing ratio, p_s the surface pressure and N observed chemical constituent with mixing ratios $c_1, \dots, c_N$. The state vector in 3D Var-Chem is such that all 3D fields are grouped together, followed by the 2D field p_s . As explained in Section 2.1, the state augmentation is limited only to observed variables/species.

A flow chart of the 3D-Var-Chem (omitting some intermediate steps) is given in Figure S1 (Supplementary Material). The 3D-Var-CHEM code can be used for: (1) general assimilation, (2) identification of observation outliers (background check), (3) monitoring (determination of O-P only), (4) testing by way of single observation experiments, and (5) stand-alone analysis splitting, i.e., Equation (5).

The minimization of the cost function in Equation (4) is performed after a transformation of variables, $\xi = \mathbf{L}\mathbf{X}$ where $\mathbf{B}_{\mathbf{X}\mathbf{X}} = \mathbf{L}\mathbf{L}^T$, which simplifies the background penalty term to a simple quadratic of the form, $(\xi - \xi^f)(\xi - \xi^f)^T$ - a transformation step called preconditioning. The minimization is then performed on the transformed variable ξ using an efficient quasi-Newton algorithm adapted for large-scale problems [65]. The preconditioning used in 3DVar-Chem follows what is done for the meteorological variables [56]. The key aspect of this computation resides in the fact that $\mathbf{L}$ times a vector $\mathbf{X}$, can be obtained as a sequence of operators, without the need to store any large matrices. This property arises principally from the assumption that the horizontal error correlation is assumed to be homogeneous and isotropic on the sphere. For such correlations, the spectral representation is diagonal in spectral space (see for example [56,62,66,67]). The sequence of operations then becomes as follows: (1) We multiply the spectral representation of the state $\mathbf{X}$ with the square root of the spectral coefficient of the correlation model $\Lambda^{1/2}$ (which is a diagonal matrix for separable models and block diagonal for non-separable correlation model); (2) perform a transform from spectral to physical space. First using a spectral transform $\mathbf{S}$ which gives a field on the Gaussian grid, and then followed by an interpolation $\mathbf{G}$ from the Gaussian grid to a regular latitude–longitude grid; (3) multiply the resulting fields by the error variances (that takes the form of a diagonal matrix of error standard deviation Σ); and (4) using balance operators $\mathbf{M}$ (to be define below), transform the primary fields into fields of physical significance accounting for cross-correlations between them. This is how we obtain, for example, the velocity potential from the stream function and an unbalanced velocity potential. This last operation is obtained through a balance operator. In the end, the square root of $\mathbf{B}_{\mathbf{X}\mathbf{X}}$ is given as $\mathbf{L} = \mathbf{M}\mathbf{\Sigma}\mathbf{G}\mathbf{S}\Lambda^{1/2}$, a sequence of operators.

Before we discuss the balance operators, we should note two things: (1) The CMC 3D-Var system uses a non-separable error correlation model. This means that for each horizontal wavenumber there is a unique vertical correlation matrix, and this introduces a dependence between horizontal and vertical scales; and (2) although it is usual in meteorological applications to perform the minimization on an analysis grid of lower resolution than the model grid (e.g., [56] and in 4D-Var incremental formulation [68]), as we argued in Part I, Section 6, the meteorological model and analysis increment, as well as the chemical forecast model and the chemical analysis increment should all be on the same grid, in order to avoid a loss of information.

2.3. Balance Operators

Balance operators have been introduced in meteorological data assimilation to account implicitly for the balance between mass and momentum in the background error covariance either through deterministic relationships (e.g., linear balance equation) [56,69–71] or through statistical regression [64, 72]. The multilinear regression approach [72] is particularly useful to define balance operators involving chemical species. In particular the streamfunction $\delta\psi$, velocity potential $\delta\chi$, temperature δT , and ozone δO_3 which are known to be correlated, can be transformed into a set of *uncorrelated* background error variables (denoted with a superscript u), as follows:

$$\begin{aligned}\delta\psi^u &= \delta\psi \\ \delta\chi^u &= \delta\chi - \mathbf{B}_{\chi\psi} \mathbf{B}_{\psi\psi}^{-1} \delta\psi \\ \delta T^u &= \delta T - \mathbf{B}_{T\psi} \mathbf{B}_{\psi\psi}^{-1} \delta\psi - \mathbf{B}_{T\chi^u} \mathbf{B}_{\chi^u\chi^u}^{-1} \delta\chi^u \\ \delta O_3^u &= \delta O_3 - \mathbf{B}_{O_3\psi} \mathbf{B}_{\psi\psi}^{-1} \delta\psi - \mathbf{B}_{O_3\chi^u} \mathbf{B}_{\chi^u\chi^u}^{-1} \delta\chi^u - \mathbf{B}_{O_3T^u} \mathbf{B}_{T^uT^u}^{-1} \delta T^u.\end{aligned}\quad (6)$$

The transformation from any set of correlated errors to uncorrelated error variables, as in Equation (6), can also be explained geometrically by adopting a Hilbert space representation of the random variables [73,74] and followed by Gram–Schmidt orthogonalization procedure (see Appendix B for a geometrical derivation).

Back-substituting, we recover the transformation from uncorrelated variables to correlated variables, in the form:

$$\begin{pmatrix} \delta\psi \\ \delta\chi \\ \delta T \\ \delta O_3 \end{pmatrix} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ \mathbf{E} & \mathbf{I} & 0 & 0 \\ \mathbf{N} & 0 & \mathbf{I} & 0 \\ \mathbf{G} & 0 & \mathbf{F} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \delta\psi^u \\ \delta\chi^u \\ \delta T^u \\ \delta O_3^u \end{pmatrix} = \mathbf{M} \begin{pmatrix} \delta\psi^u \\ \delta\chi^u \\ \delta T^u \\ \delta O_3^u \end{pmatrix}\quad (7)$$

where:

$$\mathbf{E} = \mathbf{B}_{\chi\psi} \mathbf{B}_{\psi\psi}^{-1} \quad \mathbf{N} = \mathbf{B}_{T\psi} \mathbf{B}_{\psi\psi}^{-1} \quad \mathbf{G} = \mathbf{B}_{O_3\psi} \mathbf{B}_{\psi\psi}^{-1} \quad \mathbf{F} = \mathbf{B}_{O_3T^u} \mathbf{B}_{T^uT^u}^{-1}\quad (8)$$

In the above we have set to zero (i.e., neglected) all cross-covariances involving the uncorrelated velocity potential, χ^u . Equation (8) consists of the balance operators involving the main variables only. The list of variables and their associated balance operator in Equation (7) is actually incomplete. To be complete it should include surface pressure, which follows the same structure as temperature, and humidity, which is assumed to be uncorrelated with any other meteorological variable. In chemistry, we could have also introduced a cross-covariance between long-lived species such as (N_2O , CH_4) or chemically-related species such as (O_3 , NO_2), but we have not done so here.

From Equation (7) we obtain the background error covariances, which can be rewritten by splitting the covariances into variances and correlations as follows:

$$\mathbf{B}_{\mathbf{xx}} = \mathbf{M} \begin{pmatrix} \mathbf{B}_{\psi\psi} & 0 & 0 & 0 \\ 0 & \mathbf{B}_{\chi^u\chi^u} & 0 & 0 \\ 0 & 0 & \mathbf{B}_{T^uT^u} & 0 \\ 0 & 0 & 0 & \mathbf{B}_{O_3^uO_3^u} \end{pmatrix} \mathbf{M}^T = \mathbf{M} \Sigma_{\mathbf{xx}} \begin{pmatrix} \mathbf{C}_{\psi\psi} & 0 & 0 & 0 \\ 0 & \mathbf{C}_{\chi^u\chi^u} & 0 & 0 \\ 0 & 0 & \mathbf{C}_{T^uT^u} & 0 \\ 0 & 0 & 0 & \mathbf{C}_{O_3^uO_3^u} \end{pmatrix} \Sigma_{\mathbf{xx}}^T \mathbf{M}^T \quad (9)$$

where $\Sigma_{\mathbf{xx}} = \text{diag}(\Sigma_{\psi\psi}, \Sigma_{\chi^u\chi^u}, \Sigma_{T^uT^u}, \Sigma_{O_3^uO_3^u})$ is a diagonal matrix of error standard deviations for all uncorrelated variables. Note that each correlation matrix $\mathbf{C}$, is actually represented spectrally as $\mathbf{C} = \mathbf{S} \Lambda \mathbf{S}^{-1}$ where $\mathbf{S}$ and $\mathbf{S}^{-1}$ are the spectral direct and inverse transform and Λ is a diagonal or block-diagonal ($n_{lev} \times n_{lev}$) matrices of spectral coefficients. For computational efficiency, the balance operators in $\mathbf{M}$ are simplified as block diagonal matrices ($n_{lev} \times n_{lev}$) for each latitude with an error variance that depends on height and latitude (using a Legendre polynomial expansion).

To start a 3D-Var or 4D-Var analysis using preconditioning requires the inverse of the square root of $\mathbf{B}_{\chi\chi}$, and, thus, we need to know the inverse of $\mathbf{M}$, which turns out to be easy to obtain as:

$$\mathbf{M}^{-1} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ -\mathbf{E} & \mathbf{I} & 0 & 0 \\ -\mathbf{N} & 0 & \mathbf{I} & 0 \\ \mathbf{NF} - \mathbf{G} & 0 & -\mathbf{F} & \mathbf{I} \end{pmatrix} \quad (10)$$

we should also note that computation using the balance operator never involve the computation of an inverse of a large matrix. For example, to compute $\delta\chi = \mathbf{E}\delta\psi = \mathbf{B}_{\chi\psi}\mathbf{B}_{\psi\psi}^{-1}\delta\psi$ we begin by first solving for $\delta\psi$ in the equation $\delta\psi = \mathbf{B}_{\psi\psi}\delta y$, then followed by computing $\delta\chi = \mathbf{B}_{\chi\psi}\delta y$.

2.4. D-Var Tracer Extension

The 3D-Var algorithm can be extended to 4D-Var by including the model integration as part of the observation operator [58,59]. The minimization of the 4D-Var cost function with the adjoint of the original model including the full physics can be difficult and computationally demanding. Instead, an incremental formulation of 4D-Var [68] can be used where the minimization of the 4D-Var cost function is approximated by a series of minimizations involving the adjoint of a tangent linear model with simplified physics and at a lower resolution [68,75,76], called the inner-loop and where its solution is used to update the full model trajectory in an outer-loop. The outer-loop trajectories define new innovations and a new cost function and the method cycles through several outer loops, each of which requires the minimization in an inner loop. At CMC, the physics component of the adjoint model includes only the vertical diffusion, surface drag, orography blocking, stratiform condensation and convection. The simplified adjoint model is also run at a resolution of $1.5^\circ \times 1.5^\circ$, which is the same resolution as that of the GEM-BACH model. For the chemistry component of GEM-BACH, the adjoint is simplified by considering only the adjoint of advection transport. There is no adjoint of chemistry. The tangent-linear model of semi-Lagrangian advection was discussed in Polavarapu [77] and the properties of the adjoint in Tanguay and Polavarapu [78]. The key element in the implementation of 4D-Var for GEM-BACH is that the minimization is performed within the inner loop which uses the tracers of observed species only (with the simplified physics). The outer-loop uses the full chemistry and physics.

4D-Var assimilation of ozone was conducted between 300 hPa and 10 hPa where ozone is long-lived and, thus, be approximated as a passive tracer. To illustrate the validity of the incremental tracer approach for ozone, Figure 1 shows the observation cost function J_o as a function of iteration. The solid black is the result of the first inner loop (up to iterate 42), while the dashed line refers to the cost function after the first update of the outer loop, during the second inner loop. We observe that the observation cost function is continuous, except for a small adjustment between the last iterate of the first inner loop and the beginning of the second inner loop, thus, indicating no significant issues with the incremental tracer approach.

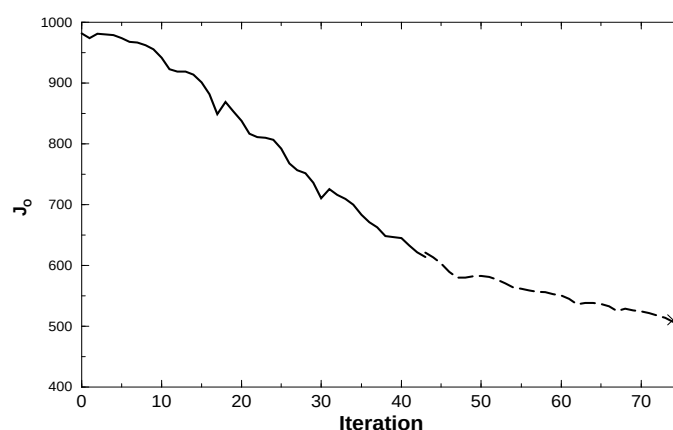


Figure 1. Observation component of the cost function for ozone assimilation as a function of iteration. Solid line is associated with the value of J_o of the first inner loop and the dashed line the value of J_o of the second inner loop.

3. Error Statistics

An accurate estimation of the observation and background error statistics is important in data assimilation as these control (at analysis time) the weight of the observations and the structure functions that spread information in space and to other model variables. The innovations contain the basic information to estimate the observation and background error statistics but this information is actually combined, i.e., not separated in its respective components. Under the assumption that observation errors are spatially uncorrelated and background errors are spatially correlated it is, however, possible to separate the observation and background error statistics. The Hollingsworth–Lönnberg (HL) method [79,80] does precisely this and is based on computing the distance between pair of observations usually at fixed locations. Here, we demonstrate that this method can also be used with a polar orbiting limb sounder, such as the MIPAS instrument on board Envisat using the distance between observations profiles along the orbit. With this approach, we were able to derive the observation and background error statistics of the observed chemical species. We should add that there are other methods in observation space that can provide observation and background error statistics, such as Desroziers [81] and Desroziers and Ivanov [82] methods, but these are based on different assumptions (see [83]).

Any of these methods are limited as they can only estimate error statistics of the observed variables in the observation space, which is insufficient to prescribe the error statistics needed for an assimilation system. Additional information can be obtained by using model output methods, such as the ensemble methods and the lagged-forecast method also known as the NMC (National Meteorological Center) method. Ensemble methods require an ensemble of model forecasts, but conducting an ensemble of the GEM-BACH model runs would be computationally demanding, and would require tuning of model error (i.e., inflation) and localization. The lagged-forecast method, widely used in meteorology, uses differences of forecast valid at the same time(s) and is based on having a complete observational coverage. However, Bouttier [84] argued that the lagged-forecast method is strongly related to the difference between the forecast error covariance and analysis error covariance, and not specifically on the forecast error covariance. Consequently, the lagged-forecast method cannot be used in extensive regions where there are no observations, as the difference between the forecast and analysis error covariances would then be close to zero.

Additionally, we should note that the lagged-forecast method is generally used to obtain the background error correlations, not the error covariances. The error variances are obtained through other means by using the innovation variance or estimates obtained by the HL method. In atmospheric chemistry, the observational coverage is generally not uniform and often has large data voids in each analysis. In this study, in particular, our main chemical observational source comes from a single polar orbiting satellite, i.e., MIPAS. The horizontal coverage of MIPAS in 6 h (analysis time window) is

limited to about a quarter or a third of the global domain. In addition, some chemical components have a strong diurnal cycle. The use of the lagged-forecast method in this context is, thus, questionable. An alternative method that has been used in stratospheric and mesospheric data assimilation consists of obtaining statistical information from six-hour differences of a single model output. This method, originally developed by Yves Rochon (personal communication) is known as the Canadian Quick Covariance (CQC) [63].

In this study, we used a combination of these methods depending on the variable type, i.e., meteorological or chemical, as summarized in Table 1. The newer approaches, such as the CQC method and the HL method used with MIPAS, will be described in the following subsections.

Table 1. Summary of error estimation methods.

Variable Type	Statistical Parameters	Statistical Assumption and Methods	
		Observation Error	Background Error
meteorological	variances	innovation-based	combination of innovation-based and lagged-forecast (NMC) methods
	correlations	spatially uncorrelated	lagged-forecast method
chemical	variances	Hollingsworth–Lönnberg (HL) method as function of height	Hollingsworth–Lönnberg (HL) method as function of height
	correlations	spatially uncorrelated	Six-hour difference (CQC) method

3.1. Estimation of Error Variances by Autocorrelation of Innovations along the Satellite Track

The observation error obtained from innovations is comprised of the following: the instrument error, the forward modeling and retrieval errors, the error due to the interpolation from observation location to model grid point, and the error due in part to the subgrid scale variability not resolved by the atmospheric model [85]. Aside from the instrument error, the sum of these other errors is called the representativeness error. The model forecast error is generally correlated horizontally over large distances, typically about 1000 km. As we shall see, we can assume that observation error is either spatially uncorrelated or correlated over much shorter distances, allowing us to estimate the observation error variance and forecast error variance by constructing spatial autocorrelation function of innovations. The spatially correlated part at zero distance of the innovation autocovariance (or the intercept) can be attributed to the model forecast error variance while and the remaining part measures the spatially uncorrelated part attributed to the observation error variance, which is in essence the HL method.

To illustrate the use of the HL method with chemical species we have conducted an assimilation of methane observations from MIPAS over a period of three weeks in August–September 2003, using 10% error for the background error and the retrieval error provided by the instrument team for observation error. We shall refer to these first guess error statistics as the old error statistics. These are not be taken as the true error statistics but are used only to derive a first set of innovation statistics from the assimilation cycle. Since MIPAS observational profiles are spaced uniformly at about 530 km along the satellite track [7], we construct an along-track spatial auto-covariance of the innovations, which is illustrated in Figure 2 at 63 hPa.

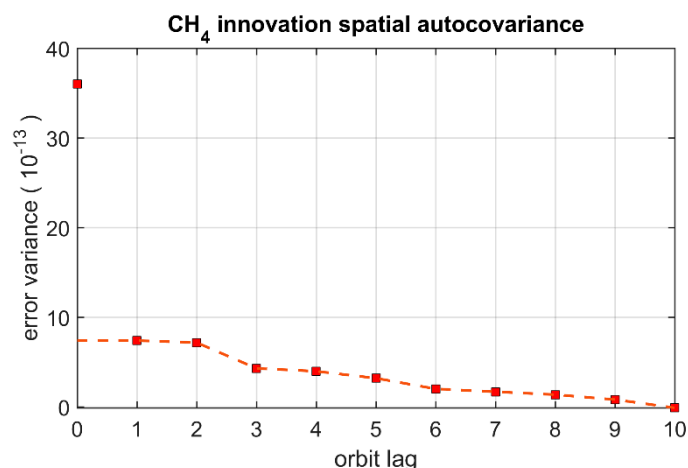


Figure 2. Spatial autocovariance of innovation for MIPAS CH₄ at 63 hPa. The abscissa shows the lags between orbits (each orbits are separated by about 530 km along the satellite track). The red squares represent the sample autocovariance values, and the dashed curve is a piecewise linear interpolation between the sample points. Note that the sample covariance at zero distance separation is at the top of the graph (near 36×10^{-13}), and the dashed line is extrapolated at lag 0.

The units of the horizontal axis are profile lags along the satellite track, with spacing of 530 km. We note that the innovation variance at zero separation is 36×10^{-15} and is distinctively different from the extrapolated intercept of the spatially correlated part, estimated to be around 8×10^{-15} . Such a separation of values at zero distance is observed at all levels and for all species. This supports our assumption that the observation error is either spatially uncorrelated or that the spatial correlation length is much shorter than the background error correlation.

The estimates of observation and background error variances for MIPAS CH₄, obtained from HL method are displayed in the left panel of Figure 3. We note that the MIPAS CH₄ observation error variances are significantly larger than those of the (model) background errors except in the region between 2 and 0.5 hPa. This indicates that MIPAS CH₄ observations will have a small impact on the analysis, the main impact region is limited to 2–0.5 hPa and also in the lower stratosphere between 100 and 50 hPa.

Comparison of three different estimates of observation error variance is shown in Figure 3 (right panel). One is the observation error variance provided by the instrument, i.e., the blue curve. We note that the instrument error variance is always smaller than or equal to the estimated observation error variance obtained with the HL method. This is consistent with the fact that the estimated observation error using innovation statistics includes the representativeness error, which is usually significant. The estimate of observation error variance using the Desroziers method [81] is shown in green and is close to the HL estimate in the mid-to-lower stratosphere from 100 hPa to 3 hPa. However, at higher altitudes important differences are noted. Accurate error variance estimates with the Desroziers method [81] relies on: (1) the assumption that the Kalman gain is nearly optimal (i.e., the analysis error is the minimum), and (2) that the innovation covariance matches the sum of the prescribed observation and background error covariances in observation space [83] (i.e., the innovation covariance consistency). Since by construction the HL variance estimates respect the innovation covariance consistency (at least in terms of variance), that may explains the discrepancy between HL and Desroziers estimates of observation error variance. We also note that at 0.1 hPa the Desroziers estimate seems to be smaller than the instrument error variance. However, in the lower stratosphere, near 100 hPa, both HL and Desroziers estimates are nearly equal, but much larger than the instrument error variances, indicates that representativeness error is significant in the lower stratosphere.

One simple way to quantify the weight given to the observation is to look at the scalar form of the Kalman gain, which involves only the ratio of estimated error variances. A scalar Kalman gain

close to one indicates that the solution is determined mostly by the observations while a gain of zero implies that the observations have no influence. In the supplementary material (Figure S2) the reader will find the scalar Kalman gain for O₃, CH₄, N₂O, NO₂, HNO₃, and H₂O that were assimilated in the course of this study. We note for instance that for O₃ the gain is about 0.2 in the lower stratosphere and steadily increases to about 0.6 in the upper stratosphere. A similar situation was found for the long-lived species CH₄ and N₂O. However, the NO₂ gain is close to one in the upper stratosphere, indicating that the model has a small impact at these altitudes. As for HNO₃, the gain increases with height and reaches a maximum value of 0.8 at 4 hPa, then decreases with altitude. Chemical water vapor (H₂O) is presented in terms of the log of concentration.

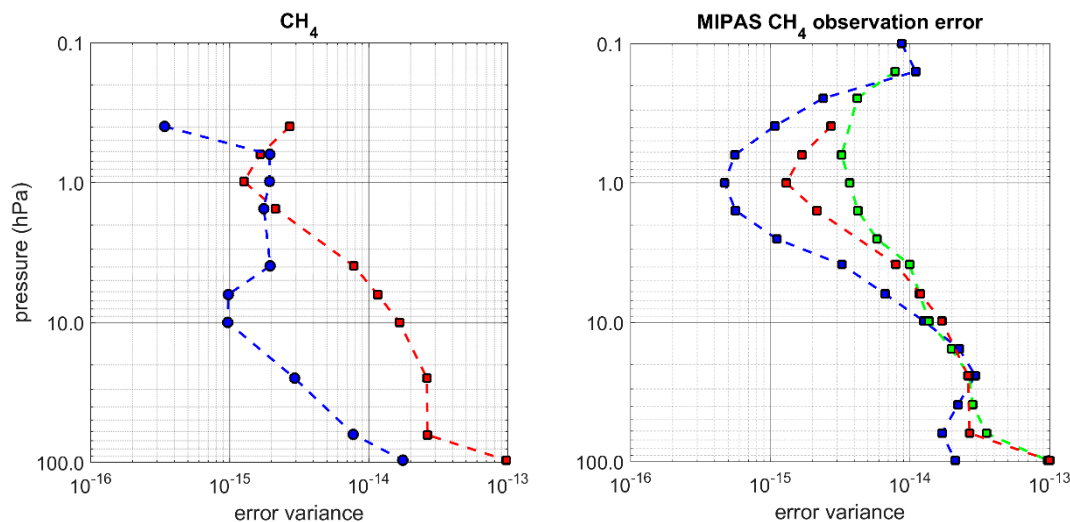


Figure 3. Estimated error variance for CH₄/MIPAS as a function of height (pressure in hPa). (**Left panel**) The estimated background error variance (blue with circles) and observation error variance (red with squares) using the HL method. (**Right panel**) Three different estimates of observation error variance. The blue curve with squares is the estimate given by the instrument team (i.e., the instrument error), the red curve with squares is the observation error variance obtained from the HL method (note it is identical to the red curve in the left panel), and the green curve with squares is the observation error variance estimated from the Desroziers method [81] (single iterate).

3.2. The Canadian Quick Covariance Method

Let us first recall that the NMC method consists of obtaining a homogeneous isotropic and horizontal/vertical non-separable correlation model on a sphere using a spherical harmonics representation of 48-hour minus 24-hour model forecasts valid at the same time (see Errera and Ménard [62] for a description on the use of spherical harmonics and how to construct error correlations, and Gauthier [72] for aspects related to meteorological applications). The Canadian Quick Covariance (CQC) method [63] is similar to the NMC method except that it uses six-hour differences of pure model forecasts. The CQC method does not involve an assimilation cycle and, thus, does not depend on observation density, and can be obtained for any variables, observed or not. This latter feature is particularly interesting for atmospheric chemistry, where many species are unobserved, or the observational coverage is limited. It should be stressed that each difference is computed using forecast valid at two different times. The information that the CQC method represents is actually the tendency, comprising advection and model physics or chemistry. Writing a model equation in the form:

$$\frac{\partial X}{\partial t} + \mathbf{V} \cdot \nabla X = F(X) \quad (11)$$

where X is a prognostic model variable and V is the wind, the CQC method, thus, derives its spatial error statistics from the six-hour differences which represent model tendencies:

$$X(x, t + 6) - X(x, t) \approx \int_0^6 \{-\mathbf{V} \cdot \nabla X + F(X)\} d\tau \quad (12)$$

where x is the spatial coordinate.

It has been argued [63] that since the large-scale motion does not change in a 6 h time period, it may explain why the stream function and unbalanced temperature correlation obtained from the CQC method have less signal in wavenumbers 10 and lower in comparison with the correlation using the NMC method. However, it is known that the NMC method has a tendency to give too much spectral error variance at these wavenumbers for meteorological correlations fields [86]. It is, thus, unclear whether the CQC method has an actual deficiency at large scales. The correlation power spectra of the species that were used in the assimilation are shown in Figure S3 (Supplementary Material) indicate generally a maximum in power at the large scales (wavenumbers 8–10) as one would expect.

To compute the background error correlation with the CQC method we first compute the variance of six-hour differences of pure model forecasts. These zonal-mean variances as a function of height are presented in Figure S4 (Supplementary Material). We then normalize the six-hour differences by the square root of these error variances to obtain an ensemble set of model variables that will be used to represent the error correlations. This ensemble set is then represented spectrally, as with the NMC method, from which by using the spectral representation of a non-separable correlation model we obtain a vertical correlation matrix $nlev \times nlev$ (see [56] or [62] for details) for each total wave number n . In a non-separable correlation model, we can compute a power spectrum as a function of the horizontal wavenumber n and vertical level, as shown in Figure S3 (Supplementary Material). A horizontal-vertical separable correlation model has a (horizontal) power spectrum that does not change with height. The non-separability of correlation models in meteorology has been attributed to dynamical processes, such as baroclinic instability [56,59]. The chemical constituent correlation models shown in Figure S3 indicate that CH_4 and H_2O above 70 hPa and NO_2 above 10 hPa are (roughly) separable as well as N_2O in the entire stratosphere (although the correlation field is very noisy). However, the correlations for O_3 and HNO_3 at large scales (for wavenumbers smaller than 20) are not (horizontally/vertically) separable. These results suggest that a chemical specie has a non-separable correlation model if its chemical time-scale lies within the statistical sampling time-range which here is anything larger than 6 h and smaller than one month, otherwise the correlation of the species is separable.

The resulting correlation length can also be computed. Figure S5 of the Supplementary Material, shows the horizontal correlation lengths for six constituents, which typically varies from 200 km (in the troposphere) to 400 km in the upper stratosphere. These correlation length-scales seem to be much smaller than the correlation length we obtain (visually) from the spatial autocovariance of innovations. Figure 2 shows that the background error correlation decreases by a value of about 0.6 at about three profile lags along an orbit, thus, indicating a correlation length of roughly 1600 km. Despite the fact the HL method has a tendency in practice to overestimate the spatial correlation length scale compared with length-scale obtained with the maximum likelihood method [87], the correlation length scale obtained with the CQC method for chemical constituents seem too small. However, since we have no means to correct for these deficiencies, we continue to use the spectral coefficients as is in the correlation models of the chemical constituents.

The background error covariance matrix is then obtained by using the background error variance estimated by the HL method and the correlation estimated using the CQC method. Thus, the background error covariance matrix is identical at all latitudes and longitudes and varies only as a function of height. We conducted a series of univariate constituent data assimilation experiments, using the background error covariance above and the observation error variance obtained from the HL

method and computed the mean analysis increment over the period of 17 August to 5 September 2003. During this time period a strong energetic particle precipitation from the mesosphere affected the polar region down to the middle stratosphere and created large NO_2 and HNO_3 mixing ratio increments [88]. Figure 4 presents the mean analysis increment for HNO_3 .

We note that the analysis increment of HNO_3 with the new statistics is larger and self-organized, indicating vertical descent of HNO_3 in the polar vortex [88], while the old statistics give rather random results with numerous small-scale features. The analysis increments for chemically active species such as O_3 and NO_2 (Figure S6 in Supplementary Material) also appear to be larger and physically coherent, while those of passive tracer (CH_4 , N_2O) are not changed significantly, remaining spatially random with both old and new statistics, with the difference that the increments with the new statistics are of somewhat larger scales.

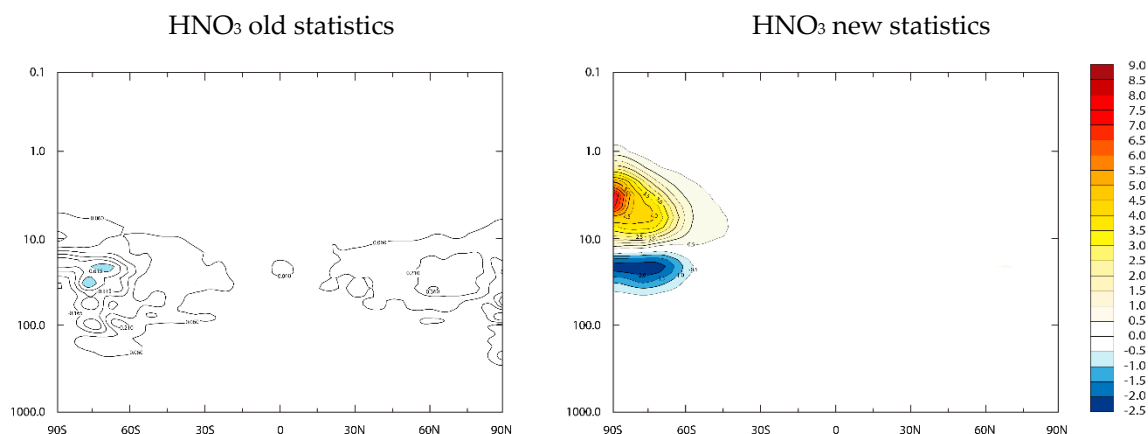


Figure 4. Zonal mean analysis increment for HNO_3 as function of height (in hPa). (**Left panel**) Using the first guess or old statistics. (**Right panel**) Using the new statistics consisting of CQC correlation and HL error variances. The value of the increment should be scaled by 10^{-9} volume mixing ratio. Note that the contour interval on the left panel are much smaller than on the right panel, and only one shaded blue contour (values between -0.5 and -1.0) appears in the left panel.

3.3. Cross-Covariance Estimates

The use of cross-covariances between meteorological and chemical variables in a 3D-Var assimilation is a distinctive feature of our study. As discussed in Part I (Section 2.1), ozone and temperature are related by photochemistry above 10 hPa. Empirical relations of the form given by Equation (1) Part I, show that temperature perturbations are negatively correlated with ozone perturbations, and this adjustment takes place on time scale of less than 20 days (see Figure 2, Part I). In the lower stratosphere, between 10 and 30 hPa, the relation between ozone and temperature is due to the infrared cooling, which takes places on a time scale of about a month. Below 10 hPa, the photochemical lifetime of ozone is so long that it can be considered as a tracer. Interestingly, these correlations clearly show up with the CQC method.

To compute the cross-correlation between two variables, u and v , using the six-hour model differences method (i.e., the CQC method) a number of simplifications of the cross-correlation representation are required. In principle, collecting statistics of six-hour differences at each 6 h time period in a month (for a total of N time periods), the cross-covariance is obtained as:

$$\mathbf{B}_{uv} = \frac{1}{N-1} \sum_k^N \left[(\mathbf{u}_{t(k)+6} - \mathbf{u}_{t(k)}) - \overline{(\mathbf{u}_{t(k)+6} - \mathbf{u}_{t(k)})} \right] \left[(\mathbf{v}_{t(k)+6} - \mathbf{v}_{t(k)}) - \overline{(\mathbf{v}_{t(k)+6} - \mathbf{v}_{t(k)})} \right]^T = \Sigma_u \mathbf{C}_{uv} \Sigma_v^T \quad (13)$$

where Σ_u, Σ_v are diagonal matrices of error standard deviations of the six-hour differences, and the index k is the index of one of the six-hour time period in a month. In principle, $\mathbf{C}_{uv}$ is a full matrix of dimension $\dim(\mathbf{u}) \times \dim(\mathbf{v})$ and needs to be considerably simplified in order to derive it from time

statistics. The cross-correlation is generally represented as a zonal field of point correlations, thus neglecting the horizontal and vertical correlations. It was found that neglecting the vertical correlation has a small impact on the zonal-mean representation of C_{uv} .

The cross-correlation between ozone and temperature computed for the month of July 2003 is shown in Figure 5. The pattern for August and October 2003 are very similar (result not shown). As discussed previously, the region above the 10 hPa is photochemistry-dominated, while below 10 hPa the ozone behaves like a tracer although its radiative effect is important on a time-scale of 20 days to a month. At around 10 hPa the photochemistry time scale is about 10 days and decreases to one day at 3 hPa, and to half a day at 2 hPa. At this altitude the photochemical time-scale decreases with latitude in the northern hemisphere summer (as shown in Figure 2, Part I). We observe in Figure 5, that the maximum anti-correlation between temperature and ozone occurs at about 2 hPa, a region in which six-hour differences are able to capture the photochemical signal of half a day. The maximum anti-correlation is also not centered at the equator, but rather in the northern hemisphere summer due to the asymmetry between hemispheres in the photochemical time-scales (see Figure 2, Part I). We note a weaker but positive correlation between temperature and ozone below 10 hPa. However, this positive correlation is not very different between interactive and non-interactive runs, with the caveat that the interactive run shows a stronger positive correlation in the northern hemisphere summer between 10 and 100 hPa. At those altitudes the radiative time-scale is on the order of 20 days to a few months. The six-hour differences method clearly cannot capture a signal on time-scales of weeks and months, and this is why there is little difference between the interactive and non-interactive runs. The differences between interactive and non-interactive runs in the northern hemisphere between 10 and 100 hPa are slightly larger if instead of six-hour differences we use 24-hour differences to derive the cross-correlation (Figure S8, Supplementary Material).

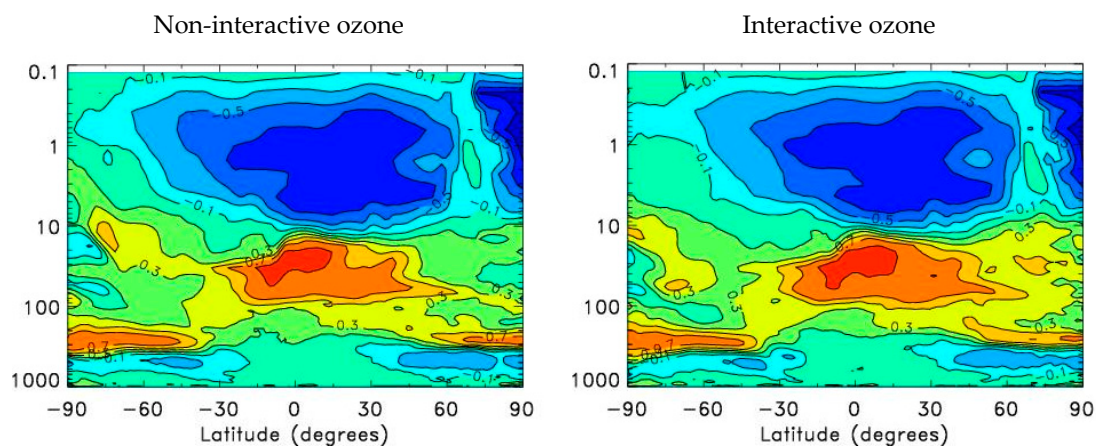


Figure 5. Cross-correlation between ozone and temperature derived from six-hour differences (i.e., CQC method) for July 2003. (**Left panel**) The non-interactive ozone-radiation run of GEM-BACH and (**right panel**) for an interactive ozone-radiation run.

Generally, the positive correlation between temperature and ozone below 10 hPa is not of radiative origin but is due to the impact of short-term (e.g., 6 h) temperature effects on ozone transport. We note for example that large positive correlations are observed near the NH and SH tropopause and in the equatorial region around 20–70 hPa, a region clearly dominated by transport.

Construction of the balance operator F (see Section 2.3) requires the unbalanced component of temperature. However, the unbalanced temperature is not directly accessible from the six-hour model differences, and would require a sequential reprocessing respecting the Gram–Smith orthogonalization of the model differences, which we have not attempted to do here. Instead, we used correlation with temperature as an approximation. We recall that what needs to be computed is $F = B_{O_3T^u} B_{T^uT^u}^{-1}$ but what we have readily available from the statistics is $A = B_{O_3T} B_{TT}^{-1}$ so we end up approximating F by A .

To understand the nature of this approximation, we first note that:

$$\begin{aligned}\mathbf{B}_{TT}\mathbf{A} = \mathbf{B}_{O_3T} &= \mathbf{E}[\delta O_3(\delta T)^T] = \mathbf{E}[\delta O_3(\delta T^u + \mathbf{B}_{T\psi}\mathbf{B}_{\psi\psi}^{-1}\delta\psi)^T] \\ &= \mathbf{E}[\delta O_3(\delta T^u)^T] + \mathbf{E}[\delta O_3(\delta\psi)^T]\mathbf{B}_{\psi\psi}^{-1}\mathbf{B}_{T\psi}^T = \mathbf{B}_{O_3T^u} + \mathbf{B}_{O_3\psi}\mathbf{B}_{\psi\psi}^{-1}\mathbf{B}_{T\psi}^T\end{aligned}\quad (14)$$

where we used Equation (6) and $\mathbf{E}$ denotes the expectation operator. The correlation between O_3 as a tracer and the streamfunction relates to the tracer-wind coupling discussed in Section 2.3 Part I, which has been an elusive coupling relationship to obtain diagnostically [89–91] (see also discussion in Section 7). It has been argued that in regions of Rossby wave breaking, potential vorticity is correlated with ozone in the lower stratosphere where O_3 behaves as a chemical tracer. Figure S7 (Supplementary Material) shows scatter plots of O_3 concentration with streamfunction values between 10 and 100 hPa for March 2003 for different latitude bands. We note, however, that *streamfunction and ozone* have no significant correlation except at the highest latitudes in the Northern Hemisphere. We, thus, make the simplification and approximation that globally, the correlation between O_3 and streamfunction can be neglected. This also implies that the balance operator $\mathbf{G}$ (Equation (8)) can be neglected. Regarding Equation (14) we, thus, make the approximation that:

$$\mathbf{B}_{O_3T} \approx \mathbf{B}_{O_3T^u} \quad (15)$$

Now, concerning the temperature error covariance, it can be calculated by taking the outer-product of δT using the temperature equation from Equation (6) while neglecting $\delta\chi^u$ as in Equation (7), which yields:

$$\mathbf{B}_{TT} = \mathbf{E}[\delta T(\delta T)^T] = \mathbf{E}[(\delta T^u + \mathbf{B}_{T\psi}\mathbf{B}_{\psi\psi}^{-1}\delta\psi)(\delta T^u + \mathbf{B}_{T\psi}\mathbf{B}_{\psi\psi}^{-1}\delta\psi)^T] = \mathbf{B}_{T^uT^u} + \mathbf{B}_{T\psi}\mathbf{B}_{\psi\psi}^{-1}\mathbf{B}_{T\psi}^T \quad (16)$$

In this study, however, for practical reasons, we will use the approximation:

$$\mathbf{B}_{TT}^{-1} \approx \mathbf{B}_{T^uT^u}^{-1} \quad (17)$$

to compute the ozone-temperature balance operator. Thus, with the approximations in Equations (15) and (17) and the limitations imposed by the statistics (as discussed at the beginning of this section) the balance operator $\mathbf{F}$ (Equation (8)) which is a function of latitude and pressure only, is approximated as:

$$\mathbf{F}(\lambda, p) = \mathbf{B}_{O_3T^u}\mathbf{B}_{T^uT^u}^{-1} \approx \mathbf{A}(\lambda, p) = \mathbf{B}_{O_3T}\mathbf{B}_{TT}^{-1} \quad (18)$$

The corresponding ozone error covariance, using the formulation in Equation (6) and taking into account the cross-covariance between ozone and temperature, yields:

$$\mathbf{B}_{O_3O_3} = \mathbf{B}_{O_3^uO_3^u} + \mathbf{F}\mathbf{B}_{T^uT^u}\mathbf{F}^T \approx \mathbf{B}_{O_3^uO_3^u} + \mathbf{A}(\lambda, p)\mathbf{B}_{T^uT^u}\mathbf{A}^T(\lambda, p) \quad (19)$$

To construct the operator $\mathbf{A}$ we used the ozone-temperature cross-correlation coefficients obtained from point-wise statistics derived from the CQC method, the ozone error variance from the HL method depending only on height as described above in Section 3.1, and the temperature error variances from the NMC method with adjustments from innovations [56]. This contrasts with the balance operators introduced by Derber and Bouttier [64] where the regression statistics are derived in spectral space—an approach used for the balance operator between meteorological variables used here in the CMC 3D-Var meteorology. The point-wise statistics used for $\mathbf{A}$ are dependent on latitude and pressure (the model hybrid vertical coordinate to be precise). We have investigated the use of a vertical correlation (but not horizontal correlation) in the operator $\mathbf{A}$ and observed only small differences (results not shown). Figure 6 (left panel) illustrates the cross-covariance thus obtained, which we will denote by $\mathbf{A}^{\text{CQC-NMC}}$.

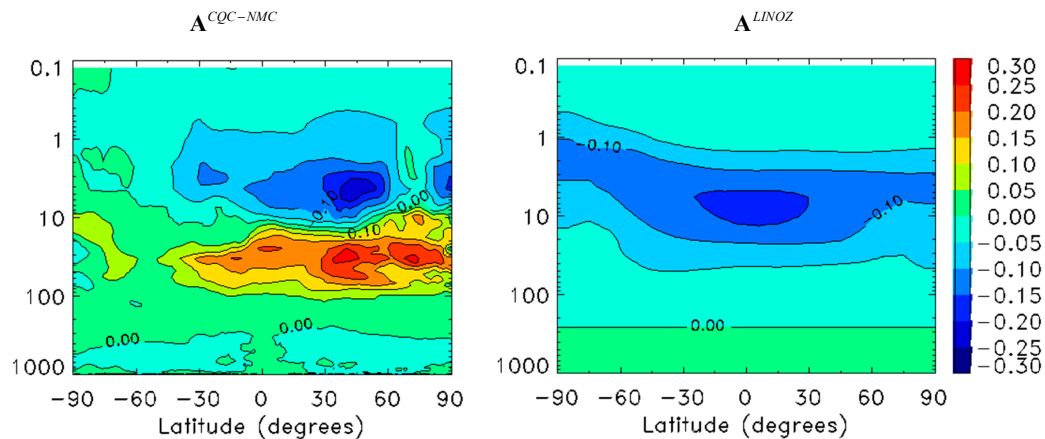


Figure 6. Balance operator between ozone and temperature for July 2003. (Left panel) $A^{CQC-NMC}$, which uses CQC for correlations and the NMC method for temperature error variance, and (right panel) A^{LINOZ} , as derived from the LINOZ scheme (derivation in Appendix C).

We also calculated the cross-covariance obtained using the LINOZ model, which is derived in Appendix C using the stationary solution of the cross-covariance evolution equation between ozone and temperature (Equation (A47)), and which we denote by A^{LINOZ} displayed in the right panel of Figure 6. The most important feature of the cross-covariance of the LINOZ model is that it contains only the effects due to photochemistry (radiative effects are absent). The cross-covariance is negative as we would expect, but in general nearly matches the ozone climatology (as explained in Appendix C), with $A^{LINOZ} \approx -2 \times 10^{-2} \overline{O_3}$.

4. Harmonization of AMSU-A Radiances with MIPAS Temperatures

The microwave sounder AMSU-A (Advanced Microwave Sounding Unit) on board several operational NOAA satellites has been the main source of temperature-sensitive measurements for NWP in the stratosphere (for the time period considered in these experiments). AMSU is a nadir-looking and horizontally-scanning instrument. The coverage of AMSU-A on board NOAA-15 and NOAA-16 during any six-hour window is almost entirely global (Figure S9 Supplementary Material). The horizontal coverage is, in fact, too dense to apply all profiles with horizontally uncorrelated observation errors, and so thinning (i.e., discarding profiles) is usually performed in operational data assimilation (also illustrated in Figure S9). Channels 10–14 are sensitive to stratospheric temperature but have rather coarse vertical resolution (Supplementary Material Figure S10). Limb sounding instruments such as MIPAS are another important source of temperature measurements in the stratosphere. A description of MIPAS and the HALOE instrument on board UARS [2–4] is given in Part I, Section 7.1. For the time period we have considered (i.e., 2003) AMSU-A and MIPAS were the two most important sources of stratospheric temperature measurements, with the exception of radiosondes that rise up to 30 km in tropical regions (and lower altitudes elsewhere).

The main issue with AMSU-A radiances is that the geolocated and calibrated radiances (i.e., level 1B) need to be bias-corrected and this is usually done by using the meteorological model short-term forecast as an “unbiased” estimate. This procedure is well adapted in the troposphere where other unbiased observations have a significant effect on the analysis, thus, by comparing model-simulated radiances with observed radiances can be used effectively to separate model bias from observational bias. Such observations are often referred to as “anchor” observations in a bias correction scheme. Observation bias-correction schemes can be either static or online with the analysis, as in the Variation Bias Correction scheme [92]. However, it is found that the application of bias correction in the upper stratosphere is problematic in the absence of “anchor” observations [93]. DiTomaso and Bormann [93] have proposed assimilating AMSU-A channel 14 without any bias correction as a way to anchor the meteorological analysis in the mid to upper stratosphere. Here, we propose another approach, which

consists of assimilating MIPAS temperature observations to anchor the stratospheric analysis and derive from it a new set of AMSU-A bias corrected radiances. This also has the effect of harmonizing these two sets of observations.

MIPAS-retrieved temperatures in the stratosphere are considered to be of good quality and compares well with HALOE temperatures (see Part I, Section 7.2.1). We, thus, conducted an assimilation of MIPAS temperature observations without AMSU-A (stratospheric) channel 10–14 as an “anchor” run. To generate this assimilation run, we used as observation error for MIPAS temperatures the estimates obtained from the HL method as described in Section 3.1, and for the meteorological error statistics a combination of innovation variance consistency with the NMC method as summarized in Table 1. From this anchor run, a new set of bias correction coefficients was obtained, as well as a new set of AMSU-A radiances with a bias correction based on MIPAS temperature.

The results are compared for 12–31 August 2003 in Figure 7. Radiance innovations based on AMSU-A stratospheric channels and using the standard bias correction used at CMC are shown in blue, and using only the model in the stratosphere and the new bias correction using an assimilation of MIPAS temperatures are shown in red. This evaluation was also conducted over other time periods; 14–31 January, and 12–18 October 2003 with similar results (not shown).

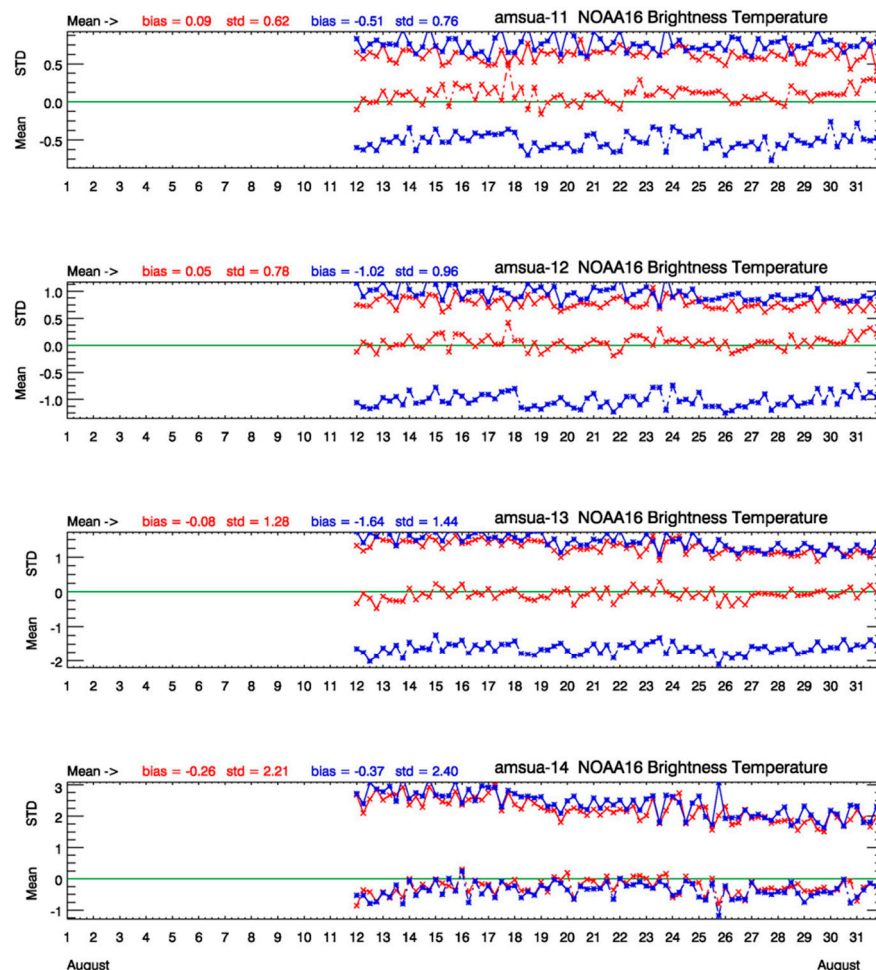


Figure 7. Mean (lower curves) and standard deviation (upper curves) of the AMSU-A radiance observations minus the forecast (6 h) for channels 11 to 14 (**upper panel to lower panel**). In blue are the results using the standard CMC bias correction scheme, which uses only the model in the stratosphere, and in red using only MIPAS temperature in the stratosphere.

We observe a net reduction in radiance bias for channels 11–14 with the new bias correction based on MIPAS temperatures, with a slight reduction in the standard deviation. The mean analysis increment at 10 hPa are presented in Figure S11 (Supplementary Material) for September 2003 and zonal mean analysis increments in Figure S12. These results indicate a significant reduction in the mean analysis increment everywhere except in the polar regions in the upper levels of the model (1 hPa and higher), which may be due to the model pole problem or the sponge layer. Because of these results, all following assimilation experiments were conducted using the new AMSU-A bias correction based on the assimilation of MIPAS temperatures.

5. The Added Value of the Assimilation of Limb Sounding (MIPAS) Temperatures

Let us first examine the benefit of assimilating MIPAS temperatures in addition to AMSU-A radiances, with the new bias correction (Section 4). An assimilation from 17 August to 30 September 2003, was conducted and the global verification results are presented in Figure 8. In blue is the assimilation of AMSU-A only, and in red the assimilation of MIPAS temperature and AMSU-A.

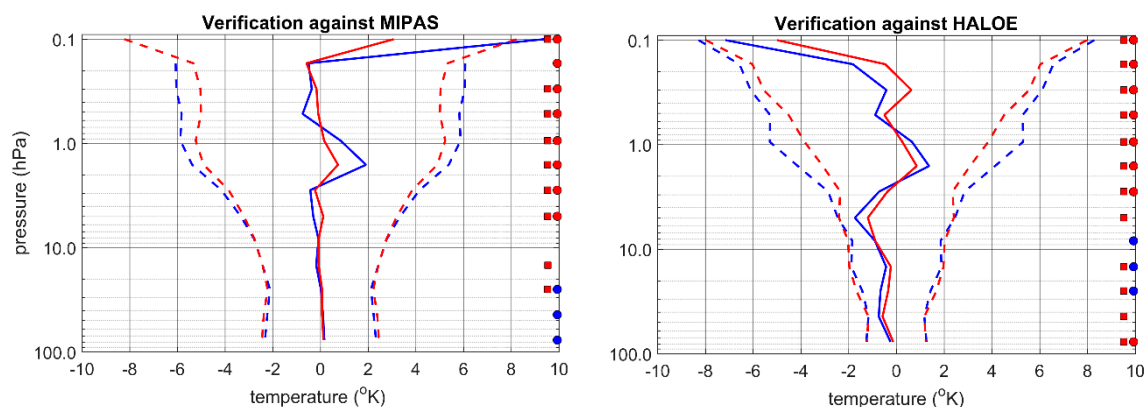


Figure 8. Global verification (observation-minus-forecast) of temperatures as function of height (in hPa) for two assimilation runs. In blue is the assimilation of AMSU-A only, and in red the assimilation of MIPAS temperature and AMSU-A. All AMSU-A data are processed with the new bias correction. The left panel illustrates verification against MIPAS temperatures, and (right panel) verification against HALOE temperatures. The squares on the far right of the panels indicate a significant Student's *t*-test with 95% (or higher) confidence interval, and the dots on the far right of the panels indicate a significant the Fisher test of variances with a 95% (or higher) confidence interval. These markers are red (blue) if the red (blue) experiment shows an improvement.

We observed an improved bias and a reduction in error variances in the mid to upper stratosphere (from 10 hPa to 0.3 hPa) with the combined assimilation of MIPAS and AMSU-A, whether the verification is performed against MIPAS or HALOE as independent observations. The larger impact in the mid- to upper stratosphere may be due to having more AMSU-A channels sensitive to the lower stratosphere, or that the limb sounding observations (thus vertical observation information) provided by MIPAS have a definite advantage in the mid- to upper-stratosphere where only one channel of AMSU-A (i.e., channel 14) provides information (note that there is about 22,000 thinned AMSU-A profiles and 300 MIPAS profiles in 6 h). To address this question we have performed an assimilation of AMSU-A (blue) only versus MIPAS (red) only, with results presented in Figure 9.

Verification against HALOE temperatures (Figure 9) shows very little difference with respect to the combined assimilation results (right panel of Figure 8), but are more pronounced in the lower stratosphere. Similar results for individual latitude regions were found in both experiments (results not shown). Thus, we see the importance of height resolving observations such as MIPAS in the stratosphere.

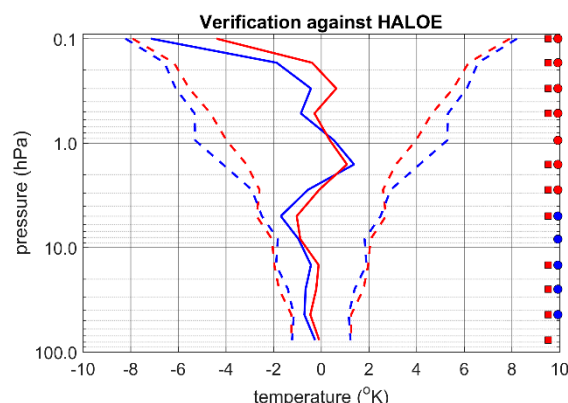


Figure 9. Global verification (observation-minus-forecast) against HALOE temperatures for two assimilation runs. In blue is the assimilation of AMSU-A, and in red is the assimilation of MIPAS temperatures only. Tests of significant differences are the same as in Figure 8.

Next, we conducted another set of experiments that directly illustrated the impact of the assimilation of limb-sounding temperature observations on model temperature and on transport of ozone. In this set of experiments, and contrary to the results presented in Figures 8 and 9, we activate ozone–radiation interaction in the model, i.e., having transported ozone affecting radiation instead of a (prescribed) ozone climatology. But as we shall see in the following section (Section 6), the ozone–radiation interaction has very little impact on the verification of six-hour forecasts. The impact actually develops over a time period of several days, so that for all practical purposes we can consider the following results to be essentially independent of the presence of ozone–radiation interaction.

To better illustrate the impact of limb sounding observations, we conducted a meteorological assimilation of MIPAS temperatures where stratospheric AMSU-A channels (11–14) are excluded (in red) and compared it with an assimilation of MIPAS temperatures where all AMSU-A channels are retained (in black). The new AMSU-A bias correction scheme was applied in both cases.

Figure 10 displays the global verification results of assimilation runs from 17 August to 31 October 2003. In general, for the mid- and upper-stratosphere, both in terms of bias and error standard deviations, the assimilation of MIPAS data with no stratospheric channels of AMSU-A performs better than assimilation using all stratospheric channels. This conclusion is valid whether the verification is made against MIPAS temperatures (left panel) or against independent temperature measurements from HALOE (right panel). This positive impact is also seen in temperature forecasts but gradually disappears over a forecast period of 10 days (see Figure S13 in Supplementary Material). Note that when the statistic exceeds the range of value potted, the line that connect this point to another point below is not plotted—that is why, for example, the blue line connecting 0.16 hPa to 0.1 hPa on the right panel is not shown.

For the same set of experiments, the impact on ozone is illustrated in Figure 11. We observe a reduction in the standard deviations of observation-minus-forecast (6 h) error in the lower and upper stratosphere whether it is verified against MIPAS ozone (left panel) or HALOE ozone (right panel).

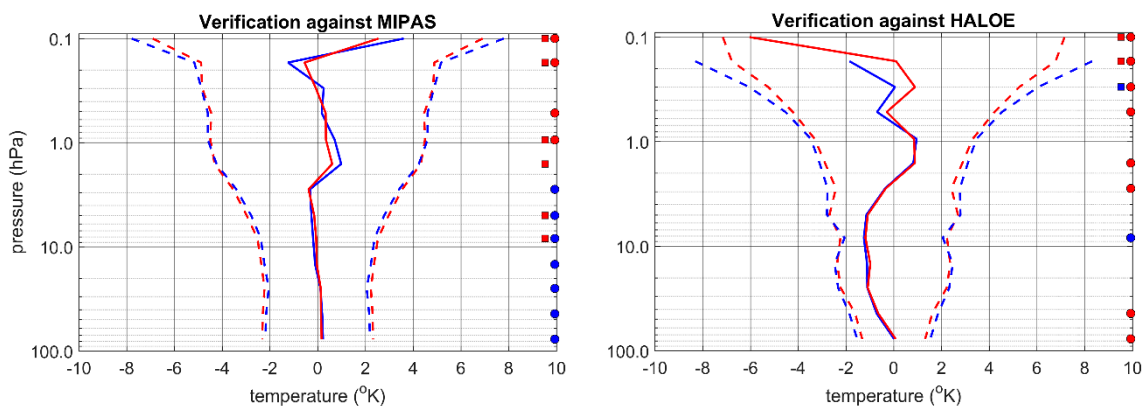


Figure 10. Global verification (observation-minus-forecast) of temperature as for two assimilation runs. In red, is the assimilation of MIPAS temperature and AMSU-A but no stratospheric channels, and in blue is the assimilation of MIPAS temperatures with all the AMSU-A channels. (**Left panel**) The verification against MIPAS temperatures, and (**right panel**) the verification against HALOE temperatures. Tests of significant differences are the same as in Figure 8.

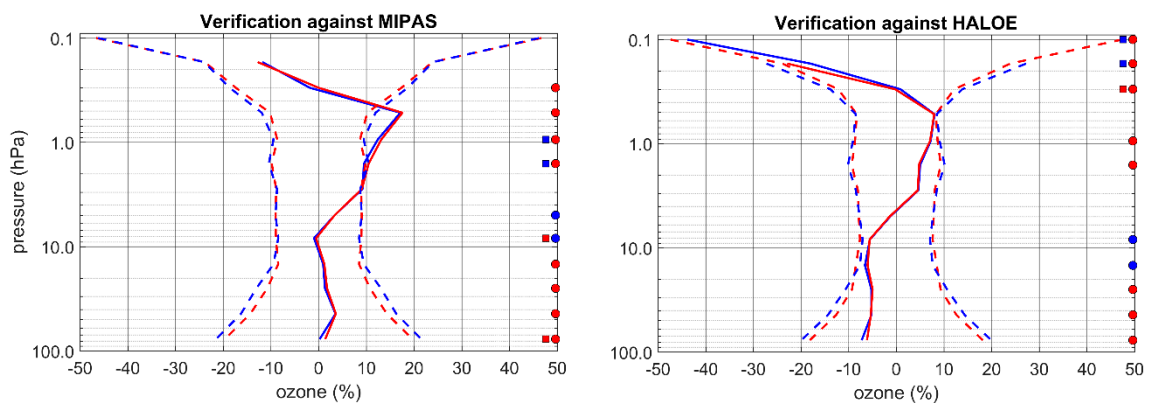


Figure 11. Same as Figure 10, but for verification of ozone MIPAS on the left and ozone HALOE on the right.

The reduction of random errors is remarkably larger in the lower stratosphere (below 20 hPa) is where transport and the vertical gradient of ozone are important. A larger reduction in standard deviations is observed over Antarctica (results not shown). A reduction in the error standard deviations is also observed for CH₄ above 3 hPa. Thus, we see that the presence of AMSU-A temperatures in the assimilation actually degrades the vertical structure, because of the coarse vertical resolution sensitivity of the associated channels, which is apparent in the transport of chemical species in regions of strong vertical concentrations.

6. Weak Coupling Assimilation due to Ozone–Radiation Interaction

We know that the ozone–radiative interaction time-scale varies from about a week at 1 hPa to about a month at 25 hPa, while the ozone photochemical time-scale is a few hours at 1 hPa and is on the order of three months at 25 hPa (see Part I, Section 2.1). At about 10 hPa, these two interactions have comparable time-scales, i.e., about two weeks. This implies that the assimilation of ozone will have little impact on temperatures above 10 hPa, but the impact, which is radiative in nature, will be noticeable in the lower stratosphere and will build up slowly over time. The assimilation of temperature on the other hand will influence the photochemistry of ozone above 10 hPa, and will influence ozone transport in the lower stratosphere.

To examine these effects in the context of assimilation we will focus on assimilating only limb sounding observations. As stated in Section 5, the assimilation of limb sounding temperatures while

excluding stratospheric AMSU-A channels has a stronger impact on both temperature and ozone transport than using the stratospheric AMSU-A channels, which tends to spread out the temperature information vertically.

An assimilation of MIPAS temperatures without stratospheric AMSU-A channels (i.e., using only channels 1–8) was performed for the period of 17 August–5 September 2003. The global verification of observation-minus-forecast (6 h) temperatures and ozone is presented in Figure 12.

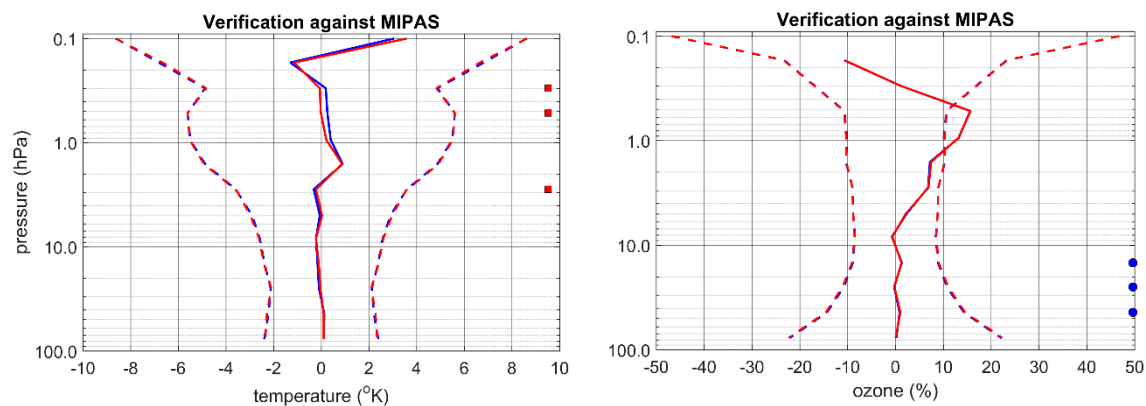


Figure 12. Global verification of the impact of ozone radiation interaction. All runs with the assimilation of MIPAS temperatures and AMSU-A channels 1–8 only. (**Left panel**) The temperature impact, (**right panel**) and the ozone impact. Runs with no ozone–radiation interaction are in blue, and with ozone–radiation interaction in red. Test of significant differences of statistics are the same as in Figure 8.

Red curves correspond to assimilation with the GEM-BACH model with ozone–radiation interaction activated while black curves correspond to an experiment where the radiation is computed from a monthly ozone climatology, not the transported ozone. We note that in these temperature-only assimilation experiments, ozone–radiation interaction creates very little change in the temperature and ozone analyses (or 6 h forecasts) except for small differences in the upper-stratospheric mean temperature and the variance of lower stratospheric ozone (most of the blue curves lie under the red curves).

The small mean difference in temperature between the two experiments around 5 hPa and above can be explained by the fact the GEM-BACH model has an ozone deficit of 15% at those altitudes (as suggested by the right panel of Figure 12, and discussed in Section 7.2.3, Part I). Thus, with the interactive model, the lower model ozone concentrations leads to cooler temperatures, which the assimilation of temperature can only partially correct since it is a systematic error.

In the lower stratosphere, the O-P variances are increased in the case of ozone–radiation interaction. We recall that there is no assimilation of ozone in these experiments, and the impact on ozone can be understood by considering ozone as an unobserved variable as defined in Section 2.1. The impact on unobserved variables can be computed from the cross-variable increments, Equation (5), and here in particular, the balance operator **A** between ozone and temperature. The associated background (or model) error covariance is given by Equation (19) and using the operator **A**. We have shown already in Figure 5 that ozone–radiation interaction increases the correlation between temperature and ozone between 10 and 100 hPa (in the northern latitude summer). Consequently, the error cross-covariance and its effect on the error variance of ozone is increased, and this is what it is observed in the right panel of Figure 12.

Although the impact of ozone-interaction is nearly absent in analyses (or six-hour forecasts), it gradually accumulates in forecasts. De Grandpré et al. [42] have reported results of assimilation of temperature and ozone on the temperature predictability using the GEM-BACH with essentially the same experimental setup has discussed here. A gradual increase in the anomaly correlation for the period of 11 August–5 September 2003 was shown reaching nearly half a day in the lower

stratosphere as a result of ozone–radiation interaction (either with assimilation of temperatures only or with assimilation of temperature and ozone). Here we show anomaly correlation results in which the assimilation of MIPAS temperature was conducted over a longer time period from 15 August to 5 October 2003 that essentially corroborate the published results. For a description of the calculation of the anomaly correlation (i.e., the correlation between the forecast and analysis valid at the same time) we refer the reader to de Grandpré et al. [42].

The above improvement comes from a better representation of ozone radiative heating in the lower stratosphere region. This radiative forcing persists throughout the forecast period due to the long photochemical lifetime of ozone which is much longer than the radiative time-scale in this region.

The precise chemistry model used has in fact little impact on these results. To show this we have conducted a similar ozone–radiation interaction assimilation experiment using a linearized chemistry model LINOZ [60] for daily mean values (with semi-Lagrangian transport [61]):

$$\frac{dO_3}{dt} = c_1 + c_2(O_3 - \overline{O_3}) + c_3(T - \overline{T}) + c_4(O_3^\uparrow - \overline{O_3^\uparrow}) \quad (20)$$

where the coefficients c_1, c_2, c_3, c_4 are determined using a chemical box model, the overbar $\overline{(\)}$ represents climatological values, and $\uparrow$ represents the overhead column. The coefficient $c_2 = 1/\tau_{O_3}$ is related to the photochemical time-scale of ozone (see also Section 2.1, Part I). Figures 13 and 14 show the impact of assimilating temperature and ozone with ozone–radiation interaction activated using the LINOZ ozone model and the BASCOE chemistry model.

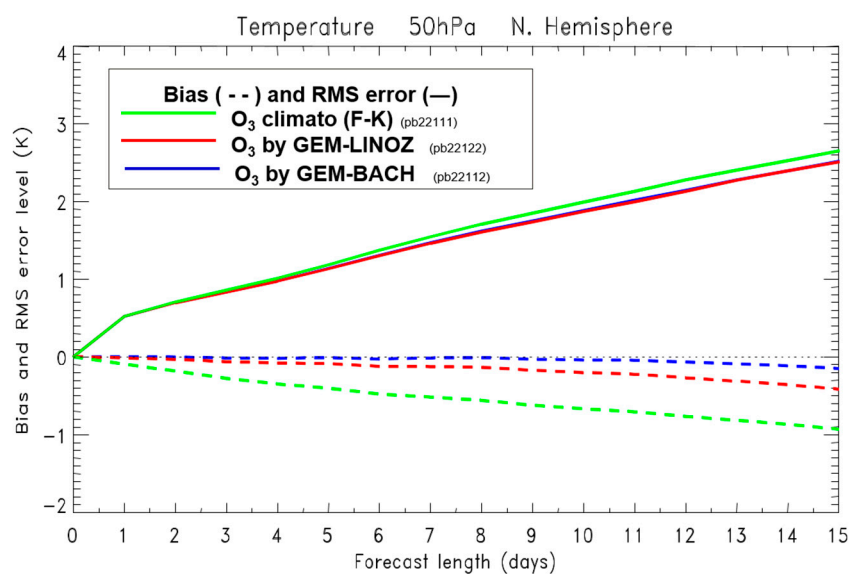


Figure 13. Fifteen-day forecast of temperatures at 70 hPa verified against analyses over the South Pole region resulting from the assimilation of MIPAS temperature and ozone. The curves in blue are from using the BASCOE chemistry, in red using the LINOZ linearized ozone chemistry, and in green using a climatological ozone. Note that, here, the bias is depicted with dash lines, while the solid lines represent the root mean square (RMS) error (not the standard deviation).

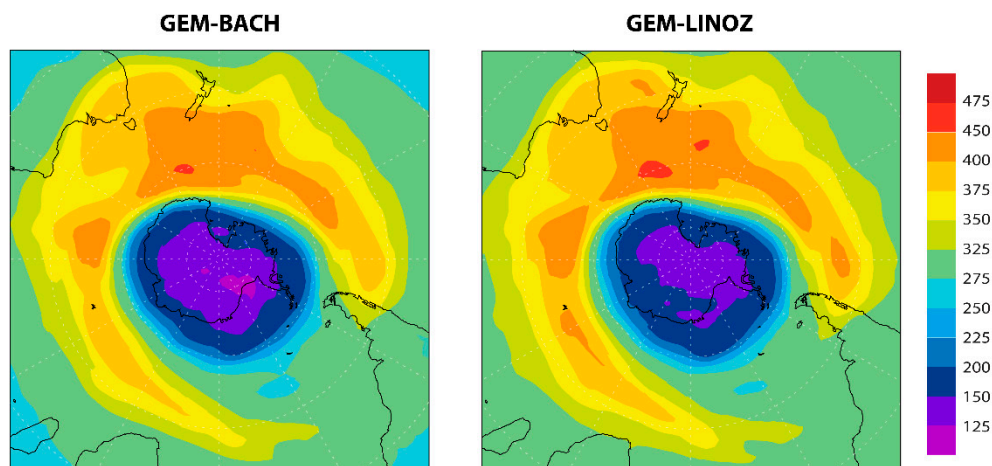


Figure 14. Total column ozone (DU) (analysis) over the South Pole region on 3 October 2003 resulting from the assimilation of MIPAS temperature and ozone. **(Left panel)** Experiment using the BASCOE chemistry scheme. **(Right panel)** Experiment using the LINOZ linearized ozone chemistry scheme.

We conclude from these figures that the analysis and time evolution of ozone over the South Pole region with GEM-BACH ozone–radiation interaction are similar whether we use the comprehensive (BASCOE) chemistry or the linearized (LINOZ) chemistry. Figure 15 shows the temperature forecast, bias, and the root mean square (RMS) error at 50 hPa over the Northern Hemisphere in comparison with MIPAS temperature analyses.

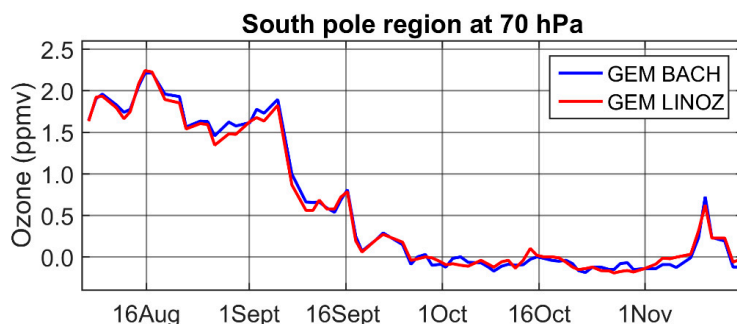


Figure 15. Time series of ozone at 70 hPa over the South Pole region resulting from the assimilation of MIPAS temperature and ozone. The blue curve is from using the BASCOE chemistry scheme, and the red curve using the LINOZ linearized ozone chemistry scheme.

Although there is a drift in the temperature forecast in all experiments, we note that the interactive runs using BASCOE and LINOZ chemistry both exhibit relatively slow growth of temperature random errors, while a faster growth of errors is seen from using the ozone climatology. This result is coherent with the anomaly correlation results presented in Figure 16 (and [42]) which indicate greater forecast skill with ozone–radiation interaction than using climatological ozone. The results illustrated in Figure 15, also suggest that an anomaly correlations computed using the LINOZ chemistry should lead to improvements over the climatological ozone run.

Thus, we conclude that weak coupling due to ozone–radiation interaction does not change significantly the analysis. However, it has an effect on the forecast skill that is observed with using either the full chemistry or a simplified (linearized) ozone chemistry schemes.

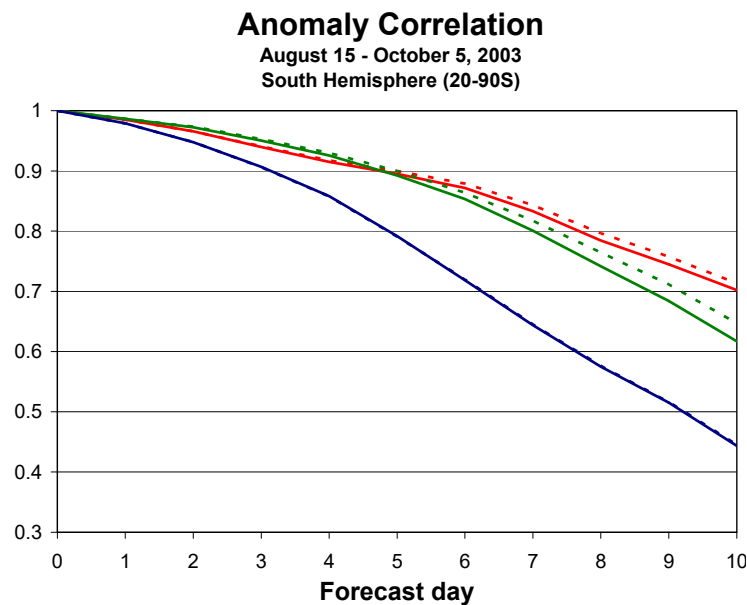


Figure 16. Anomaly correlations at 10 (red), 50 (green) and 100 (purple) hPa in the southern hemisphere (20–90S) for ozone-radiation interactive (dashed lines) and non-interactive ozone (solid lines) experiments.

7. Strongly Coupled Temperature–Ozone Assimilation with 3D-Var

The 3D-Var-Chem developed in this study allows for cross-covariances between meteorological and chemical variables and between chemical variables themselves. To examine the effect of adding cross-covariances between temperature and ozone in the context of 3D-Var, we have conducted experiments using the balance operators $\mathbf{A}^{CQC-NMC}$ and $\mathbf{A}^{LINOZ}$ described in Section 3.3.

We have conducted three assimilation experiments using MIPAS O₃ and AMSU-A temperatures (all channels) for a period of two weeks, from 17 August to 4 September 2003. Figure 17 shows the verification over the globe in the three case: univariate (blue), multivariate with the balance operator $\mathbf{A}^{CQC-NMC}$ (red), and multivariate with the LINOZ balance operator $\mathbf{A}^{LINOZ}$ (grey dots).

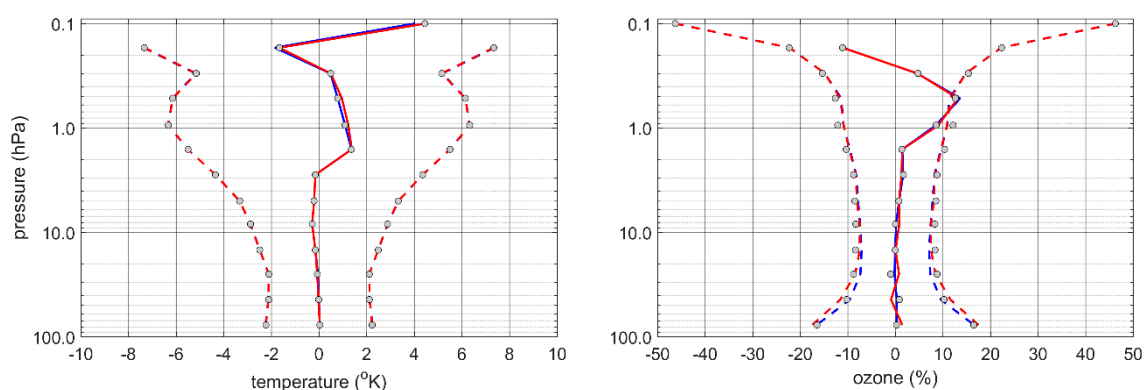


Figure 17. Multivariate temperature-ozone assimilation. Univariate ozone and temperature assimilation (blue), multivariate assimilation performed with the CQC-NMC balance $\mathbf{A}^{CQC-NMC}$ (red) and with LINOZ balance $\mathbf{A}^{LINOZ}$ (grey dots). The solid lines denote average differences (biases) and the dashed lines indicate the standard deviations (by $\pm\sigma$), except for $\mathbf{A}^{LINOZ}$ where we use only grey dots for both metrics. (**Left panel**) The temperature O-P (observation minus six-hour forecast) statistics from comparisons to MIPAS observations. (**Right panel**) The ozone O-P statistics. Note that to simplify we have not presented the significant tests of differences of statistics because it would have been cumbersome to illustrate with three pairs of experiments.

We note that in general there is little change between all three experiments, indicating no advantage in using multivariate cross-covariances between temperature and ozone in a 3D-Var assimilation system. The exceptions are for the upper stratosphere temperature where the LINOZ operator reduces slightly the temperature bias (although not significantly), and for ozone with an increase of variances for both the LINOZ and CQC-NMC operators in the lower stratosphere.

These results can be explained by the fact that, as we showed in Section 6, the ozone–radiation interaction increases the error cross-covariance in the region between 10 and 100 hPa, while above 2 hPa the photochemical time-scale of ozone is so short that any adjustment due to the analysis is lost in a six-hour time period.

Additionally, the lack of significant impact of using balance operators suggest that the modeling of the balance operators may not be adequate. For the CQC-NMC balance operator the assumption of using the temperature instead of the unbalanced temperature, i.e., Equation (7), in the CQC-NMC may have a detrimental effect. The error correlations below 10 hPa (see Figure 5) are dominated by transport—thus contain the balanced temperature. For the LINOZ balance operator, there is no account for radiative effects which would influence the mid- and lower-stratosphere where the analysis of ozone would have a persistent effect (contrary to the upper stratosphere where the model relaxation time for ozone is very short). Although we have not continued these experiments further, it seems necessary to re-examine the balance operators between ozone and temperature to better account for ozone–radiation interaction: (1) for the CQC-NMC operator by using unbalanced temperature to truly isolate ozone–radiation from transport in the lower stratosphere; and (2) to add an ozone–radiation interaction of the simplified ozone modeling, i.e., LINOZ operator.

8. Strongly Coupled Tracer–Meteorology Assimilation with 4D-Var

A second strongly coupled data assimilation experiments was conducted by using chemical tracer observations to infer information about the winds. We should first note that the information about winds inferred from tracers can either be mechanistic or statistical in nature. The evolution of quantities transported by the atmospheric flow field contains implicit information about the underlying winds. This is the basis for a mechanistic inference. As 4D-Var considers a time series of observations, it extracts wind information from time series of quantities like humidity and passive tracer concentrations [44,93,94]. On the other hand, Daley [43] has alluded to the fact that the spatial variation of error variances can also provide information about the winds (this is related to statistical inference). To understand how this works, let us consider a steady state example of a two-dimensional non-divergent flow. In steady state the streamfunction is identical to the trajectories or streamlines.

We recall that streamlines $\mathbf{X}$ are solutions of:

$$\frac{d\mathbf{X}}{dt} = \mathbf{V}(x, t) \quad (21)$$

where $\mathbf{V}$ is the horizontal velocity vector at coordinate x and time t . $\mathbf{X}$ is the Lagrangian solution of the flow, and since a tracer is a Lagrangian-conserved quantity, the concentration of a chemical tracer c depends only on $\mathbf{X}$, i.e., $c = c(\mathbf{X})$. On the other hand, a non-divergent flow can be described entirely by a streamfunction ψ :

$$u = \frac{\partial \psi}{\partial y}; \quad v = -\frac{\partial \psi}{\partial x} \quad (22)$$

where $\mathbf{V} = (u, v)$, such that $\nabla \cdot \mathbf{V} = 0$. However, since streamfunctions also have the property that $\mathbf{V} \cdot \nabla \psi = 0$, in a steady-state case where $\partial_t \psi = 0$, the material derivative of ψ is zero. In this case, the streamfunction ψ is constant following the material particles, and, thus, the streamline and streamfunction coincides, and one could, thus, use the streamfunction as a proxy for the concentrations.

The cross-covariance between the streamfunction and the wind is obtained from:

$$\langle u\psi \rangle = \frac{1}{2} \frac{\partial}{\partial y} \langle \psi^2 \rangle; \quad \langle v\psi \rangle = -\frac{1}{2} \frac{\partial}{\partial x} \langle \psi^2 \rangle \quad (23)$$

and, thus, is clearly dependent on the spatial variation of the streamfunction error variance.

Given our steady state assumption with non-divergent winds, the streamfunction and the tracer concentrations are related through the Lagrangian coordinate $\mathbf{X}$. From a statistical point of view, the cross-covariance $\langle uc \rangle$, $\langle vc \rangle$ between wind and the concentration plays a fundamental role in our ability to infer information about wind from concentration. If these cross-covariances are zero, statistical inference is not possible. Thus, we can see that statistical inference of winds from tracer in a steady-state non-divergent flow depends on gradients of concentration error variance.

The above argument stresses the importance of having correct error statistics to be able to infer correct winds. In preliminary experiments using the old error statistics (Section 3.1) with the assimilation of MIPAS methane data in 4D-Var, the impact on the wind increments was small. It was noticed that the weight given to these observations was small. The observation and background error statistics of Polavarapu [63] were reevaluated using the HL method described in Section 3.1 and this experiment was repeated with the new error statistics in order to examine the sensitivity to changes in the error statistics.

The 4D-Var assimilation of MIPAS methane data with the old error statistics resulted in the wind analysis increment shown in Figure 18 (left panel), while Figure 18 (right panel) shows the equivalent from an experiment that used the revised background error statistics for chemical species. Note that the wind magnitude, plotted as contours, is more intense with the HL (i.e., new) statistics than with the old (first-guess) statistics, although mechanistically there is no difference between the two cases, since the initial concentrations and the wind trajectories are the same in both cases. The results are shown at 100 hPa, a level where methane induces the most significant wind corrections. The north-south concentration gradient of methane is weak at 100 hPa which gives small wind correction, but in the Southern Hemisphere the gradient is larger creating significant wind corrections. One also has to keep in mind that the wind analysis increments shown in this figure are limited to regions where observations are available, and depend on the concentration analysis increments themselves.

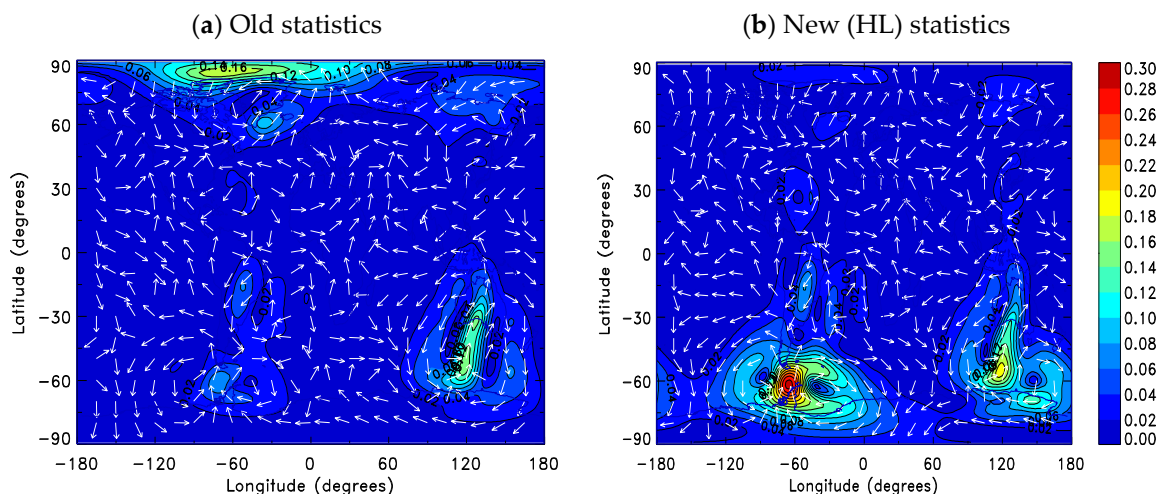


Figure 18. Wind analysis increments in response to MIPAS CH_4 observations obtained with (a) the first estimate of background-error statistics for chemical constituents, and (b) the new statistics estimated using the Hollingsworth–Lönnberg method. Contours are the wind amplitude in m/s and arrows indicate the wind direction. Results are shown here at the 100-hPa level.

Next, a set of experiments was carried out using the new HL statistics where we produced wind analysis increments generated by assimilating individually the three species O_3 , CH_4 , N_2O , and all three together. The results shown at 10 hPa in Figure 19 indicate the additive nature of the wind increments as the three species lead to different impacts at different locations. Analysis wind increments obtained at 50 and 100 hPa are displayed in Figures S16 and S17 (Supplementary Material). The differences in the increments can be explained “mechanistically” by differences in the distribution

of the constituents at different levels. Figure S18 (Supplementary Material) shows that the distribution of N_2O is more homogenized than that of O_3 at 100 hPa. Ozone, generated in the tropical lower stratosphere, is transported in the Southern Hemisphere on a relatively short time scale. Gradients in the ozone field are more important than the gradient of N_2O , and, thus, provide more information about the underlying winds. When observations are present, the presence of these gradients yields the most significant wind increments. Nitrous oxide observations (N_2O) are also involved, but the weaker wind gradients in this field make it more difficult to accurately locate the displacement, which contains the wind information.

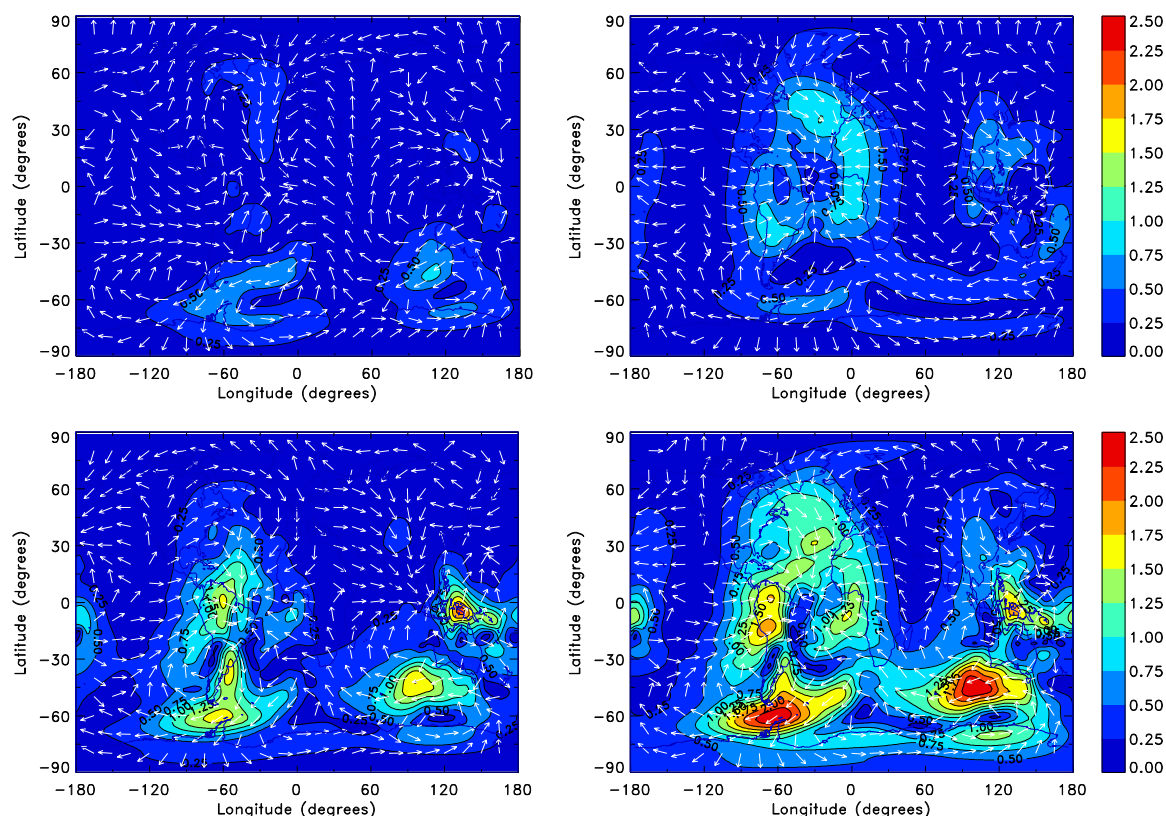


Figure 19. Wind analysis increments at 10 hPa obtained by assimilating CH_4 (top left), O_3 (top right), N_2O (bottom left), and all three species (bottom right). Contours are the wind amplitude in m/s and arrows indicate the wind direction.

The above results indicate that the assimilation of ozone, methane and nitrous oxide yields significant wind increments. A validation of the winds was performed by comparing it with wind measurements from radiosondes in the troposphere. The results shown in Figure 20 for the zonal wind component indicate a reduction in the zonal wind bias all the way to the free troposphere. The results shown here are based on assimilation cycles covering the period 15 August to 5 October 2003, over which the results were then averaged. The impact on the meridional wind are, however, negligible (result not shown).

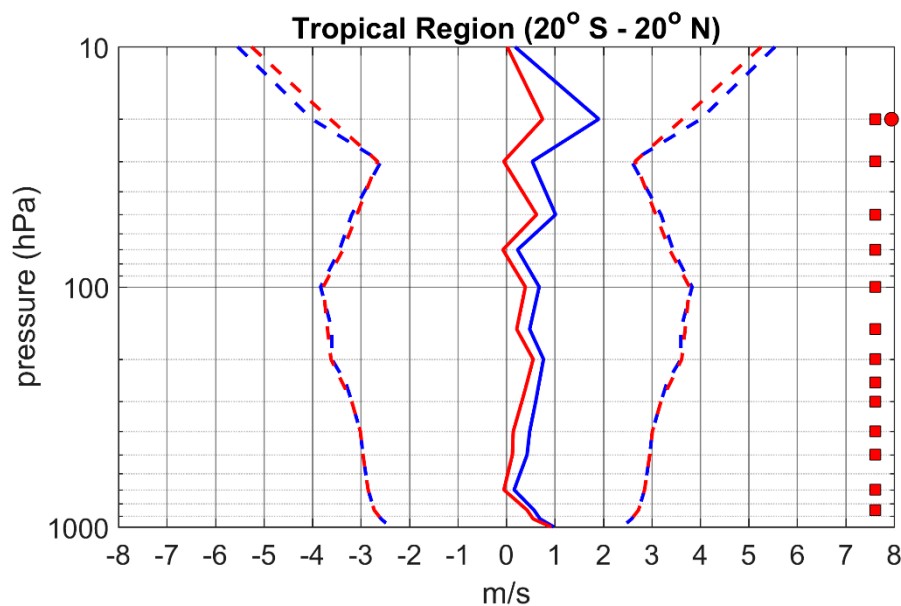


Figure 20. Impact of assimilating stratospheric chemical tracers on tropospheric tropical zonal winds. Verification against radiosonde observations over the tropical region (20° S– 20° N) of observation minus 6-h forecast for the period 15 August to 5 October 2003. The results in red correspond to a 4D-Var assimilation experiment with assimilation of ozone, methane, and nitrous oxide. Results in blue are 4D-Var experiments but without assimilation of the long-lived species. The squares on the far right of the figure indicate a significant Student t-test of means with 95% confidence interval, and the dots indicate a significant the Fisher test of variances with a 95% confidence interval. These markers are red, indicating an improvement.

The difference between the 4D-Var wind magnitude with and without the assimilation of passive chemical tracers (i.e., no chemistry) is shown in Figure 21. We note that the wind magnitude correction in the tropical troposphere and lower stratosphere is about 0.5 ms^{-1} to 3 ms^{-1} agrees, for the most part, with the zonal wind correction in Figure 20, except near 20 hPa where the difference between 4D-Var experiments is about 3 ms^{-1} while the difference with radiosonde observations indicate a correction of about 1 ms^{-1} . We also note large mid-latitude corrections especially in the Southern Hemisphere just outside the polar vortex, in the surf-zone, i.e., the region of Rossby wave breaking.

Higher up near 2–3 hPa and near the equator, there is a large and possibly suspicious wind increment. The origin of this large difference is not clear. For instance, we have noted that lower down at about 10 hPa (and exactly at the equator) the model dynamics has developed very large vertical diffusion coefficients due to the non-staggered vertical discretization. However, it is not clear if the two are related, the maximum wind magnitude correction is not exactly centered at the equator nor is it at the same altitude. These large wind magnitude corrections in the Tropics (at 3 hPa and at 20 hPa) could also be related to the treatment of the background error covariance in the tropical region, which assumes that wind and temperature analyses become univariate close to the equator. Having no wind observations (above 30 hPa) and since the temperature increments are not balanced the wind increments, whereas the winds increments are created only by tracer observations, it may result in inconsistent analyses between wind and temperatures.

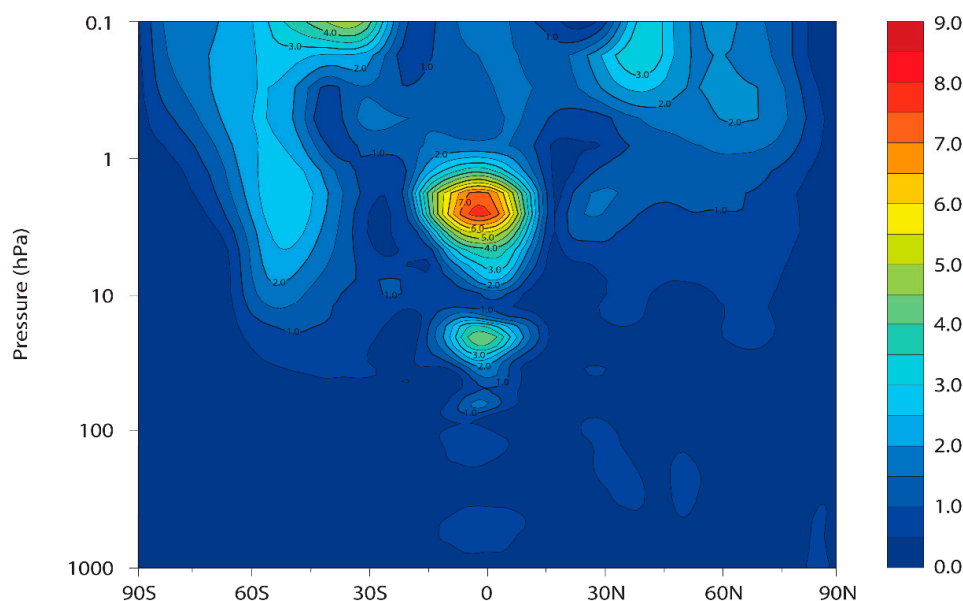


Figure 21. Difference between the (analysis) wind magnitude (in m/s) obtained from two 4D-Var assimilation cycles executed with and without the assimilation of ozone, methane and nitrous oxide. The results are averaged over the period 15 August to 5 October 2003. The zonal mean average is shown here.

We also observe small changes in the temperatures through the lower stratosphere below 10 hPa where constituents are assimilated. Slight differences appear in the tropics but also in the Northern Hemisphere lower stratosphere. For example, Figure 22 depicts the O-P temperature time series at 20 hPa in the Northern Hemisphere and indicates a buildup of systematic differences between a 3D-Var and 4D-Var+tracer assimilation systems throughout the period.

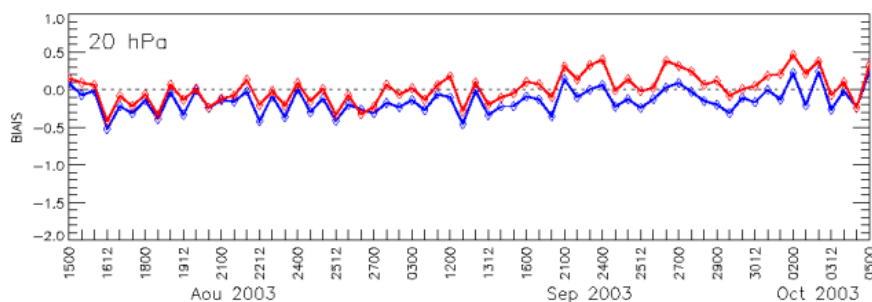


Figure 22. OmP temperature time series between the radiosondes and the 3D-Var (blue) and 4D-var (red) assimilation cycles at 20 hPa in the North Hemisphere.

The impact of the wind correction on the transport of chemical species (e.g., O_3 , CH_4 , N_2O) is a second-order effect and is more difficult to assess, as the 4D-Var assimilated those constituents to produce a wind correction. For ozone, differences appear mainly in the winter hemisphere (Southern Hemisphere) where dynamical processes are more important. Figure S19 (Supplementary Material) shows the comparison of ozone from the assimilation against MIPAS O_3 observations for the period 20 September to 5 October. Below 10 hPa in the Southern Hemisphere mid-latitudes and polar regions, the results of 4D-Var shows a slight degradation at about 100 hPa. In the case of methane and nitrous oxide, differences between both analyses appear in the tropics and are mainly driven by changes in the zonal wind (not shown).

Finally we may comment on the fact the wind corrections are mainly zonal, and not meridional. This can be attributed to zonal structure of the background error variance. Since there is no zonal

variation of the background error variance (only meridional), according to Equation (21) the statistical inference of the concentration observations can only infer information on the u (zonal) component of the wind, not v . It is only through mechanistic inference that information on the meridional component of the wind can be obtained, but the meridional wind component is generally small, giving rise to little corrections. Having more realistic error statistics (i.e., varying also in the zonal direction) may give rise to a more accurate wind inference.

9. Summary and Conclusions

We investigated the issues and particularities of coupled meteorology–chemistry data assimilation in the context of the stratosphere where there is an abundance of vertically-resolved observations and we performed a number of weak and strong coupling data assimilation experiments.

One of the key issues in these assimilation problems is the difference between the large number of prognostic model variables compared and the number of observed variables. In a variational data assimilation formulation, the following question arises: “Do we make an approximation with a J_B term containing only observed variables, or should we consider the full state vector?”. In the case of the coupled model considered here (i.e., GEM-BACH) where there are 61 prognostic variables and 10 observed variables. We showed, through the preconditioning of the variational minimization problem, that there is a split between observed and unobserved variables. The minimization of the cost function can be carried out involving only the observed variables in the J_B term, with the “analysis increment” of the unobserved variables deduced offline using the analysis increment of the observed variables, provided we have knowledge of the cross-covariance between observed and unobserved variables.

We extended the concept of balance operators in a 3D-Var context to include any variable (here chemical species variables) in addition to meteorological variables. In fact, any set of correlated random variables $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_k\}$ can be transformed into a set of uncorrelated random variables $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \dots, \mathbf{u}_k\}$ via a Gram–Schmidt orthogonalization procedure provided we define a proper inner product for random variables by using the mathematical expectation operator. The Hilbert space representation for random variables is a powerful tool that can be used in other contexts such as in cross-validation and optimization of covariance parameters [73]. The procedure to construct balance operators is not new, but the way presented here captures the general nature of the approach applicable to any geophysical context.

Obtaining error statistics of chemical variables that are observed with a single polar orbiting satellite requires some adjustment and modification of the standard methods used in the data-rich meteorological context. For example, we adapted the Hollingsworth–Lonnberg method assuming that the statistics is homogeneous on a sphere, depends only on height, and using the distance between consecutive profiles along the satellite track as a measure of distance to construct the innovation autocovariance function. We showed that we can extract spatially correlated and uncorrelated components, from which we assign the observation error and the background error variances. Since the NMC method assumes implicitly a wide coverage of observations per analyses, we used instead the Canadian Quick Covariance method (CQC) [63], which consists in using six-hour differences of the forecasts. The CQC method actually represents the spatial statistics due to advection with physical and chemical forcing terms. These difference fields are then used to obtain the parameters of a horizontal-vertical non-separable spectral correlations model. We show that spatial correlations of most species are in fact separable in the stratosphere except for O_3 , and HNO_3 on large scales (wavenumber 20 and smaller). However, the resulting horizontal correlation lengths appear to be too small. With the CQC approach, we also computed cross-covariances between ozone and temperature, and showed that it contains signals not only from photodissociation and ozone–radiation interaction but also transport, which is undesirable. The cross-covariance should in fact be computed between ozone and the residual temperature unexplained by the mass field (i.e., unbalanced temperature) rather than the temperature itself, but this would require additional development of the CQC method.

Despite these approximations and limitations, we conducted several assimilation experiments. First, we showed the added-value of limb sounding temperature measurements in the stratosphere. By assimilating MIPAS temperatures without the stratospheric AMSU channels, we created a model state that could effectively be used as an anchor run for a bias correction of the stratospheric AMSU channels. Secondly, the assimilation of vertically-resolved MIPAS temperatures is shown to reduce the temperature error variance and bias in the mid- and upper-stratosphere more than using the bias-corrected AMSU radiances, this despite the increased horizontal density and spatial coverage of AMSU.

We then examined further the weak coupling due to ozone–radiation interaction and showed that the impact on analysis is nearly negligible but develops over the forecast time. Additionally, it arises with a simplified linearized ozone chemistry model and does not require a full chemical representation.

We also conducted a strong coupling assimilation experiment between ozone and temperature using a 3D-Var assimilation scheme with a balance operator between ozone and temperature increments by information from using the CQC method. The strong 3D-Var data assimilation coupling experiment has virtually no impact in the upper stratosphere because of the very fast time-scale of the model adjustment process (photochemical and radiative), while the impact in the lower stratosphere is a small degradation in error variances. We suspect that the use of a balance operator using temperature instead of the residual temperature unexplained by the mass field (i.e., unbalanced temperature) could be responsible for this degradation.

Finally, we used a strong constraint 4D-Var to assimilate long-lived chemical species (O_3 , CH_4 and N_2O) observations from the limb sounder MIPAS to infer winds in the stratosphere. Inference on winds can be mechanistic in nature, that means recovering wind information from a time series of concentrations (e.g., a uniform concentration has no mechanistic capability in inferring winds). The inference can also be statistical in nature, where gradients in concentration error variances introduces cross-covariances between winds and chemical tracers [43,44,48]. Our experiments demonstrated the importance of having correct chemical background and observation error covariances, thus, supporting the statistical inference nature of the problem. The use of multiple tracers was also shown to be complementary, as the horizontal distribution of concentration gradients and vertical distribution of background error is different for different chemical tracers. Overall, an improvement in the tropical zonal winds was found in the lower stratosphere and a large portion of the troposphere, as assessed with radiosonde observations. A zonal-wind increment of about 2.5 ms^{-1} was also found in the surf-zone above 5 hPa but it is unclear if this helped the transport of chemical constituents, possibly due to the fact that assimilating chemical tracers result in a second-order effect, which is not easily detectable. We also observed the buildup of a temperature bias in the tropical lower stratosphere (at 20 hPa) associated with the tropical wind correction—a wind correction that is supported by the radiosonde observations.

Overall, the coupled meteorology–chemistry data assimilation experiments have shown some interesting results in both weak and strong coupling experiments, but also in terms of added value of limb sounding observations and in the importance of correctly prescribing the error covariances.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2073-4433/10/12/798/s1>, **Figure S1.** Flow chart covering the main steps and options of the 3D-Var-Chem. **Figure S2.** Scalar gain for O_3 , CH_4 , N_2O , HNO_3 , NO_2 , and $\ln(H_2O)$. **Figure S3.** Background vertical error correlation power spectra from 6h-difference method. **Figure S4.** Background error variance from 6hr-difference method. **Figure S5.** Horizontal correlation length. **Figure S6.** Mean analysis increment for O_3 , CH_4 , N_2O , NO_2 . **Figure S7.** Scatter of O_3 and streamfunction values between 10 and 100 hPa for the month of March 2003. **Figure S8.** Cross-correlation between ozone and temperature derived from 24-h difference method for July 2003. **Figure S9.** Horizontal coverage of AMSU-A profiles in six hours. **Figure S10.** Sensitivity matrix of brightness temperature over temperature for channels 10–14 of AMSU-A. **Figure S11.** Mean analysis increment at 10 hPa for the month of September 2003. **Figure S12.** Zonal mean analysis increment for September 2003. **Figure S13.** Global verification of observation-minus-forecast temperatures for different forecast lead time. **Figure S14.** Coefficient of the LINOZ scheme for September. **Figure S15.** LINOZ climatology for September. **Figure S16.** Same as Figure 19 but at 50 hPa. **Figure S17.** Same as Figure S19 but at 100 hPa. **Figure S18.** Analysis of N_2O and O_3 at 100 hPa on 11 August 2003, 00 UTC. **Figure S19.** OmP

ozone comparison against MIPAS for the 3D-Var assimilation cycle and 4D-Var for the period of 20 September to 5 October 2003 over the South Pole region and Southern Hemisphere mid-latitudes.

Author Contributions: Conceptualization: R.M.; methodology: R.M., P.G., Y.R., A.R., J.d.G.; software: Y.R., A.R., J.d.G., Y.Y., C.C., S.C.; validation: P.G., Y.R., A.R., J.d.G., Y.Y., C.C., S.C.; formal analysis: R.M., P.G., Y.R., A.R., J.d.G.; investigation: R.M., P.G., Y.R., A.R., J.d.G., writing—original draft preparation: R.M.; writing—review and editing: R.M., Y.R., A.R., J.d.G., S.C., visualization: Y.R., A.R., J.d.G., Y.Y., C.C., S.C.; supervision: R.M. and P.G.; project administration: R.M.; funding acquisition: R.M.

Funding: This work was funded by the European Space Agency/ESTEC contract no. 18560/04/NL/FF “Coupled Chemistry–Dynamics Data Assimilation” with the Contract Officer Tobias Wehr (ESA/ESTEC) This work was also supported in kind by the Atmospheric Science and Technology Directorate of Environment and Climate Change Canada (ECCC).

Acknowledgments: We wish to thank Saroja Polavarapu for her numerous advice in data assimilation and help identifying some particular issues in data assimilation. We wish to thank Mateusz Reszka and Jacek Kaminski for the review of the manuscript. We wish to thank Bernard Bilodeau and Sergey Skachko for discussions on coupled data assimilation. We wish to thank Michel Béland, Director of Climate and Atmospheric Research Directorate and the late Keith Puckett, Director of the Air Quality Research Division both at Environment and Climate Change Canada for their continuous support in making this project a reality. We thank Dominique Fonteyn from the Belgian Institute for Space Aeronomy for the initiation of our fruitful partnership. We thank Paul Vaillancout, Mike Neish and Cathy Xie from ECCC and the late John C. McConnell from York University for their help and advice on specific issues. Finally, we are grateful to the two anonymous reviewers.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors, ESA/ESTEC and the government of Canada, had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Abbreviations

3D-Var	Three-dimensional variational analysis
3D-Var-Chem	3D-Var coupled meteorology–chemistry
4D-Var	Four-dimensional variational analysis
AMSU	Advanced Microwave Sounding Unit
BASCOE	Belgian Assimilation System for Chemical Observations
CMC	Canadian Meteorological Center
CQC	Canadian Quick Covariance method
DU	Dobson Unit
ECCC	Environment and Climate Change Canada
ECMWF	European Center for Medium Range Forecasting
EOS	Earth Observing System
FGAT	First guess at appropriate time
GEM	Global Environmental Multiscale
GEM-BACH	GEM Belgian Atmospheric Chemistry
HALOE	HALogen Occultation Experiment
HL	Hollingsworth–Lönnberg method
IR	Infrared
LINOZ	LINearized model for Ozone
MDPI	Multidisciplinary Digital Publishing Institute
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
NASA	National Aeronautics and Space Administration
NH	Northern Hemisphere
NMC	National Meteorological Center method
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical Weather Prediction
O-P	Observation minus Prediction (or forecast)
RMS	Root Mean Square
SH	Southern Hemisphere
TOMS	Total Ozone Mapping Spectrometer
UARS	Upper Atmosphere Research Satellite
WRF-CHEM	Weather and Research Forecasting model coupled with Chemistry

Appendix A Derivation of Analysis Splitting between Observed and Unobserved Variables

Mathematically, the problem is posed as follows (the derivation was first published in a conference proceedings [95]). Let us find a change of variable that would simplify the J_b term to a simple inner product term. The way to accomplish this transformation of variable is by factoring $\mathbf{B}$ into square root and invertible matrix $\mathbf{S}$:

$$\mathbf{B} = \mathbf{S}\mathbf{S}^T \quad (\text{A1})$$

Defining $\boldsymbol{\zeta}$ such that:

$$\mathbf{z} = \mathbf{S}\boldsymbol{\zeta} \quad (\text{A2})$$

the J_b term then simplifies to:

$$(\mathbf{z} - \mathbf{z}^f)^T \mathbf{B}^{-1} (\mathbf{z} - \mathbf{z}^f) = (\boldsymbol{\zeta} - \boldsymbol{\zeta}^f)^T (\boldsymbol{\zeta} - \boldsymbol{\zeta}^f) \quad (\text{A3})$$

Introducing a representation of observed and unobserved variables in the $\mathbf{B}$ covariance matrix leads to a decomposition of the form:

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{xx} & \mathbf{B}_{xu} \\ \mathbf{B}_{ux} & \mathbf{B}_{uu} \end{pmatrix} \quad (\text{A4})$$

The inverse of $\mathbf{B}$ is then of the form:

$$\mathbf{B}^{-1} = \begin{pmatrix} \mathbf{D} & \mathbf{E} \\ \mathbf{F} & \mathbf{G} \end{pmatrix} \quad (\text{A5})$$

where:

$$\begin{aligned} \mathbf{D} &= (\mathbf{B}_{xx} - \mathbf{B}_{xu} \mathbf{B}_{uu}^{-1} \mathbf{B}_{ux})^{-1} \\ \mathbf{E} &= -\mathbf{B}_{xx}^{-1} \mathbf{B}_{xu} \mathbf{G} \\ \mathbf{F} &= -\mathbf{B}_{uu}^{-1} \mathbf{B}_{ux} \mathbf{D} \\ \mathbf{G} &= (\mathbf{B}_{uu} - \mathbf{B}_{ux} \mathbf{B}_{xx}^{-1} \mathbf{B}_{xu})^{-1}. \end{aligned} \quad (\text{A6})$$

To obtain the square root $\mathbf{S}$, let it first be represented in the form:

$$\mathbf{S} = \begin{pmatrix} \mathbf{d} & \mathbf{e} \\ \mathbf{f} & \mathbf{g} \end{pmatrix} \quad (\text{A7})$$

Then from Equation (A1) we obtain:

$$\begin{aligned} \mathbf{B}_{xx} &= \mathbf{d}\mathbf{d}^T + \mathbf{e}\mathbf{e}^T \\ \mathbf{B}_{xu} &= \mathbf{f}\mathbf{d}^T + \mathbf{g}\mathbf{e}^T \\ \mathbf{B}_{ux} &= \mathbf{d}\mathbf{f}^T + \mathbf{e}\mathbf{g}^T \\ \mathbf{B}_{uu} &= \mathbf{f}\mathbf{f}^T + \mathbf{g}\mathbf{g}^T. \end{aligned} \quad (\text{A8})$$

There is more than one solution that satisfies these four equations. One of the solutions that leads to a triangular form consists of letting $\mathbf{e} = 0$. We can then easily invert $\mathbf{S}$. So letting $\mathbf{e} = 0$ in Equation (A8) we first get:

$$\mathbf{B}_{xx} = \mathbf{d}\mathbf{d}^T \quad (\text{A9})$$

This is the square-root form of the background error covariance matrix used in 3D-Var, which is usually denoted as $\mathbf{L}$, thus, we have:

$$\mathbf{d} = \mathbf{L} \quad (\text{A10})$$

From the second line of Equation (A8) we obtain:

$$\mathbf{B}_{\mathbf{u}\mathbf{x}} = \mathbf{f}\mathbf{L}^T \Rightarrow \mathbf{f} = \mathbf{B}_{\mathbf{u}\mathbf{x}}\mathbf{L}^{-T} \quad (\text{A11})$$

and the third equation is satisfied trivially. Finally the fourth equation of Equation (A8) takes the form:

$$\mathbf{B}_{\mathbf{u}\mathbf{u}} = \mathbf{B}_{\mathbf{u}\mathbf{x}}\mathbf{L}^{-T}\mathbf{L}^{-1}\mathbf{B}_{\mathbf{x}\mathbf{u}} + \mathbf{g}\mathbf{g}^T \Rightarrow \mathbf{g}\mathbf{g}^T = \mathbf{B}_{\mathbf{u}\mathbf{u}} - \mathbf{B}_{\mathbf{u}\mathbf{x}}\mathbf{B}_{\mathbf{x}\mathbf{x}}^{-1}\mathbf{B}_{\mathbf{x}\mathbf{u}} = \mathbf{G}^{-1} \quad (\text{A12})$$

where we used the fact that $\mathbf{B}_{\mathbf{x}\mathbf{x}}^{-1} = \mathbf{L}^{-T}\mathbf{L}^{-1}$, thus, $\mathbf{g}$ is the square-root of the inverse of $\mathbf{G}$:

$$\mathbf{g} = \sqrt{\mathbf{G}^{-1}} \quad (\text{A13})$$

Now let:

$$\tilde{\mathbf{z}} = \begin{pmatrix} \tilde{\mathbf{x}} \\ \tilde{\mathbf{u}} \end{pmatrix} \quad (\text{A14})$$

where the tilde variables are departures from the forecast, i.e., $\tilde{\mathbf{x}} = \mathbf{x} - \mathbf{x}^f$, $\tilde{\mathbf{u}} = \mathbf{u} - \mathbf{u}^f$, $\tilde{\mathbf{z}} = \mathbf{z} - \mathbf{z}^f$. Then consider the transformed variable:

$$\tilde{\boldsymbol{\zeta}} = \sqrt{\mathbf{B}^{-1}}\tilde{\mathbf{z}} \quad (\text{A15})$$

which allows us to write:

$$\tilde{\mathbf{z}}^T \mathbf{B}^{-1} \tilde{\mathbf{z}} = \tilde{\boldsymbol{\zeta}}^T \tilde{\boldsymbol{\zeta}} \quad (\text{A16})$$

Specifically, we have:

$$\tilde{\mathbf{z}} = \begin{pmatrix} \tilde{\mathbf{x}} \\ \tilde{\mathbf{u}} \end{pmatrix} = \sqrt{\mathbf{B}}\tilde{\boldsymbol{\zeta}} = \mathbf{S}\tilde{\boldsymbol{\zeta}} = \begin{pmatrix} \mathbf{L} & 0 \\ \mathbf{B}_{\mathbf{u}\mathbf{x}}\mathbf{L}^{-T} & \sqrt{\mathbf{G}^{-1}} \end{pmatrix} \begin{pmatrix} \tilde{\zeta}_1 \\ \tilde{\zeta}_2 \end{pmatrix} \quad (\text{A17})$$

and this system is easily inverted to give:

$$\begin{aligned} \tilde{\zeta}_1 &= \mathbf{L}^{-1}\tilde{\mathbf{x}} \\ \tilde{\zeta}_2 &= \sqrt{\mathbf{G}}(\tilde{\mathbf{u}} - \mathbf{B}_{\mathbf{u}\mathbf{x}}\mathbf{B}_{\mathbf{x}\mathbf{x}}^{-1}\tilde{\mathbf{x}}) \end{aligned} \quad (\text{A18})$$

The cost function Equation (2) (main text) written in terms of these variables yields:

$$\begin{aligned} 2J(\tilde{\mathbf{x}}, \tilde{\mathbf{u}}) &= \begin{pmatrix} \tilde{\mathbf{x}} \\ \tilde{\mathbf{u}} \end{pmatrix}^T \mathbf{B}^{-1} \begin{pmatrix} \tilde{\mathbf{x}} \\ \tilde{\mathbf{u}} \end{pmatrix} + (\mathbf{y} - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}) \\ &= (\boldsymbol{\zeta} - \boldsymbol{\zeta}^f)^T (\boldsymbol{\zeta} - \boldsymbol{\zeta}^f) + (\mathbf{y} - \mathbf{H}\mathbf{L}\boldsymbol{\zeta}_1)^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{L}\boldsymbol{\zeta}_1) \\ &= (\boldsymbol{\zeta}_1 - \boldsymbol{\zeta}_1^f)^T (\boldsymbol{\zeta}_1 - \boldsymbol{\zeta}_1^f) + (\boldsymbol{\zeta}_2 - \boldsymbol{\zeta}_2^f)^T (\boldsymbol{\zeta}_2 - \boldsymbol{\zeta}_2^f) + (\mathbf{y} - \mathbf{H}\mathbf{L}\boldsymbol{\zeta}_1)^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{L}\boldsymbol{\zeta}_1) \\ &= 2J_1(\boldsymbol{\zeta}_1) + 2J_2(\boldsymbol{\zeta}_2) \end{aligned} \quad (\text{A19})$$

which has the interesting property that the *minimization with respect to $\boldsymbol{\zeta}_1$ is independent of the minimization with respect to $\boldsymbol{\zeta}_2$* . The minimization with respect to $\boldsymbol{\zeta}_1$ is given by minimizing the cost function:

$$J_1(\boldsymbol{\zeta}_1) = \frac{1}{2}(\boldsymbol{\zeta}_1 - \boldsymbol{\zeta}_1^f)^T (\boldsymbol{\zeta}_1 - \boldsymbol{\zeta}_1^f) + \frac{1}{2}(\mathbf{y} - \mathbf{H}\mathbf{L}\boldsymbol{\zeta}_1)^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{L}\boldsymbol{\zeta}_1) \quad (\text{A20})$$

and that with respect to $\boldsymbol{\zeta}_2$ with the cost function:

$$J_2(\boldsymbol{\zeta}_2) = \frac{1}{2}(\boldsymbol{\zeta}_2 - \boldsymbol{\zeta}_2^f)^T (\boldsymbol{\zeta}_2 - \boldsymbol{\zeta}_2^f) \quad (\text{A21})$$

The minimization of Equation (A21) has the trivial solution:

$$\boldsymbol{\zeta}_2 - \boldsymbol{\zeta}_2^f = 0 \quad (\text{A22})$$

Now, assuming that $\sqrt{\mathbf{G}}$ is invertible, the Equation (A22) yields:

$$\mathbf{u}^a - \mathbf{u}^f = \mathbf{B}_{\mathbf{u}\mathbf{x}} \mathbf{B}_{\mathbf{x}\mathbf{x}}^{-1} (\mathbf{x}^a - \mathbf{x}^f) \quad (\text{A23})$$

The minimization of the cost function $J_1(\zeta_1)$ Equation (A20) is actually identical to the form 3D-Var takes written after preconditioning. Indeed in normal form with the non-transformed variables, Equation (A20) takes the form:

$$J(\mathbf{x}) = \frac{1}{2} (\mathbf{x} - \mathbf{x}^f)^T \mathbf{B}_{\mathbf{x}\mathbf{x}}^{-1} (\mathbf{x} - \mathbf{x}^f) + \frac{1}{2} (\mathbf{y} - H\mathbf{x})^T \mathbf{R}^{-1} (\mathbf{y} - H\mathbf{x}) \quad (\text{A24})$$

Appendix B Geometric Interpretation of the Derivation of the Balance Equations

Balance between different variables occur, in fact, in many geophysical problems. Here in the context of chemistry, it occurs between long-lived species, or between ozone and temperature (which we will develop in detail below). Using the statistical regression modeling allows to formulate the balance operators in a general context for any geophysical problem.

To simplify the representation of the background error covariance $\mathbf{B}_{\mathbf{x}\mathbf{x}}$, the set of correlated variables is transformed via a Gram–Smidt orthogonalization procedure into a set of uncorrelated variables whose covariance representation is then block-diagonal. The transformation from uncorrelated variables back to the original variables is achieved through what is called, a *balance operator* or, in fact, linear regression.

Random variables (and random vectors) can be represented as a Hilbert space provided we use the mathematical expectation to define the inner product [73] (or see [74] Section 1.2). For random variables (vectors) that have a non-zero mean, the proper definition of an inner product is:

$$\langle \mathbf{x}, \mathbf{y} \rangle = E[(\mathbf{x} - E(\mathbf{x}))(\mathbf{y} - E(\mathbf{y}))^T] = \text{cov}(\mathbf{x}, \mathbf{y}) \quad (\text{A25})$$

where $\mathbf{x}$ and $\mathbf{y}$ are random vectors. The effect of an inner product in a Hilbert space of random variables is, thus, to create a non-random variable. In Equation (A25), $\langle \mathbf{x}, \mathbf{y} \rangle$ is a matrix where each entry is non-random. The square of the norm is then the variance, $\|\mathbf{x}\|^2 = \text{var}(\mathbf{x})$, and the correlation matrix θ , between variables $\mathbf{x}$ and $\mathbf{y}$ is obtained as $\cos(\theta) = \langle \mathbf{x}, \mathbf{y} \rangle / (\|\mathbf{x}\| \|\mathbf{y}\|)$. Therefore, the uncorrelated random variables $\text{cov}(\mathbf{x}, \mathbf{y}) = 0$ are orthogonal, i.e., $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

A set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_k\}$ of variables of a Hilbert space can always be transformed into a set of orthogonal variables $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \dots, \mathbf{u}_k\}$ via the Gram–Schmidt orthogonalization procedure as follows:

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{v}_1 \\ \mathbf{u}_2 &= \mathbf{v}_2 - \text{proj}_{\mathbf{u}_1}(\mathbf{v}_2) \\ \mathbf{u}_3 &= \mathbf{v}_3 - \text{proj}_{\mathbf{u}_1}(\mathbf{v}_3) - \text{proj}_{\mathbf{u}_2}(\mathbf{v}_3) \\ &\dots \end{aligned} \quad (\text{A26})$$

where the projection (proj) is defined as:

$$\text{proj}_{\mathbf{u}}(\mathbf{v}) \triangleq \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u} \quad (\text{A27})$$

Applying this procedure to random vectors using Equation (A25) and specifically to (unbiased) model background errors of the streamfunction $\delta\psi$, velocity potential $\delta\chi$, temperature δT , and ozone δO_3 which are known to be correlated, we obtain the *transformed uncorrelated* background error variables (denoted with a superscript u), Equation (6).

Appendix C Error Covariance from the LINOZ Scheme

The coefficients c_1, c_2, c_3, c_4 of the LINOZ scheme for September, determined using a box model, are illustrated in Figure S14 (Supplementary Material), and the mean state in Figure S15. Note that in the Equation (20) the concentration is expressed as a mixing ratio (in ppmv) and is, thus, typically on the order of 10^{-6} .

The overhead column in Dobson units (DU) is calculated as follows. By definition one DU is equivalent to 0.01 mm of ozone at standard temperature and pressure and is equal to 2.69×10^{16} molecules cm^{-2} . The overhead number of molecules of ozone is:

$$\int_z^\infty n_{\text{O}_3}(z') dz' \quad (\text{A28})$$

where n is the *number density*, expressed generally in molecules- m^{-3} . The *volume mixing ratio* is the ratio of the number density of the gas over the number density of (dry) air, i.e.,

$$\text{O}_3 = \frac{n_{\text{O}_3}}{n_A} \quad (\text{A29})$$

Using the relationship:

$$n_A M_A = \rho_A N_a \quad (\text{A30})$$

where M_A is the molecular weight of air (equal to $0.028964 \text{ kg mol}^{-1}$), N_a is Avogadro's number (equal to $6.02252 \times 10^{23} \text{ molecules mol}^{-1}$), and ρ_A is the density of air, the overhead number of ozone molecules can then be rewritten as:

$$\int_z^\infty n_{\text{O}_3}(z') dz' = \frac{N_a}{M_A} \int_z^\infty \text{O}_3(z') \rho_A(z') dz' = -\frac{N_a}{M_A} \int_p^0 \frac{\text{O}_3(p')}{g} dp' \quad (\text{A31})$$

Taking perturbations of ozone, $\tilde{\text{O}}_3$, and temperature, $\tilde{T}$, in Equation (20) gives the following evolution equation for the perturbations:

$$\frac{D\tilde{\text{O}}_3}{Dt} = c_2 \tilde{\text{O}}_3 + c_3 \tilde{T} + kc_4 \int_p^0 \tilde{\text{O}}_3(p') dp' \quad (\text{A32})$$

where k is a constant that accounts for expressing the overhead column in DU. In Equation (A32) we have neglected the changes in wind due to temperature perturbation. The last term of Equation (A32) contributes primarily in the lower stratosphere.

To establish an error cross-covariance between temperature and ozone let us first neglect the overhead ozone component in Equation (A32), and let us assume for now that the material derivative of temperature perturbations is small compared to that of ozone, which agrees with the fact that temperature changes to ozone perturbations that occur on a much longer time-scale than photochemical perturbations (ozone changes due to temperature perturbations), i.e., let us assume:

$$\frac{D\tilde{T}}{Dt} = 0 \quad (\text{A33})$$

$$\frac{D\tilde{\text{O}}_3}{Dt} = c_2 \tilde{\text{O}}_3 + c_3 \tilde{T} \quad (\text{A34})$$

Multiplying Equation (A33) by $\tilde{O}_3$ and Equation (A34) by $\tilde{T}$, adding the resulting equations and taking the expectation gives:

$$\frac{D\langle\tilde{O}_3\tilde{T}\rangle}{Dt} = c_2\langle\tilde{O}_3\tilde{T}\rangle + c_3\langle\tilde{T}^2\rangle. \quad (A35)$$

In Equation (A34) the error cross-covariance is between any pressure levels. Carrying out the derivation more explicitly with pressure levels, Equation (A35) can be rewritten as:

$$\frac{D\langle\tilde{O}_3(p)\tilde{T}(p')\rangle}{Dt} = c_2(p)\langle\tilde{O}_3(p)\tilde{T}(p')\rangle + c_3(p)\langle\tilde{T}(p)\tilde{T}(p')\rangle \quad (A36)$$

where p and p' are two distinct pressure levels. Multiplying Equation (A33) by T' also gives:

$$\frac{D\langle\tilde{T}(p)\tilde{T}(p')\rangle}{Dt} = 0 \quad (A37)$$

Similarly a covariance evolution for ozone can be derived as follows. Multiplying Equation (A34) at pressure level p with $\tilde{O}_3(p')$, and multiplying Equation (A34) at pressure level p' with $\tilde{O}_3(p)$, adding the equations and taking the expectation gives:

$$\frac{D\langle\tilde{O}_3(p)\tilde{O}_3(p')\rangle}{Dt} = [c_2(p) + c_2(p')]\langle\tilde{O}_3(p)\tilde{O}_3(p')\rangle + c_3(p)\langle\tilde{T}(p)\tilde{O}_3(p')\rangle + c_3(p')\langle\tilde{O}_3(p)\tilde{T}(p')\rangle \quad (A38)$$

In matrix form Equation (A36) is rewritten as:

$$\frac{D\mathbf{P}_{O_3T}}{Dt} = \mathbf{C}_2\mathbf{P}_{O_3T} + \mathbf{C}_3\mathbf{P}_{TT} \quad (A39)$$

where $\mathbf{C}_2$ and $\mathbf{C}_3$ are diagonal matrices, i.e., of the form:

$$\mathbf{C}_2 = \begin{bmatrix} c_2(p_1) & 0 & 0 \\ 0 & \cdots & 0 \\ 0 & 0 & c_2(p_N) \end{bmatrix} \quad (A40)$$

and:

$$\mathbf{P}_{O_3T} = \begin{bmatrix} \langle\tilde{O}_3(p_1)\tilde{T}(p_1)\rangle & \cdots & \langle\tilde{O}_3(p_1)\tilde{T}(p_N)\rangle \\ \cdots & \cdots & \cdots \\ \langle\tilde{O}_3(p_N)\tilde{T}(p_1)\rangle & \cdots & \langle\tilde{O}_3(p_N)\tilde{T}(p_N)\rangle \end{bmatrix} \quad (A41)$$

$$\mathbf{P}_{TT} = \begin{bmatrix} \langle\tilde{T}(p_1)\tilde{T}(p_1)\rangle & \cdots & \langle\tilde{T}(p_1)\tilde{T}(p_N)\rangle \\ \cdots & \cdots & \cdots \\ \langle\tilde{T}(p_N)\tilde{T}(p_1)\rangle & \cdots & \langle\tilde{T}(p_N)\tilde{T}(p_N)\rangle \end{bmatrix} \quad (A42)$$

In matrix form Equation (A37) is written as:

$$\frac{D\mathbf{P}_{TT}}{Dt} = 0 \quad (A43)$$

and Equation (A38) can be rewritten as:

$$\frac{D\mathbf{P}_{O_3O_3}}{Dt} = \mathbf{C}_2\mathbf{P}_{O_3O_3} + \mathbf{C}_3\mathbf{P}_{O_3T} + (\mathbf{C}_2\mathbf{P}_{O_3O_3} + \mathbf{C}_3\mathbf{P}_{O_3T})^T \quad (A44)$$

where the superscript T is the matrix transpose.

At this point, it is hypothetical what kind of assumption is needed to derive a balance in these covariance evolution equations. Based on time-scales, one might argue that the derivative of the error

cross-covariance evolves slowly, and to a first approximation, we may want to consider the following balance:

$$\mathbf{C}_2 \mathbf{P}_{O_3 T} + \mathbf{C}_3 \mathbf{P}_{TT} = 0 \quad (\text{A45})$$

from which we obtain:

$$\mathbf{P}_{O_3 T} = -\mathbf{C}_2^{-1} \mathbf{C}_3 \mathbf{P}_{TT} \quad (\text{A46})$$

and the balance operator is then of the form:

$$\mathbf{A} = \mathbf{P}_{O_3 T} \mathbf{P}_{TT}^{-1} = -\mathbf{C}_2^{-1} \mathbf{C}_3 \quad (\text{A47})$$

which is a diagonal matrix. This is what is being used to create Figure 6 (right panel) in the main document.

In Figure A1 we plot the ratio $c_3(p)/c_2(p)$.

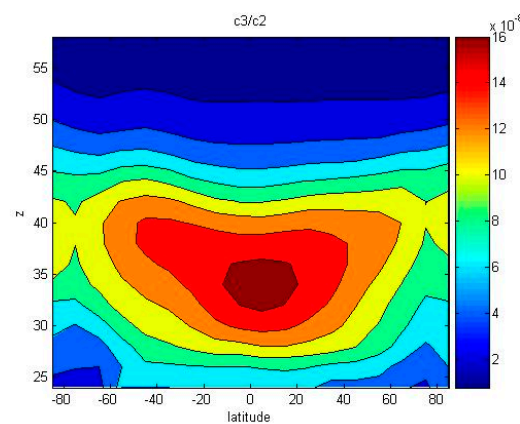


Figure A1. Ratio $c_3(p)/c_2(p)$ for the month of September.

Limiting the plot below 24 km was necessary because both coefficients c_2, c_3 change by several orders of magnitude from top to bottom, with very small values of c_2 in the lower stratosphere, and are, thus, prone to numerical error amplification. Surprisingly, the isolines of the ratio follow the general pattern of zonal mean ozone. Figure A2 depicts a point by point scatter plot between the ratio $c_3(p)/c_2(p)$ and the zonal mean ozone $\bar{O}_3$. A very high correlation is, thus, observed.

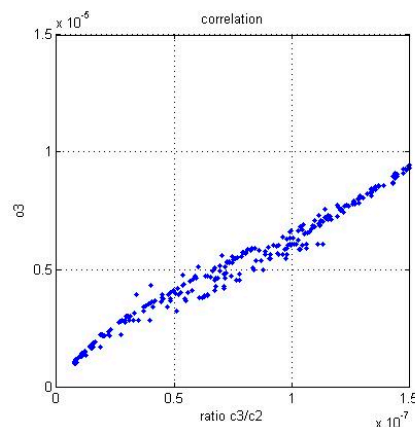


Figure A2. Point-by-point, (λ, p) , scatter of $c_3(p)/c_2(p)$ with $\bar{O}_3$.

With such a balance model, temperature increments may produce realistic ozone increments, and may be an avenue worth investigating further.

Extension with the Photochemical Term c_4

Let us now add the contribution from the photochemical term c_4 . The equation for the ozone perturbation, Equation (A32) is now written as:

$$\frac{D\tilde{O}_3(p)}{Dt} = c_2(p)\tilde{O}_3(p) + c_3(p)\tilde{T}(p) + kc_4(p) \int_p^0 \tilde{O}_3(p'') dp'' \quad (A48)$$

From Equations (A48) and (A33) we obtain:

$$\frac{D\langle \tilde{O}_3(p)\tilde{T}(p') \rangle}{Dt} = c_2(p)\langle \tilde{O}_3(p)\tilde{T}(p') \rangle + c_3(p)\langle \tilde{T}(p)\tilde{T}(p') \rangle + kc_4(p) \int_p^0 \langle \tilde{O}_3(p'')\tilde{T}(p') \rangle dp'' \quad (A49)$$

In discrete form, the last term of the r.h.s. of Equation (A49) takes the following form:

$$\begin{aligned} &\text{For } p = p_N \text{ and } p' = p_j \quad kc_4(p_N) \langle \tilde{O}_3(p_N)\tilde{T}(p_j) \rangle \Delta p_N \\ &\text{For } p = p_{N-1} \text{ and } p' = p_j \quad kc_4(p_{N-1}) \{ \langle \tilde{O}_3(p_N)\tilde{T}(p_j) \rangle \Delta p_N + \langle \tilde{O}_3(p_{N-1})\tilde{T}(p_j) \rangle \Delta p_{N-1} \} \\ &\dots \\ &\text{For } p = p_i \text{ and } p' = p_j \quad kc_4(p_i) \sum_{k=i}^N \langle \tilde{O}_3(p_k)\tilde{T}(p_j) \rangle \Delta p_k \end{aligned}$$

and in matrix form:

$$\Sigma \mathbf{P}_{O_3T} \quad (A50)$$

where:

$$\Sigma = \begin{bmatrix} \Delta p_1 & \Delta p_2 & \dots & \Delta p_N \\ 0 & \Delta p_2 & \dots & \Delta p_N \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \Delta p_N \end{bmatrix} \quad (A51)$$

In matrix form, similarly, to Equation (A39) we obtain:

$$\frac{D\mathbf{P}_{O_3T}}{Dt} = \mathbf{C}_2\mathbf{P}_{O_3T} + \mathbf{C}_3\mathbf{P}_{TT} + k\mathbf{C}_4\Sigma\mathbf{P}_{O_3T} = (\mathbf{C}_2 + k\mathbf{C}_4\Sigma)\mathbf{P}_{O_3T} + \mathbf{C}_3\mathbf{P}_{TT} \quad (A52)$$

The balance operator is then of the form:

$$\mathbf{A} = -(\mathbf{C}_2 + k\mathbf{C}_4\Sigma)^{-1}\mathbf{C}_3 \quad (A53)$$

To compute the Δp_k term we use a centered formula:

$$\Delta p_k = p_{k+\frac{1}{2}} - p_{k-\frac{1}{2}} \quad (A54)$$

Data tabulated at discrete heights can be transformed into pressure by integrating the hydrostatic equation and gas law:

$$\frac{dp}{p} = -\frac{g}{RT}dz \quad (A55)$$

giving, in discrete form:

$$p_k = p_0 \exp \left\{ -\frac{g\Delta z}{R} \left(\frac{1}{\bar{T}_0} + \frac{1}{\bar{T}_1} + \dots + \frac{1}{\bar{T}_{k-1}} \right) \right\} \quad (A56)$$

where $\bar{T}_k$ is the mean layer temperature. The Figure A3 below illustrates the vertical staggering:

$$\begin{array}{ccc} p_{k+1} & \text{---} & z_{k+1} \\ & \bullet & \bar{T}_k \\ p_k & \text{---} & z_k \end{array}$$

Figure A3. Vertical staggering of temperature and height.

Defining:

$$p_{k+\frac{1}{2}} = \frac{p_k + p_{k+1}}{2} \quad p_{k-\frac{1}{2}} = \frac{p_k + p_{k-1}}{2} \quad (\text{A57})$$

we obtain:

$$\Delta p_k = \frac{p_0}{2} \exp\left\{-\frac{g\Delta z}{R}\left(\frac{1}{\bar{T}_0} + \dots + \frac{1}{\bar{T}_{k-1}}\right)\right\} \left[\exp\left(-\frac{g\Delta z}{R\bar{T}_k}\right) - \exp\left(+\frac{g\Delta z}{R\bar{T}_{k-1}}\right)\right] \quad (\text{A58})$$

References

1. Lahoz, W. Research Satellites. In *Data Assimilation: Making Sense of Observations*; Lahoz, W., Khattatov, B., Ménard, R., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; pp. 301–321. [\[CrossRef\]](#)
2. Reber, C.A.; Trevathan, C.E.; McNeal, R.J.; Luther, M.R. The Upper Atmosphere Research Satellite (UARS) mission. *J. Geophys. Res.* **1993**, *98*, 10643–10647. [\[CrossRef\]](#)
3. UARS Science Team; Rood, R.B.; Geller, M.A. UARS Data and Scientific Results (Special Issue). *J. Atmos. Sci.* **1994**, *51*, 2781–3105.
4. UARS Science Team; Gille, J.C.; Massie, S.T.; Mankin, W.G. Evaluation of the UARS Data (Special Issue). *J. Geophys. Res.* **1996**, *6*, 9539–10473.
5. Louet, J. The Envisat Mission and System. Available online: http://www.esa.int/esapub/bulletin/bullet106/bul106_1.pdf (accessed on 20 August 2001).
6. Envisat Science Team. Validation Workshop Proceedings. 2002. Available online: https://envisat.esa.int/pub/ESA_822DOC/envisat_val_1202/proceedings/ (accessed on 18 October 2019).
7. MIPAS Science Team. MIPAS Geophysical Validation (Special Issue). *Atmos. Chem. Phys.* **2002**, *9*, 413–442.
8. EOS Aura. Available online: <http://aura.gsfc.nasa.gov/> (accessed on 18 October 2019).
9. EOS Aura Science Team. EOS Aura (Special Issue). *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1063–1379. [\[CrossRef\]](#)
10. Dee, D.P.; Coupled, D.A. Presentation at the WMO Symposium on Data Assimilation. Available online: http://das6.umd.edu/program/Daily/slides/9.4-Dee_Dick.pdf (accessed on 18 October 2019).
11. Penny, S.G.; Akella, S.; Alves, O.; Bishop, C.; Buehner, M.; Chevallier, M.; Counillon, F.; Draper, C.; Frolov, S.; Fujii, Y.; et al. *Coupled Data Assimilation for Integrated Earth System Analysis and Prediction: Goals, Challenges and Recommendations*; WWRP 2017-3 Report; WMO: Geneva, Switzerland, 2017; p. 50.
12. Zupanski, M. Data Assimilation for Coupled Modeling Systems. In *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications*; Park, S.E., Xu, L., Eds.; Springer: Berlin/Heidelberg, Germany, 2017. [\[CrossRef\]](#)
13. Han, G.; Wu, X.; Zhang, S.; Li, W. Error covariance estimation for coupled data assimilation using the Lorenz atmosphere and a simple pycnocline ocean model. *J. Clim.* **2013**, *26*, 218–231. [\[CrossRef\]](#)
14. Tardif, R.; Hakim, G.J.; Snyder, C. Coupled atmosphere-ocean data assimilation experiments with a low-order climate model. *Clim. Dyn.* **2014**, *43*, 1631–1643. [\[CrossRef\]](#)
15. Tardif, R.; Hakim, G.J.; Snyder, C. Coupled atmosphere-ocean data assimilation experiments with a low-order model and CMIP5 model data. *Clim. Dyn.* **2015**, *45*, 1415–1427. [\[CrossRef\]](#)
16. Lu, F.; Liu, Z.; Zhang, S.; Liu, Y. Strongly coupled data assimilation using leading averaged coupled covariance (LACC). Part I: Simple model study. *Mon. Weather Rev.* **2015**, *143*, 3823–3827. [\[CrossRef\]](#)
17. Smith, P.J.; Fowler, A.M.; Lawless, A.S. Exploring strategies for coupled 4D-var data assimilation using an idealized atmosphere-ocean model. *Tellus A* **2015**, *67*, 27025. [\[CrossRef\]](#)
18. Fowler, A.M.; Lawless, A.S. An idealized study of coupled atmosphere-ocean 4D-var in the presence of model error. *Mon. Weather Rev.* **2016**, *144*, 4007–4030. [\[CrossRef\]](#)
19. Sluka, T.C.; Penny, S.G.; Kalnay, E.; Miyoshi, T. Assimilating atmospheric observations into the ocean using strongly coupled ensemble data assimilation. *Geophys. Res. Lett.* **2016**, *43*, 752–759. [\[CrossRef\]](#)

20. Penny, S.G.; Hamil, T.M. Coupled data assimilation for integrated earth system analysis and prediction. *Bull. Am. Meteorol. Soc.* **2017**. [[CrossRef](#)]
21. Storto, A.; Martin, M.J.; Deremble, B.; Masina, S. Strongly coupled data assimilation experiments with linearized ocean-atmosphere balance relationship. *Mon. Weather Rev.* **2018**, *146*, 1233–1257. [[CrossRef](#)]
22. Laloyaux, P.; Frolov, S.; Ménétrier, B.; Bonavita, M. Implicit and explicit cross-correlations in coupled data assimilation. *Q. J. R. Meteorol. Soc.* **2018**, *144*, 1851–1863. [[CrossRef](#)]
23. Browne, P.A.; de Rosnay, P.; Zuo, H.; Bennett, A.; Dawson, A. Weakly coupled ocean-atmosphere data assimilation in the ECMWF NWP system. *Remote Sens.* **2019**, *11*, 234. [[CrossRef](#)]
24. Guivarc'h, C.; Roberts-Jones, J.; Harris, C.; Lea, D.J.; Ryan, A.; Ascione, I. Assessment of ocean analysis and forecast from an atmosphere ocean coupled data assimilation operational system. *Ocean Sci.* **2019**, *15*, 1307–1326.
25. Skachko, S.; Buehner, M.; Laroche, S.; Lapalme, E.; Smith, G.; Roy, F.; Surcel-Colan, D.; Bélanger, J.-M.; Garand, L. Weakly coupled atmosphere-ocean data assimilation in the Canadian global prediction system (v1). *Geosci. Model Dev. Discuss.* **2019**, *12*, 5097–5112. [[CrossRef](#)]
26. Rasmy, M.; Koike, T.; Kuria, D.; Mirza, C.R.; Li, X.; Yang, K. Development of the coupled atmosphere and land data assimilation system (CALDAS) and its application over the Tibetan plateau. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 4227–4242. [[CrossRef](#)]
27. de Rosnay, P.; Drusch, M.; Vasiljevic, D.; Balsamo, G.; Albergel, C.; Isaksen, I. A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF. *Q. J. R. Meteorol. Soc.* **2013**, *139*, 1199–1213. [[CrossRef](#)]
28. Lea, D.; Mirouze, I.; Martin, M.; King, R.; Hines, A.; Walters, D.; Thurlow, M. Assessing a new coupled data assimilation system based on the Met Office coupled atmosphere–land–ocean–sea ice model. *Mon. Weather Rev.* **2015**, *143*, 4678–4694. [[CrossRef](#)]
29. Lin, L.-F.; Pu, Z. Characteristics of background error covariance of soil moisture and atmospheric states in strongly coupled land-atmosphere data assimilation. *Mon. Weather Rev.* **2018**, *57*, 2507–2528. [[CrossRef](#)]
30. Shahabadi, M.B.; Bélair, S.; Bilodeau, B.; Carrera, M.L.; Garand, L. Impact of weak coupling between land and atmosphere data assimilation systems on Environment and Climate Change Canada's global deterministic prediction system. *Weather Forecast.* **2019**, *34*, 1741–1758. [[CrossRef](#)]
31. Lee, E.; Županski, M.; Županski, D.; Park, S.K. Impact of OMI aerosol optical depth on analysis increments through coupled meteorology-aerosol data assimilation for an Asian dust storm. *Remote Sens. Environ.* **2017**, *193*, 38–53. [[CrossRef](#)]
32. Wu, T.-C.; Županski, M.; Saleey, S.; Kliewer, A.; Grasso, L.; Bian, Q.; Atwood, S.A.; Wang, Y.; Wang, J. RAMS-MLEF atmosphere-aerosol coupled data assimilation: A case study of a dust storm even over the Arabian peninsula on 4 August 2016. *Atmos. Chem. Phys. Discuss.* **2018**. [[CrossRef](#)]
33. Županski, M.; Kliewer, A.; Wu, T.-C.; Apodaca, K.; Bian, Q.; Atwood, S.; Wang, Y.; Wang, J.; Miller, S.D. Impact of atmospheric and aerosol optical depth observations on the aerosol initial conditions in a strongly-coupled data assimilation. *Atmos. Chem. Phys. Discuss.* **2019**. [[CrossRef](#)]
34. Benedetti, A.; Vitart, F. Can the direct effect of aerosols improve subseasonal predictability? *Mon. Weather Rev.* **2018**, *146*, 3481–3498. [[CrossRef](#)]
35. Randles, C.A.; da Silva, A.M.; Buchard, V.; Colarco, P.R.; Darmenov, A.; Govindaraju, R.; Smirnov, A.; Holben, B.; Ferrere, R.; Hair, J.; et al. The MERRA-2 aerosol reanalysis, 1980 onward. Part I: System description and data assimilation evaluation. *J. Clim.* **2017**, *30*, 6823–6850. [[CrossRef](#)]
36. Bocquet, M.H.; Elbern, H.; Eskes, M.; Hirtl, R.; Žabkar, G.R.; Carmichael, J.; Flemming, A.; Inness, M.; Pagaoski, J.L.; Pérez Camacho, P.E.; et al. Data assimilation in atmospheric chemistry models; current status and future prospects for coupled chemistry meteorology models. *Atmos. Chem. Phys.* **2015**, *15*, 5325–5358. [[CrossRef](#)]
37. Saide, P.E.; Carmichael, G.R.; Spak, S.N.; Minnis, P.; Ayers, J.K. Improving aerosol distributions below clouds by assimilating satellite-retrieved cloud droplet number. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 11939–11943. [[CrossRef](#)]
38. Park, S.K.; Lim, S.; Županski, M. Structure of forecast error covariance in coupled atmosphere-chemistry data assimilation. *Geosci. Model Dev.* **2015**, *8*, 1315–1320. [[CrossRef](#)]
39. Dragani, R.; McNally, A.P. Operational assimilation of ozone-sensitive infrared radiances at ECMWF. *Q. J. R. Meteorol. Soc.* **2013**, *139*, 2068–2080. [[CrossRef](#)]

40. Engelen, R.J.; Bauer, P. The use of variable CO₂ in the data assimilation of AIRS and IASIS radiances. *Q. J. R. Meteorol. Soc.* **2014**, *140*, 958–965. [\[CrossRef\]](#)
41. Morcrette, J.-J. *Ozone–Radiation Interactions in the ECMWF Forecast System*; European Centre for Medium-Range Weather Forecasts: Reading, UK, 2003; p. 36.
42. De Grandpré, J.; Ménard, R.; Rochon, Y.J.; Charrette, C.; Chabrilat, S.; Robichaud, A. Radiative impact of ozone on temperature predictability in a coupled chemistry-dynamics data assimilation system. *Mon. Weather Rev.* **2009**, *137*, 679–692. [\[CrossRef\]](#)
43. Daley, R. Recovery of the one and two dimensional windfields from chemical constituent observations using the constituent transport equation and an extended Kalman filter. *Meteorol. Atmos. Phys.* **1996**, *60*, 119–136. [\[CrossRef\]](#)
44. Daley, R. Estimating the windfield from chemical constituent observations: Experiments with an extended Kalman filter. *Mon. Weather Rev.* **1995**, *123*, 181–198. [\[CrossRef\]](#)
45. Riishojgaard, L.P. On four-dimensional variational assimilation of ozone data in weather prediction models. *Q. J. R. Meteorol. Soc.* **1996**, *122*, 1545–1571. [\[CrossRef\]](#)
46. Peuch, A.; Thépaut, J.-N.; Pailleux, J. Dynamical impact of total ozone observations in a four-dimensional variational assimilation. *Q. J. R. Meteorol. Soc.* **2000**, *126*, 1641–1659. [\[CrossRef\]](#)
47. Semane, N.; Peuch, V.-H.; Pradier, S.; Desroziers, G.; El Amraoui, L.; Brousseau, P.; Massart, S.; Chapnik, B.; Peuch, A. On the extraction on wind information from the assimilation of ozone profiles in Meteo-France 4D-var operational NWP suite. *Atmos. Chem. Phys.* **2009**, *9*, 4855–4867. [\[CrossRef\]](#)
48. Milewski, T.; Bourqui, M.S. Assimilation of stratospheric temperature and ozone with an ensemble Kalman filter in a chemistry-climate model. *Mon. Weather Rev.* **2011**, *139*, 3389–3404. [\[CrossRef\]](#)
49. Milewski, T.; Bourqui, M.S. Potential of an ensemble Kalman smoother for stratospheric chemical-dynamical data assimilation. *Tellus* **2013**, *65*, 18541. [\[CrossRef\]](#)
50. Allen, D.R.; Hoppel, K.W.; Nedoluha, G.E.; Baker, N.L.; Xu, L.; Rosmond, T.E. Limitations of wind extraction from 4D-var assimilation of ozone. *Atmos. Chem. Phys.* **2013**, *13*, 3501–3515. [\[CrossRef\]](#)
51. Allen, D.R.; Hoppel, K.W.; Kuhl, D.D. Wind extraction potential from 4D-var assimilation of stratospheric O₃, N₂) and H₂O using a global shallow water model. *Atmos. Chem. Phys.* **2014**, *14*, 3347–3360. [\[CrossRef\]](#)
52. Allen, D.R.; Hoppel, K.W.; Kuhl, D.D. Wind extraction potential from ensemble Kalman filter assimilation of stratospheric ozone using a global shallow water model. *Atmos. Chem. Phys.* **2015**, *15*, 5835–5850. [\[CrossRef\]](#)
53. Bocquet, M.; Sakov, P. Joint state and parameter estimation with an iterative ensemble Kalman smoother. *Nonlinear Process. Geophys.* **2013**, *20*, 803–818. [\[CrossRef\]](#)
54. Haussaire, J.-M.; Bocquet, M. A low-order coupled chemistry meteorology model for testing online and offline data assimilation schemes: L95-GRS (v1.0). *Geosci. Model Dev.* **2016**, *9*, 393–412. [\[CrossRef\]](#)
55. Ménard, R.; Chabrilat, S.; Charette, C.; Gauthier, P.; de Grandpré, J.; Robichaud, A.; Rochon, Y.; Yan, Y. *Coupled Chemical-Dynamical Data Assimilation*; Final Report ESA/ESTEC Contract No. 18560/04/NL/FF; ESA/ESTEC: Noordwijk, The Netherlands, 2007; p. 486.
56. Gauthier, P.; Charette, C.; Fillion, L.; Koclas, P.; Laroche, S. Implementation of a 3D variational data assimilation system in the Canadian Meteorological Centre: Part 1: The global analys. *Atmos. Ocean* **1999**, *37*, 103–156. [\[CrossRef\]](#)
57. Lorenc, A.; Rawlins, F. Why does 4D-Vat beat 3D-Var? *Q. J. R. Meteorol. Soc.* **2005**, *131*, 3247–3257. [\[CrossRef\]](#)
58. Courtier, P. Dual formulation of four-dimensional variational assimilation. *Q. J. R. Meteorol. Soc.* **1997**, *123*, 2449–2461. [\[CrossRef\]](#)
59. Gauthier, P.; Tanguay, M.; Laroche, S.; Pellerin, S. Extension of 3DVAR to 4DVAR: Implementation of 4DVAR at the meteorological service of Canada. *Mon. Weather Rev.* **2007**, *135*, 2339–2354. [\[CrossRef\]](#)
60. McLinden, C.; Olsen, S.; Hannegan, B.; Wild, O.; Prather, M.; Sundet, J. Stratospheric ozone in 3-D models: A simple chemistry and cross-tropopause flux. *J. Geophys. Res.* **2000**, *105*, 14653–14665. [\[CrossRef\]](#)
61. de Grandpré, J.; Tanguay, M.; Qaddouri, A.; Zerroukat, M.; McLinden, C.A. Semi-Lagrangian Advection of Stratospheric Ozone on a Yin-Yang Grid System. *Mon. Weather Rev.* **2016**, *144*, 1035–1050. [\[CrossRef\]](#)
62. Errera, Q.; Ménard, R. Technical Note: Spectral representation of spatial correlations in variational assimilation with grid point models and application to the Belgian Assimilation System for Chemical Observations (BASCOE). *Atmos. Chem. Phys.* **2012**, *12*, 10015–10031. [\[CrossRef\]](#)
63. Polavarapu, S.; Ren, S.; Rochon, Y.; Sankey, D.; Ek, N.; Koshyk, J.; Tarasick, D. Data assimilation with the Canadian middle atmosphere model. *Atmos. Ocean* **2005**, *43*, 77–100. [\[CrossRef\]](#)

64. Derber, J.; Bouttier, F. A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus* **1999**, *51*, 195–221. [\[CrossRef\]](#)
65. Gilbert, J.C.; Lemaréchal, C. Some numerical experiments with variable-storage quasi-Newton algorithms. *Math. Program.* **1989**, *45*, 407–435. [\[CrossRef\]](#)
66. Boer, G. Homogenous and isotropic turbulence on the sphere. *J. Atmos. Sci.* **1983**, *40*, 154–163. [\[CrossRef\]](#)
67. Gauthier, P.; Courtier, P.; Moll, P. Assimilation of simulated wind lidar data with a Kalman filter. *Mon. Weather Rev.* **1992**, *121*, 1803–1820. [\[CrossRef\]](#)
68. Courtier, P.; Thépaut, J.-N.; Hollingsworth, A. A strategy for operational implementation of 4D-Var using an incremental approach. *Q. J. R. Meteorol. Soc.* **1994**, *120*, 1367–1387. [\[CrossRef\]](#)
69. Parrish, D.F.; Derber, J.C. The National Meteorological Center’s spectral statistical-interpolation analysis system. *Mon. Weather Rev.* **1992**, *120*, 1747–1763. [\[CrossRef\]](#)
70. Heckly, W.A.; Courtier, P.; Pailleux, J.; Andersson, E. The ECMWF Variational Analysis: General Formulation and Use of Background Information. In *ECMWF Workshop on Variational Assimilation, with Special Emphasis on Three-Dimensional Aspects*; European Centre for Medium-Range Weather Forecasts: Reading, UK, 1992; pp. 49–94.
71. Rabier, F.; McNally, A.; Andersson, E.; Courtier, P.; Undén, P.; Eyre, J.; Hollingsworth, A.; Bouttier, F. The ECMWF implementation of the three-dimensional variational assimilation (3D-Var). II: Structure functions. *Q. J. R. Meteorol. Soc.* **1998**, *124*, 1809–1829. [\[CrossRef\]](#)
72. Gauthier, P.; Buehner, M.; Fillion, L. *Background-Error Statistics Modelling in a 3D Variational Data Assimilation Scheme: Estimation and Impact on the Analyses*; Technical Report; ECMWF: Reading, UK, 1998.
73. Ménard, R.; Deshaies-Jacques, J. Evaluation of analysis by cross-validation, Part II: Diagnostic and optimization of analysis error covariance. *Atmosphere* **2018**, *9*, 70. [\[CrossRef\]](#)
74. Caines, P.E. *Linear Stochastic Systems*; John Wiley and Sons: New York, NY, USA, 1998; p. 874.
75. Laroche, S.; Gauthier, P. A validation of the incremental formulation of 4D variational data assimilation in a nonlinear barotropic flow. *Tellus* **1998**, *50*, 557–572. [\[CrossRef\]](#)
76. Lawless, A.S.; Nichols, N.K.; Boess, C.; Bunse-Gerstner, A. Using model reduction methods within incremental four-dimensional variational data assimilation. *Mon. Weather Rev.* **2008**. [\[CrossRef\]](#)
77. Polavarapu, S.; Tanguay, M.; Ménard, R.; Staniforth, A. The tangent linear model for semi-Lagrangian schemes: Linearizing the process of interpolation. *Tellus* **1996**, *48*, 74–95. [\[CrossRef\]](#)
78. Tanguay, M.; Polavarapu, S. The adjoint of the semi-Lagrangian treatment of the passive tracer equation. *Mon. Weather Rev.* **1999**, *127*, 551–564. [\[CrossRef\]](#)
79. Rutherford, I.D. Data assimilation by statistical interpolation of forecast error fields. *J. Atmos. Sci.* **1972**, *29*, 809–815. [\[CrossRef\]](#)
80. Hollingsworth, A.; Lönnberg, P. The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus* **1986**, *38*, 111–136. [\[CrossRef\]](#)
81. Desroziers, G.; Berre, L.; Chapnik, B.; Poli, P. Diagnosis of observation-, background-, and analysis-error statistics in observation space. *Q. J. R. Meteorol. Soc.* **2005**, *131*, 3385–3396. [\[CrossRef\]](#)
82. Desroziers, G.; Ivanov, S. Diagnosis and adaptive tuning of observation-error parameters in a variational assimilation. *Q. J. R. Meteorol. Soc.* **2001**, *127*, 1433–1452. [\[CrossRef\]](#)
83. Ménard, R. Error covariance estimation methods based on analysis residuals: Theoretical foundation and convergence properties derived from simplified observation networks. *Q. J. R. Meteorol. Soc.* **2016**, *142*, 257–273. [\[CrossRef\]](#)
84. Bouttier, F. Sur la prévision de la qualité des prévisions météorologiques. Ph.D. Thesis, Université Paul Sabatier, Toulouse, France, 1994; p. 240.
85. Janjić, T.; Bormann, N.; Bocquet, M.; Carton, J.A.; Cohn, S.E.; Dance, S.L.; Losa, S.N.; Nichols, N.K.; Potthast, R.; Waller, J.A.; et al. On the representation error in data assimilation. *Q. J. R. Meteorol. Soc.* **2017**, *144*, 1257–1278. [\[CrossRef\]](#)
86. Pereira, M.B.; Berre, L. The use of an ensemble approach to study the background error covariance in a global NWP model. *Mon. Weather Rev.* **2006**, *134*, 2466–2489. [\[CrossRef\]](#)
87. Ménard, R.; Deshaies-Jacques, M.; Gassett, N. A comparison of correlation-length estimation methods for the objective analysis of surface pollutants at Environment and Climate Change Canada. *J. Air Waste Manag. Assoc.* **2016**, *66*, 874–895. [\[CrossRef\]](#)

88. Robichaud, A.; Ménard, R.; Chabrilat, S.; de Grandpré, J.; Rochon, Y.J.; Yang, Y.; Charrette, C. Impact of energetic particle precipitation on stratospheric polar constituents: An assessment using monitoring and assimilation of operational MIPAS data. *Atmos. Chem. Phys.* **2010**, *10*, 1739–1757. [[CrossRef](#)]
89. Riihshøjgaard, L.P.; Källén, E. On the correlation between ozone and potential vorticity for large scale Rossby waves. *J. Geophys. Res.* **1997**, *107*, 8793–8804. [[CrossRef](#)]
90. Li, Y.; Ménard, R.; Riishøjgaard, L.P.; Cohn, S.E.; Rood, R.B. A study on assimilating potential vorticity data. *Tellus* **1998**, *50*, 490–506. [[CrossRef](#)]
91. Allaart, M.A.F.; Kelder, H.; Heijboer, L.C. On the relation between ozone and potential vorticity. *Geophys. Res. Lett.* **1993**, *20*, 811–814. [[CrossRef](#)]
92. Ménard, R. Bias Estimation. In *Data Assimilation: Making Sense of Observations*; Lahoz, W., Khattatov, B., Ménard, R., Eds.; Springer: New York, NY, USA, 2010; pp. 113–136.
93. Di Tomaso, E.; Bormann, N. Assimilation of ATOVS Radiances at ECMWF: First Year EUMETSAT Fellowship Report. In *EUMETSAT/ECMWF Fellowship Programme Research Report 22*; ECMWF: Reading, UK, 2010; p. 27.
94. Thépaut, J.-N.; Courtier, P. Four-dimensional variational data assimilation using the adjoint of a multilevel primitive-equation model. *Q. J. R. Meteorol. Soc.* **1991**, *117*, 1225–1254. [[CrossRef](#)]
95. Ménard, R.; Polavarapu, S.; Yang, Y. Model error estimation: Its application to chemical data assimilation. In *Proceedings of the ECMWF/SPARC Workshop, Shinfield Park, UK, 23–26 June 2003*.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).