Peer-reviewed version available at *Atmosphere* **2019**, *10*, 798; doi:10.3390/atmos10120798

Article

Coupled stratospheric chemistry-meteorology data assimilation. Part II: Weak and strong coupling

Richard Ménard[1*], Pierre Gauthier[4], Yves Rochon[1], Alain Robichaud[1], Jean de Grandpré[1], Yan Yang[1], Cécilien Charrette[3], and Simon Chabrillat[2]

[1] Air Quality Research Division, Environment and Climate Change Canada
[2] Belgium Institute for Space Aeronomy, Brussels, Belgium
[3] Meteorological Research Division, Environment and Climate Change Canada
[4] –Département des sciences de la Terre et de l'atmosphère, Université du Québec à Montréal, Canada.
* Correspondence: richard.menard@canada.ca; Tel.: +1-514-421-4613,
2121 Transcanada Highway, Dorval, (QC), CANADA, H9P 1J3

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 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. Also, 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 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 observations available and these are mostly related to temperature, however, there are several research satellites measuring chemical composition in this region [1]. Important missions began in the early 1990's 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. Since 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. Here we will discuss how these research satellite observations can provide useful information. Also we will present weak and strong data assimilation coupling experiments.

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 [10-11]. In weakly coupled data assimilation system, each geophysical component (e.g. chemistry, meterology) 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 throught 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 [12-22] and atmosphere-land-surface coupled systems [23-27].

Coupled meteorology-chemistry data assimilation has primarily been examined in the context tropospheric aerosol-radiation interaction on short time-scales [28-30], on subseasonal prediction [31] and decadal time scales [32] (also see [33] for a review of chemical data assimilation). It was also used to estimate parameters in the activation of aerosols into cloud droplets [34], and in determining cross-covariance between temperature and constituents (O_3, NO_2 and SO_2) using the coupled tropospheric model WRF-CHEM and an ensemble based approach [35]. 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 [36] and a CO_2 assimilation [37],

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 [38] and at the Canadian Meteorological Center (CMC) with the experimental model GEM-BACH [39]. The experiments conducted at ECMWF 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 showed an impact on forecast in the lower stratosphere predictability.

Strong data assimilation coupling has been considered in the context of using chemical tracer observations to infer winds. In some of the earlier studies using an extended Kalman filter with a simplified two-dimensional transport model, it was shown that wind recovery is very sensitive to the accuracy of chemical observations, and to the concentration fields having sufficient horizontal gradients and small data voids [40]. It was also shown that constituents in zones of

93 convergence could only determine the winds nearby. Experiments conducted with a one-
94 dimensional model also showed that wind information can still be obtained in the case of a flat
95 concentration field if there are gradients in the concentration error covariances [41]. Using the
96 barotropic vorticity equation with a 4D-Var assimilation system, Riishøjgaard [42] examined the
97 issues of data density and length of the assimilation window, and arrived at similar conclusions.
98 Using column measurements of ozone with a NWP model and 4D-Var method, a small
99 improvement in the winds was obtained using simulated observations, but a deterioration using
100 real observations [43]. The negative impact was suspected to be result from observational bias.
101 In another study using an operational NWP model with a 4D-Var assimilation system, a small
102 impact (about 0.1 ms^{-1}) was found on zonal wind with no reduction of error standard deviation
103 [44]. These unsuccessful results conducted in an operational context suggested that additional
104 studies were necessary. Using an ensemble Kalman filter and an intermediate-complexity
105 model, Milewski and Bourqui [45] demonstrated that information about the ozone-wind cross-
106 covariance is essential in constraining dynamical fields when ozone only is assimilated.
107 Moreover they showed that a further reduction in error can be obtained with an Ensemble Kalman
108 smoother [46]. In a series of studies using 4D-Var and ensemble Kalman filter, Allen et al [47-49]
109 showed that poorly-specified observation error could lead to an increase in RMS wind error, also
110 that observational coverage is important wind so that wind extraction could be improved if
111 several chemical tracers were used, and that the balance between wind and temperature could be
112 offset by the wind recovery from tracer measurements. We should note that the wind extraction
113 from tracer observation is part of a more general class of joint state-parameter estimation
114 problems (e.g. [50-51]).

115 The present study took place in the period 2005-2009 with funding, in part, from ESA/ESTEC
116 [52]. This article, henceforth referred to as Part II summarized the data assimilation aspects.
117 First, we present the extension of the CMC variational assimilation system to include chemical
118 variables in Section 2, where we discuss in particular the analysis splitting and preconditioning,
119 the extension of balance operators with chemical variables, and the validity of an incremental
120 formulation of adjoint tracer operators for 4D-Var assimilation of long-lived species. In Section
121 3 we describe the error statistics of chemical variables using the Hollingsworth-Lönnberg method
122 to estimate the error variances, and using the Canadian Quick Covariance (CQC) method to
123 obtain non-separable error correlations. We also discuss the method and issues related to the
124 cross-covariances between temperature and ozone. In Sections 4 and 5 we illustrate the benefits
125 of using limb sounding temperatures from MIPAS to improve the AMSU-A bias correction and

better simulation of temperature and transport in the stratosphere. We then discuss weak coupling data assimilation experiment involving ozone and its impact on meteorological forecasts, which indicates that a simplified linearized ozone chemistry is sufficient to obtain most of the desired results. 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 ued here consists of a model integration step to obtain a 6-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 6-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 6-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 [53].
- 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 [54] (actually closest to a 1-hour bin).
- 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) [55-56].

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, called the *analysis*, leads 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 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 ozone-radiation interaction and no chemistry. No physical parameterization is used in the adjoint model except for diffusion and a linearized planetary boundary layer scheme. For 4D-Var long-lived species chemical assimilation experiments the adjoint model is that of passive tracer advection.

Background error covariances and observation error covariances are needed to compute the minimum variance estimate. The background error correlation model used in this study for each variable is 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 either from a linearized photochemical model such as LINOZ [57] (see [58] for its implementation with semi-Lagrangian transport), or from correlations in the model output.

Applying the above variable transformations and representing the error correlations in spherical harmonics, it is possible to completely diagonalize the error correlation matrix [59]. The covariances in physical space and between all variables can then be obtained through a series of transformations on a vector. With this formulation and a reduction due to spectral transform and vertical eigenfunctions, the background error covariance matrix can be, in principle, easily expanded to include other variables in addition to the dynamical 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 [60] 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+4 = 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 10 three-dimensional fields (horizontal winds, temperature, water vapor, O_3, CH_4, NO_2, N_2O, HNO_3 and $ClONO_2$). As far as unobserved

constituents that have a background correlated 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 prognostic meteorological and comprises all the prognostic chemical model variables, which, in our case, represents more than 61 three-dimensional fields (the chemistry model alone accounts for 57 advected species). In principle, a state estimate should be conducted on all prognostic variables. Yet, only a small fraction of these variables is observed. For example, MIPAS/ESA 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.

$$\mathbf{Z} = \begin{pmatrix} \mathbf{X} \\ \mathbf{U} \end{pmatrix}. \tag{1}$$

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

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

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,

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{\mathrm{XX}} & \mathbf{B}_{\mathrm{XU}} \\ \mathbf{B}_{\mathrm{UX}} & \mathbf{B}_{\mathrm{UU}} \end{pmatrix}, \tag{3}$$

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

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

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

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

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 in presented in Appendix A).

Analysis splitting concept is interesting and practical for a number of reasons. Consider the behavior of unobserved variables $\mathbf{U}$ in either a 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, $\mathbf{X}^a = (\boldsymbol{\mu}^a, \boldsymbol{\chi}^a)^T$, 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 [53] and extended to include chemical variables [60] was further extended in this study to include cross-covariances between observed species and between observed and unobserved variables using a balance operator. The general

framework will be explained in Section 2.3 and the associated error statistics in Section 3.3. Cross-covariances involving chemical variables was estimated point-wise, while the meteorological variable error covariances (and cross-covariances) were computed in spectral space as in by Derber and Bouttier [61].

The coupled chemical-meteorological model state used in the 3DVar-Chem in Equation (4) consists of $\mathbf{X} = [\psi,\chi,\mathbf{T},\ln(\mathbf{q}),\mathbf{c}_1,\ldots,\mathbf{c}_N,\mathbf{p}_s]^T$, where ψ is the streamfunction, χ the velocity potential, $\mathbf{T}$ the temperature, $\mathbf{q}$ the (tropospheric) water vapor mixing ratio, $\mathbf{p}_s$ the surface pressure and N observed tracers, or chemical constituent mixing ratios $\mathbf{c}_1$, …, $\mathbf{c}_N$. The state vector in 3D Var-Chem is such that all 3D fields are grouped together, followed by the 2D field $\mathbf{p}_s$. 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 $\mathbf{B}_{xx} = \mathbf{L}\,\mathbf{L}^T$, 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 [62]. The preconditioning used in 3DVar-Chem follows what is done for the meteorological variables [53]. 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 are assumed to be homogeneous and isotropic on the sphere. For such correlations, the spectral representation is diagonal in spectral space (see for example [63, 64,53,59]). The sequence of operations then becomes as follows: 1 – We multiply the spectral representation of the state with the square root of the spectral coefficient of the correlation model,

2 – Perform a transform from spectral to physical space, 3 – Multiply the resulting fields by the error variances, and 4 – Using balance operators, 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.

Before we discuss the balance operators, we should note two things: 1- The CMC 3D-Var system uses a non-separable error correlation model. It means that for each horizontal wavenumber there is a unique vertical correlation matrix, which introduces a dependence between horizontal and vertical scales; 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. [53] and in 4D-Var is called an incremental formulation [65]), 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) [53,66-68] or through statistical regression [69,61]. For coupled meteorology-chemistry the multilinear regression approach [69] can easily be extended to include chemical species. In particular the streamfunction $\delta\boldsymbol{\psi}$, velocity potential $\delta\boldsymbol{\chi}$, temperature $\delta\mathbf{T}$, and ozone $\delta\mathbf{O}_3$ which are known to be correlated, can be transformed into a set of *uncorrelated* background error variables (denoted with a superscript u), as follows

$$
\begin{aligned}
\delta\boldsymbol{\psi}^u &= \delta\boldsymbol{\psi} \\
\delta\boldsymbol{\chi}^u &= \delta\boldsymbol{\chi} - \mathbf{B}_{\boldsymbol{\chi}\boldsymbol{\psi}}\mathbf{B}_{\boldsymbol{\psi}\boldsymbol{\psi}}^{-1}\,\delta\boldsymbol{\psi} \\
\delta\mathbf{T}^u &= \delta\mathbf{T} - \mathbf{B}_{\mathbf{T}\boldsymbol{\psi}}\mathbf{B}_{\boldsymbol{\psi}\boldsymbol{\psi}}^{-1}\,\delta\boldsymbol{\psi} - \mathbf{B}_{\mathbf{T}\boldsymbol{\chi}^u}\mathbf{B}_{\boldsymbol{\chi}^u\boldsymbol{\chi}^u}^{-1}\,\delta\boldsymbol{\chi}^u \\
\delta\mathbf{O}_3^u &= \delta\mathbf{O}_3 - \mathbf{B}_{\mathbf{O}_3\boldsymbol{\psi}}\mathbf{B}_{\boldsymbol{\psi}\boldsymbol{\psi}}^{-1}\,\delta\boldsymbol{\psi} - \mathbf{B}_{\mathbf{O}_3\boldsymbol{\chi}^u}\mathbf{B}_{\boldsymbol{\chi}^u\boldsymbol{\chi}^u}^{-1}\,\delta\boldsymbol{\chi}^u - \mathbf{B}_{\mathbf{O}_3\mathbf{T}^u}\mathbf{B}_{\mathbf{T}^u\mathbf{T}^u}^{-1}\,\delta\mathbf{T}^u \,.
\end{aligned}
\tag{6}
$$

292 The transformation from any set of correlated errors to uncorrelated error variables, as in
293 Equation (6), can also be explained geometrically by adopting a Hilbert space representation of
294 the random variables [70,71] and followed by Gram-Smidt orthogonalization (see Appendix B for
295 this geometrical derivation).
296
297 Backsubstituting, we recover the transformation from uncorrelated variables to correlated
298 variables, in the form

299
$$\begin{pmatrix} \delta\boldsymbol{\psi} \\ \delta\boldsymbol{\chi} \\ \delta\mathbf{T} \\ \delta\mathbf{O}_3 \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{E} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{N} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{G} & \mathbf{0} & \mathbf{F} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \delta\boldsymbol{\psi}^u \\ \delta\boldsymbol{\chi}^u \\ \delta\mathbf{T}^u \\ \delta\mathbf{O}_3^u \end{pmatrix} = \mathbf{M} \begin{pmatrix} \delta\boldsymbol{\psi}^u \\ \delta\boldsymbol{\chi}^u \\ \delta\mathbf{T}^u \\ \delta\mathbf{O}_3^u \end{pmatrix}, \tag{7}$$

300 where

301
$$\mathbf{E} = \mathbf{B}_{\chi\psi}\mathbf{B}_{\psi\psi}^{-1} \qquad \mathbf{N} = \mathbf{B}_{\mathrm{T}\psi}\mathbf{B}_{\psi\psi}^{-1} \qquad \mathbf{G} = \mathbf{B}_{\mathrm{O}_3\psi}\mathbf{B}_{\psi\psi}^{-1} \qquad \mathbf{F} = \mathbf{B}_{\mathrm{O}_3\mathrm{T}^u}\mathbf{B}_{\mathrm{T}^u\mathrm{T}^u}^{-1} \tag{8}$$

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

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

312
$$\mathbf{B}_{\mathrm{XX}} = \mathbf{M} \begin{pmatrix} \mathbf{B}_{\psi\psi} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{\chi^u\chi^u} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{B}_{\mathrm{T}^u\mathrm{T}^u} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{B}_{\mathrm{O}_3^u\mathrm{O}_3^u} \end{pmatrix} \mathbf{M}^T = \mathbf{M}\boldsymbol{\Sigma}_{\mathrm{XX}} \begin{pmatrix} \mathbf{C}_{\psi\psi} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{\chi^u\chi^u} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{C}_{\mathrm{T}^u\mathrm{T}^u} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{C}_{\mathrm{O}_3^u\mathrm{O}_3^u} \end{pmatrix} \boldsymbol{\Sigma}_{\mathrm{XX}}^T \mathbf{M}^T \tag{9}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2019 doi:10.20944/preprints201910.0241.v1

Peer-reviewed version available at Atmosphere 2019, 10, 798; doi:10.3390/atmos10120798

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

Finally, the implementation of balance operators in a 3D-Var or 4D-Var using preconditioning requires the inverse of the square root of $\mathbf{B}_{xx}$, and thus we need to know the inverse of $\mathbf{M}$, which turns out to be easy to obtain as

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

2.4 4D-Var tracer extension

The 3D-Var algorithm can be extended to 4D-Var by including the model integration as part of the observation operator [55,56]. 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 [65] 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 [65,72,73], called the inner-loop and where its solution is used to update the full model trajectory in an outer-loop. The outer-loop trajectories defines 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° x 1.5°, 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 [74] and the properties of the adjoint in Tanguay and Polavarapu [75]. 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 it behaves as a passive tracer. To illustrate the validity of the incremental tracer approach for ozone, Figure 1 shows the observation cost function Jo 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 outerloop, during the second inner loop. We observe a nearly continuous decrease in the observation cost function except for a small adjustment between the last iterate of the first inner loop and the beginning of the second inner loop.

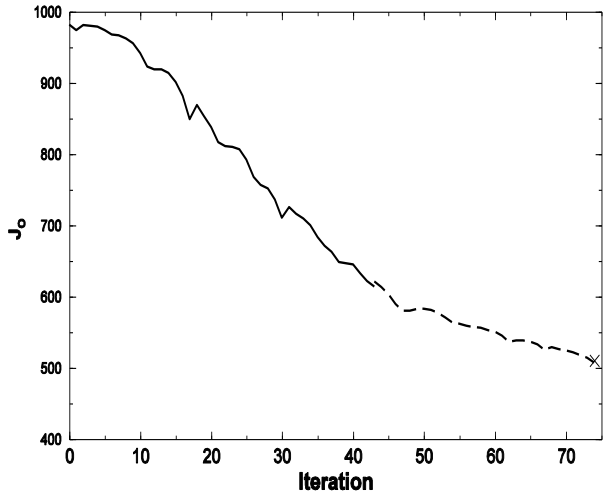


Figure 1 Observation component of the cost function for ozone assimilation as a function of iteration. Solid line is associated with the value of J_o of the first inner loop and the dashed line the value of J_o of the second inner loop.

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 [76,77] does precisely this and is based on computing the distance between pair of observations that fit well with geographically fixed observations. Here, we demonstrate that this method can also be used with a polar orbiting limb sounder such as MIPAS, because the distance between observation profiles is uniform. 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 based on innovations that can provide observation and background error statistics, such as Desroziers [78] and Desroziers and Ivanov [79], but these are based on different assumptions (see [80]).

Any of the innovation-based 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 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 parameters. The lagged-forecast method, widely used in meteorology, is based on having a complete observations coverage. Bouttier [81] has 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 a large region where there are no observations, as the difference between the forecast and analysis error covariances is close to zero.

Also, 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 observational source is a single polar orbiting satellite, i.e. MIPAS. The horizontal coverage of MIPAS in 6 hours (analysis time window) is limited to about a quarter or 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 6-hour differences of a single

model output. This method, originally developed by Yves Rochon (personal communication) is known as the Canadian Quick Covariance (CQC) [60].

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	6-hour difference (CQC) method

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.

3.1 Estimation of error variances by autocorrelation of innovations along the satellite track

The observation error obtained from innovations comprises: the instrument error, the forward modeling and retrieval errors, the error due to the interpolation from observation location to model grid point, and the representativeness error due in part to the subgrid scale variability not resolved by the atmospheric model [82]. The model forecast error is generally correlated horizontally over large distances, typically 500-1000 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 intercept of the spatially correlated part of the innovation 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.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2019 doi:10.20944/preprints201910.0241.v1

Peer-reviewed version available at Atmosphere 2019, 10, 798; doi:10.3390/atmos10120798

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 63hPa.

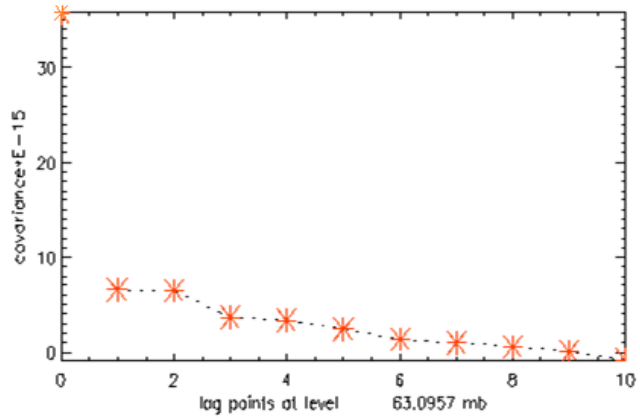


Figure 2 Spatial autocovariance of innovation for MIPAS CH_4 at 63 hPa. Abscissa are units of horizontal separation between profiles (each unit is around 530 km). The red stars represent the sample autocovariance values, and the dashed curve are linear interpolation between the sample points. Note that at zero separation the sample covariance is at the top of the graph (near 36 $x10^{-15}$), and no interpolation between the zero distance and lag-1 is done.

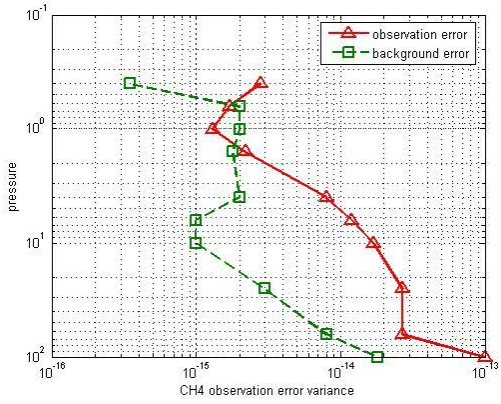
The units of the horizontal axis are profile lag points along the satellite track, with spacing of 530 km. We note that at zero separation the innovation variance is $36x10^{-15}$ and is distinctively different from the extrapolated intercept of the spatially correlated part, estimated to be around $8x10^{-15}$. 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.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2019 doi:10.20944/preprints201910.0241.v1

Peer-reviewed version available at Atmosphere 2019, 10, 798; doi:10.3390/atmos10120798

The estimates of observation and background error variances for MIPAS CH_4, obtained from HL method are displayed in the left panel of Figure 3. We note that the MIPAS CH_4 observation error variance is significantly larger than the (model) background error except in the region between 2 and 0.5 hPa. This indicates that it is not everywhere that MIPAS CH_4 observations will have an impact on the analysis, the main impact region is limited to 2 to 0.5 hPa, and perhaps also 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 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 [78] is shown in green and is close to the HL estimate in the mid-to-lower stratosphere from 100 hPa to 3hPa. However, at higher altitudes important differences are noted. Since the Desroziers method [78] relies on the assumption that the Kalman gain is nearly optimal (i.e. close to the truth) [80], explains the discrepancy between HL and Desroziers estimates of observation error variance.



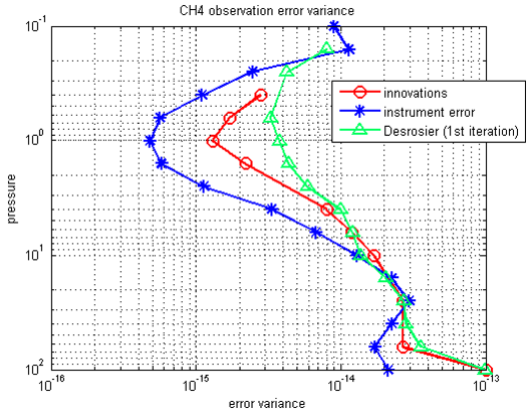


Figure 3 Estimated error variance for CH_4/MIPAS as a function of height. Left panel, shows the estimated background error variance (green with squares) and observation error variance (red with triangles) as a function of height using the HL method. Right panel illustrates three different estimates of observation error variance. Blue with stars is the estimate given by the instrument team (i.e. the instrument error), red with circles is the observation error variance obtained from the HL method (note that this is the same as the red triangle in the left panel), and green with triangles is the observation error variance estimate from the Desroziers method [78].

One way to summarize the estimated error variances 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 O_3, CH_4, N_2O, NO_2, HNO_3 and H_2O that were assimilated in the course of this study. We note for instance that for O_3 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 CH_4 and N_2O. However, the NO_2 gain is close to one in the upper stratosphere, indicating that the model as a small impact at these altitudes. As for HNO_3, the gain increases with height and reaches a maximum value of 0.8 at 4hPa, then decreases with altitude. Chemical water vapor (H_2O) is presented in terms of the log of concentration.

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 [59] for a description on the use of spherical harmonics and how to construct error correlations, and Gauthier [69] for aspects related to meteorological applications). The Canadian Quick Covariance (CQC) method [60] is similar to the NMC method except that it uses 6-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. Writing a model equation in the form,

$$\frac{\partial \psi}{\partial t} + \mathbf{V} \cdot \nabla \psi = F(\psi), \tag{11}$$

the CQC method thus derives its spatial error statistics from the 6-hour differences which represent model tendencies,

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

It has been argued [60] that since the large-scale motion doesn't change in a 6hr 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. But, it is known that the NMC method has a tendency to give too much spectral error variance at these wavenumbers for meteorological correlations fields [83]. It is thus unclear whether the CQC method has an actual deficiency at large scales. The latitudinal power spectra of the species that were used in the assimilation are shown in Figure S3 (Supplementary Material) and indicate generally a maximum in power at the large scales (wavenumber 8-10) as one would expect.

To compute the background error correlation with the CQC method we first need to compute the variance of 6-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 6-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 in the NMC method, from which by using the spectral representation of a non-separable correlation model we obtain, for each horizontal wave number n, a vertical correlation matrix $nlev \times nlev$ (see [59] or [53] for details). In a non-separable correlation model, we can compute a power spectrum as a function of the horizontal wavenumber n and vertical level, shown is illustrated in Figure S3 (Supplementary Material). A horizontal-vertical separable correlation model has a horizontal power spectrum that does not change with height. The results shown in Figure S3 indicate that for most chemical species the correlation is

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2019 doi:10.20944/preprints201910.0241.v1

Peer-reviewed version available at Atmosphere 2019, 10, 798; doi:10.3390/atmos10120798

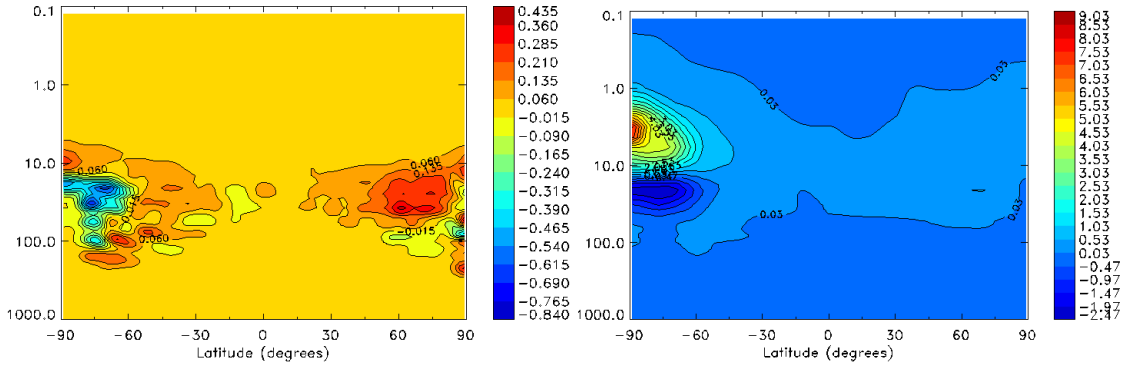
519 horizontal-vertical separable except for O_3, HNO_3 and H_2O at large scales (for wavenumbers
520 smaller than 20).

521 The resulting correlation length can also be computed. Figure S5 of Supplementary Material,
522 shows the horizontal correlation length for six constituents, that typically varies from 200 km (in
523 the troposphere) to 400 km in the upper stratosphere. These correlation length-scales seem to be
524 too small if we visually compare the correlation length we get from the spatial autocovariance of
525 innovations. Figure 2 shows a decorrelation length scale of 2 or 3 orbits, which corresponds to
526 about 1000 to 1500 km. Despite the fact the HL method has a tendency in practice to overestimate
527 the spatial correlation length scale compared with length-scale obtained with the maximum
528 likelihood method [84], the correlation length scale obtained with the CQC method for chemical
529 constituents seem too small. However, since we have no means to correct for these deficiencies,
530 we continue to use the spectral coefficients as is in the correlation models of the chemical
531 constituents.

532 The background error covariance is then obtained by using the background error variance
533 estimated by the HL method to the correlation estimated using the CQC method. Thus, the
534 background error variance is identical at all latitudes and longitudes and varies as a function of
535 height only. We conducted a series of univariate constituent data assimilation experiments,
536 using the background error covariance above and the observation error obtained from the HL
537 method and computed the mean analysis increment over the period of August 17 to September
538 5, 2003. During this time period a strong energetic particle precipitation from the mesosphere
539 affected the polar region down to the middle stratosphere and created large NO_2 and HNO_3
540 mixing ratio increments [85]. Figure 4 presents the mean analysis increment for HNO_3.

541

542 **HNO_3 old statistics** **HNO_3 new statistics**



543

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2019 doi:10.20944/preprints201910.0241.v1

Peer-reviewed version available at Atmosphere 2019, 10, 798; doi:10.3390/atmos10120798

Figure 4 Analysis increment for HNO_3. 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} vmr.

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 [85], while the old statistics give random results with numerous small-scale features. The analysis increments for chemically active species such as O_3 and NO_2 also appear to be larger and also 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 (see Figures S6 in Supplementary Material).

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 take place 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 6-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 6-hour differences over a month (assuming here 30 days), the cross-covariance is obtained as

$$\mathbf{B_{uv}} = \frac{1}{4*30}\sum_{i=1}^{4*30}\left\langle\left[(\mathbf{u}_{t(i)+6}-\mathbf{u}_{t(i)})-\overline{(\mathbf{u}_{t(i)+6}-\mathbf{u}_{t(i)})}\right]\left[(\mathbf{v}_{t(i)+6}-\mathbf{v}_{t(i)})-\overline{(\mathbf{v}_{t(i)+6}-\mathbf{v}_{t(i)})}\right]^T\right\rangle = \mathbf{\Sigma_u}\,\mathbf{C_{uv}}\,\mathbf{\Sigma_v^T} \quad (13)$$

where $\mathbf{\Sigma_u}$, $\mathbf{\Sigma_v}$ are diagonal matrices of error standard deviations of the 6-hour differences, and the index i is for the four 6-hour time periods in a day However, $\mathbf{C_{uv}}$ is a full 6-dimensional matrix and needs to be significantly simplified to be computed from statistics. We generally represented it as a zonal field of point correlations, thus neglecting the horizontal and vertical

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2019 doi:10.20944/preprints201910.0241.v1

Peer-reviewed version available at Atmosphere 2019, 10, 798; doi:10.3390/atmos10120798

correlations. It was found the neglecting the vertical correlation has a small impact on the zonal-mean representation of **B**.

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 is very similar (result not shown). As discussed previously, the region above the 10hPa is photochemistry-dominated, while below 10hPa 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 6-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 6-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 difference between interactive and non-interactive runs in the northern hemisphere between 10 and 100 hPa is slightly larger if instead of 6 -hour differences we use 24-hour differences to derive the cross-correlation (Figure S8, Supplementary Material).

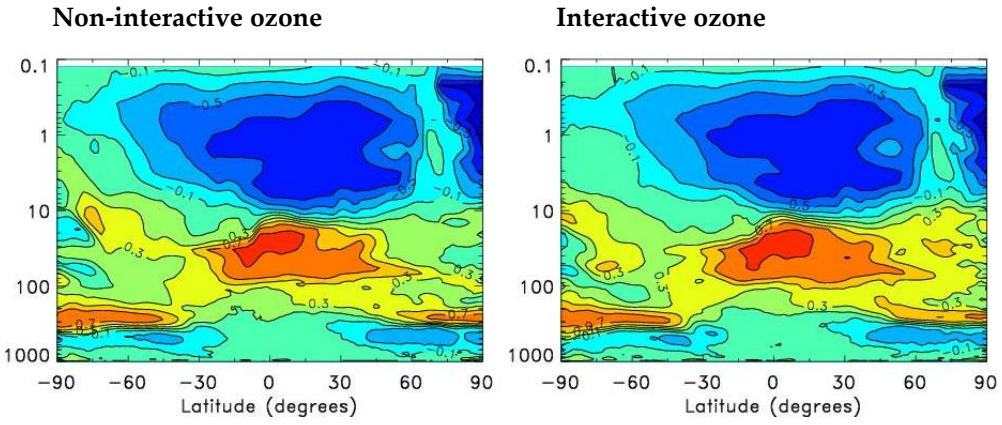


Figure 5 Cross-correlation between ozone and temperature derived from 6-hr differences (i.e. CQC method) for July 2003. Left panel refers to a non-interactive ozone-radiation run of GEM-BACH and right panel for an interactive ozone-radiation run.

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 hour) temperature effects on ozone transport. Large positive correlations are observed near the NH and SH tropopause and in the equatorial region around 20-70 hPa.

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 6-hour model differences, and would require a sequential reprocessing respecting the Gram-Smith orthogonalization of model differences, which we have not attempted to do here. Instead, we used an approximation. We recall that what needs to be computed is $\mathbf{F} = \mathbf{B}_{\mathbf{O_3T^u}} \mathbf{B}^{-1}_{\mathbf{T^uT^u}}$ but what we have readily available from the statistics is $\mathbf{A} = \mathbf{B}_{\mathbf{O_3T}} \mathbf{B}^{-1}_{\mathbf{TT}}$. We will approximate **F** by **A** in the following way.

To understand the approximation, we first note that,

$$\mathbf{B}_{\mathbf{TT}}\mathbf{A} = \mathbf{B}_{\mathbf{O_3T}} = \left\langle \delta\mathbf{O}_3 (\delta\mathbf{T})^T \right\rangle = \left\langle \delta\mathbf{O}_3 (\delta\mathbf{T}^u + \mathbf{B}_{\mathbf{T\psi}}\mathbf{B}^{-1}_{\mathbf{\psi\psi}}\delta\boldsymbol{\psi})^T \right\rangle$$
$$= \left\langle \delta\mathbf{O}_3 (\delta\mathbf{T}^u)^T \right\rangle + \left\langle \delta\mathbf{O}_3 (\delta\boldsymbol{\psi})^T \right\rangle \mathbf{B}^{-1}_{\mathbf{\psi\psi}}\mathbf{B}^T_{\mathbf{T\psi}} = \mathbf{B}_{\mathbf{O_3T^u}} + \mathbf{B}_{\mathbf{O_3\psi}}\mathbf{B}^{-1}_{\mathbf{\psi\psi}}\mathbf{B}^T_{\mathbf{T\psi}} \quad , \quad (14)$$

where we used Equation (6). The correlation between O_3 as a tracer and the streamfunction relates to the tracer-wind coupling discussed in Section 2.3 Part I. It has long been an elusive goal to obtain [86-88] (see also discussion in Section 7). It was argued that in regions of Rossby wave breaking activity, that potential vorticity is correlated with ozone as a chemical tracer in the lower stratosphere. Figure S7 (Supplementary Material) shows scatter plots of O_3 concentration and streamfunction 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 that globally the correlation between O_3 and streamfunction can be neglected*. This also implies that the balance operator **G** (Equation 8) can be neglected. Regarding Equation (5) we thus make the approximation that,

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

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

$$\mathbf{B}_{\mathbf{TT}} = \left\langle \delta\mathbf{T}(\delta\mathbf{T})^T \right\rangle = \left\langle (\delta\mathbf{T}^u + \mathbf{B}_{T\psi}\mathbf{B}_{\psi\psi}^{-1}\delta\boldsymbol{\psi})(\delta\mathbf{T}^u + \mathbf{B}_{T\psi}\mathbf{B}_{\psi\psi}^{-1}\delta\boldsymbol{\psi})^T \right\rangle = \mathbf{B}_{\mathbf{T}^u\mathbf{T}^u} + \mathbf{B}_{\mathbf{T}\boldsymbol{\psi}}\mathbf{B}_{\boldsymbol{\psi}\boldsymbol{\psi}}^{-1}\mathbf{B}_{\mathbf{T}\boldsymbol{\psi}}^T \quad . \tag{16}$$

By using the matrix inversion lemma, we then obtain its inverse as,

$$\mathbf{B}_{\mathbf{TT}}^{-1} = \mathbf{B}_{\mathbf{T}^u\mathbf{T}^u}^{-1} - \mathbf{B}_{\mathbf{T}^u\mathbf{T}^u}^{-1}\mathbf{B}_{\mathbf{T}\boldsymbol{\psi}}\left(\mathbf{B}_{\boldsymbol{\psi}\boldsymbol{\psi}} + \mathbf{B}_{\mathbf{T}\boldsymbol{\psi}}^T\mathbf{B}_{\mathbf{T}^u\mathbf{T}^u}^{-1}\mathbf{B}_{\mathbf{T}\boldsymbol{\psi}}\right)^{-1}\mathbf{B}_{\mathbf{T}\boldsymbol{\psi}}^T\mathbf{B}_{\mathbf{T}^u\mathbf{T}^u}^{-1} \quad . \tag{17}$$

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

$$\mathbf{B}_{\mathbf{TT}}^{-1} \approx \mathbf{B}_{\mathbf{T}^u\mathbf{T}^u}^{-1} \tag{18}$$

to compute the balance operator. Thus with the approximations in Equation (21) and the limitations imposed by the statistics (as discussed at the beginning of this section) the balance operator **F** (Equation 8) which is a function of latitude and pressure only, is approximated as

$$\mathbf{F}(\lambda, p) = \mathbf{B}_{\mathbf{O}_3\mathbf{T}^u}\mathbf{B}_{\mathbf{T}^u\mathbf{T}^u}^{-1} \approx \mathbf{A}(\lambda, p) = \mathbf{B}_{\mathbf{O}_3\mathbf{T}}\mathbf{B}_{\mathbf{TT}}^{-1}. \tag{19}$$

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

$$\mathbf{B}_{\mathbf{O}_3\mathbf{O}_3} = \mathbf{B}_{\mathbf{O}_3^u\mathbf{O}_3^u} + \mathbf{F}\mathbf{B}_{\mathbf{T}^u\mathbf{T}^u}\mathbf{F}^T \approx \mathbf{B}_{\mathbf{O}_3^u\mathbf{O}_3^u} + \mathbf{A}(\lambda, p)\mathbf{B}_{\mathbf{T}^u\mathbf{T}^u}\mathbf{A}^T(\lambda, p). \tag{20}$$

To construct the operator **A** we use the cross-correlation obtained from point-wise statistics derived from the CQC method for ozone-temperature cross-correlation, and point-wise statistics derived from the NMC method for temperature covariance. This contrasts with the balance operators introduced by Derber and Bouttier [61] 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 **A** are dependent on latitude and pressure (the hybrid vertical coordinate to be precise). We have investigated the use of a vertical correlation (but not horizontal correlation) in the operator **A** and observed little difference (results not shown). However, the important aspect is which error variances are considered to obtain the cross-covariance $\mathbf{B}_{\mathbf{O}_3\mathbf{T}}$ and covariance $\mathbf{B}_{\mathbf{TT}}$. For temperature error variance, we adopted the procedure used in the meteorological CMC 3D-Var which consist of renormalizing the NMC error variances using the innovation statistics [53]. For ozone, we used the error variance estimates obtained from the HL method and made it dependent only on height as described above in Section 3.1. Figure 6 (left panel) illustrate the cross-covariance thus obtained, which we will denote by $\mathbf{A}^{CQC-NMC}$.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2019 doi:10.20944/preprints201910.0241.v1

Peer-reviewed version available at Atmosphere 2019, 10, 798; doi:10.3390/atmos10120798

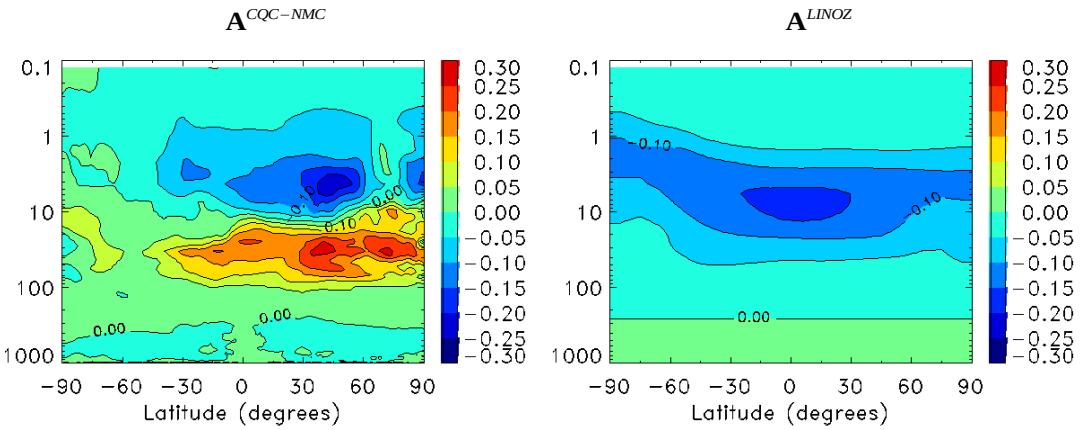


Figure 6 Balance operator between ozone and temperature for July 2003. Left panel, $\mathbf{A}^{CQC-NMC}$, which uses CQC and NMC methods, and right panel, $\mathbf{A}^{LINOZ}$, as derived from the LINOZ scheme.

We also calculated the cross-covariance obtained using the LINOZ model, which is derived in Appendix B using the stationary solution of the cross-covariance evolution equation between ozone and temperature, and which we denote by $\mathbf{A}^{LINOZ}$, displayed in the right panel of Figure 6. The ratio of the unexplained variance to the total variance for the operator $\mathbf{A}^{CQC-NMC}$ is shown in Figure S16 (Supplementary Material). 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 B), with $\mathbf{A}^{LINOZ} \approx -2\times10^{-2}\ \overline{O_3}$.

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 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 6-hour window is almost entirely global (Figure S9 Supplementary Material). The horizontal coverage is in fact too dense to consider 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 (Figure S10 – Supplementary Material). Limb sounding instruments such as MIPAS are another important source of

685 temperature measurements in the stratosphere. A description of MIPAS and HALOE is given in
686 Part I, Section 7.1. For the time period we have considered (i.e. 2003) AMSU-A and MIPAS are
687 the two most important source of stratospheric temperature measurements, with the exception of
688 radiosondes that rise up to 30 km in tropical regions (and lower altitudes elsewhere).

689 The main issue with AMSU-A radiances is that the geolocated and calibrated radiances (i.e.
690 level 1B) need to be bias-corrected and this is usually done by using the meteorological model
691 short-term forecast as an "unbiased" estimate. This procedure is well adapted in the troposphere
692 where other unbiased observations have a significant effect on the analysis, thus by comparing
693 model-simulated radiances with observed radiances can be used effectively to separate model
694 bias from observational bias. Such observations are often referred to as "anchor" observations
695 in a bias correction scheme. Observation bias-correction schemes can be either static or online
696 with the analysis, as in the Variation Bias Correction scheme [89]. However, it is found that the
697 application of bias correction in the upper stratosphere is problematic in the absence of "anchor"
698 observations [90]. DiTomaso and Bormann [90] have proposed assimilating AMSU-A channel
699 14 without any bias correction as a way to anchor the meteorological analysis in the mid to upper
700 stratosphere. Here, we propose another approach, which consists of assimilating MIPAS
701 temperature observations to anchor the stratospheric analysis and derive from it a new set of
702 AMSU-A bias corrected radiances. This also has the effect of harmonizing these two sets of
703 observations.

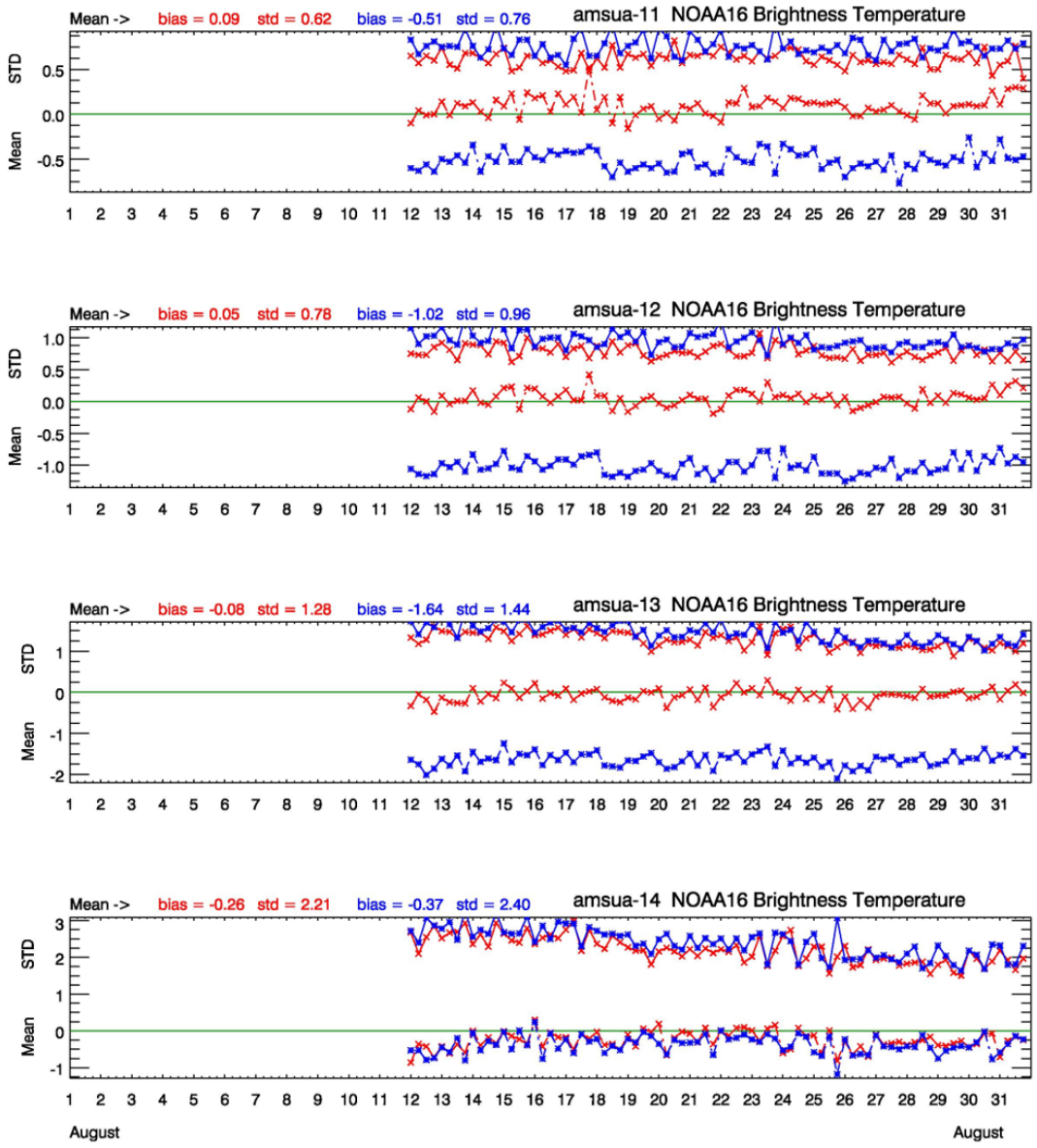
704 MIPAS-retrieved temperatures in the stratosphere are considered to be of good quality and
705 compare well with HALOE temperatures (see Part I, Section 7.2.1). We thus conducted an
706 assimilation of MIPAS temperature observations *without* AMSU-A (stratospheric) channel 10-14
707 as an "anchor"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

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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2019 doi:10.20944/preprints201910.0241.v1

Peer-reviewed version available at Atmosphere 2019, 10, 798; doi:10.3390/atmos10120798

718



719

720

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

726

727 We observe a net reduction in radiance bias for channels 11-13 with the new bias correction
728 based on MIPAS temperatures, with a slight reduction in the standard deviation. The mean

analysis increment at 10 hPa is presented in Figure S11 (Supplementary Material) for September 2003 and a zonal mean analysis increment 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. Following the above results, all further 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 temperature in addition to AMSU-A radiances, with the new bias correction (Section 4). An assimilation from August 17 to September 30, 2003 was conducted and the global verification results are presented in Figure 8. In green is the assimilation of AMSU-A only, and in black the assimilation of MIPAS temperature and AMSU-A.

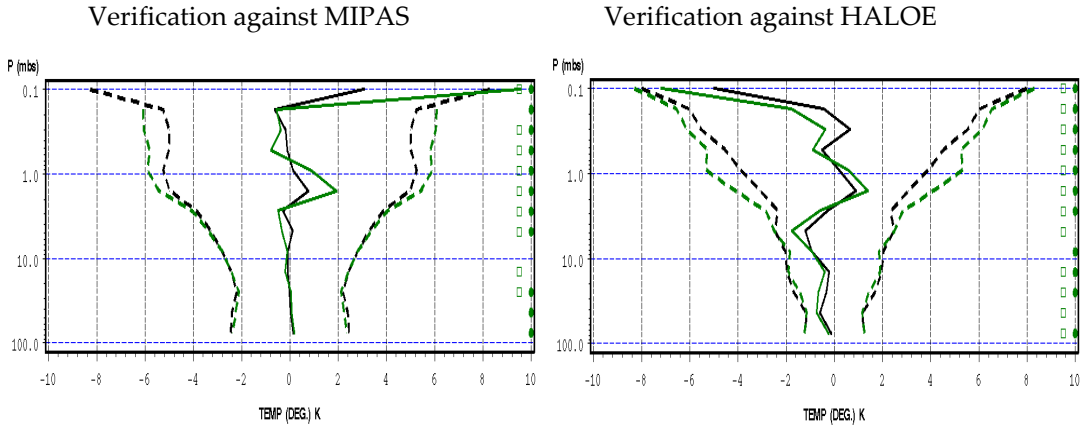


Figure 8 Global verification (observation-minus-forecast) of temperatures for two assimilation runs. All AMSU-A data are processed with the new bias correction. The left panel illustrates verification against MIPAS temperatures, and the right panel, verification against HALOE temperatures. The green squares on the far right of the panels indicate significance according the the Student t-test of means, and the green dots on the far right of the panels indicate significance according to the Fisher test of variances.

We observe an improved bias and reduction in error variance 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 and HALOE as independent observations. The larger impact in the mid to upper stratosphere may be due to the fact that there are more AMSU-A channels sensitive to the lower stratosphere, or that the limb sounding observations 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. To address this question we have performed an assimilation of AMSU-A only versus MIPAS only.

Verification against HALOE temperatures (Figure 9) shows very little difference with respect the combined assimilation results (right panel of Figure 8), but 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 in the stratosphere.

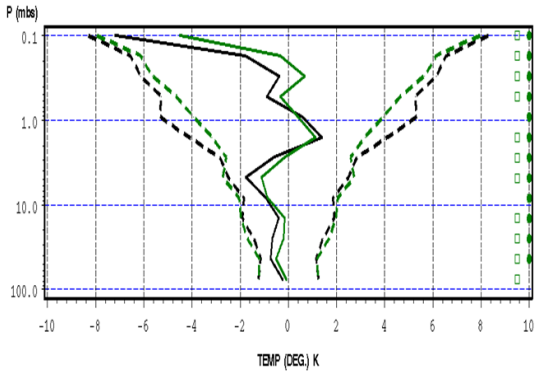


Figure 9 Global verification (observation-minus-forecast) against HALOE temperatures for two assimilation runs. In black is the assimilation of AMSU-A, and in green is the assimilation of MIPAS temperatures only. Otherwise similar to Figure 8.

Next, we conducted another set of experiment that directly illustrate 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 Figure 8 and 9, we activate ozone-radiation interaction in the model. But as we shall see in the following section (Section 6), the ozone-radiation interaction has very little impact on verification of 6-hour forecast. The impact actually develops over a time period of several days, so that for all practical purposes

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2019 doi:10.20944/preprints201910.0241.v1

Peer-reviewed version available at Atmosphere 2019, 10, 798; doi:10.3390/atmos10120798

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 temperature where all AMSU-A channels are retained (in black). The new AMSU-A bias correction scheme was applied in both cases.

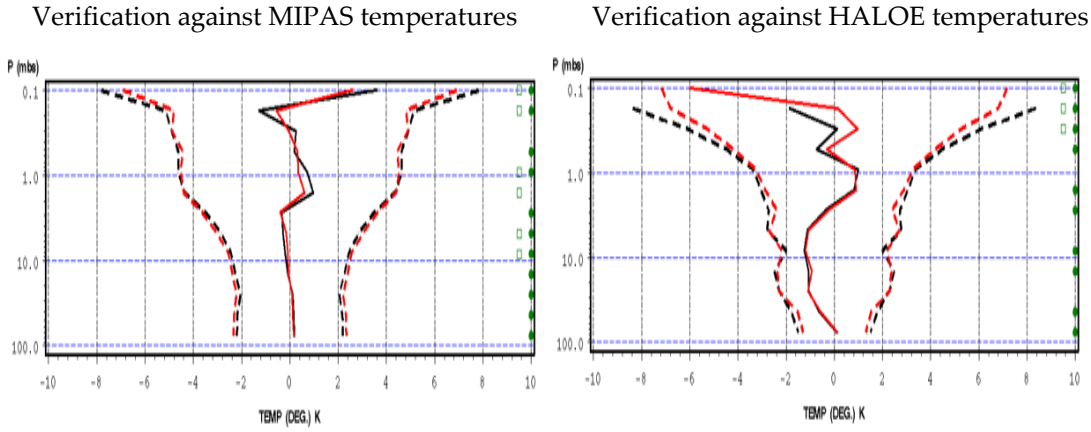


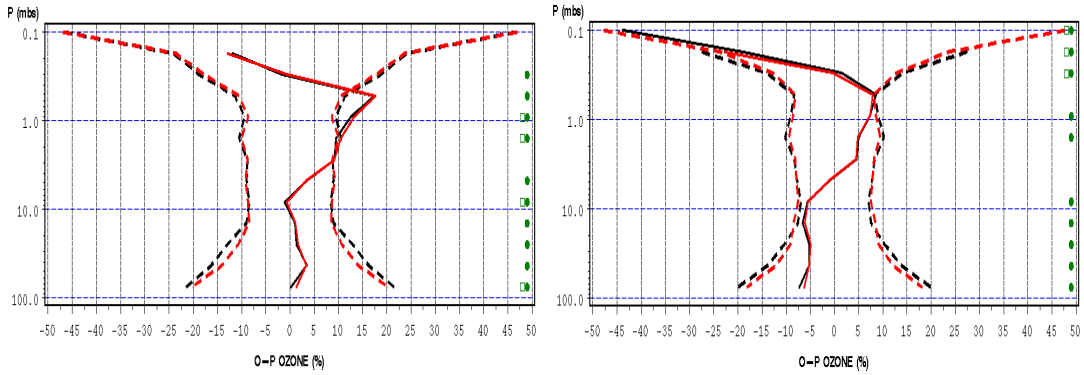
Figure 10 Global verification (observation-minus-forecast) of temperature for two assimilation runs. In red, is the assimilation of MIPAS temperature and AMSU-A with no stratospheric channels, and in black is the assimilation of MIPAS temperatures with all the AMSU-A channels. The left panel illustrates the verification against MIPAS temperatures, and on the right panel, the verification against HALOE temperatures.

Figure 10 displays the global verification results of assimilation runs from August 17 to October 31, 2003. In general, for the mid and upper stratosphere, both in terms of bias and error standard deviation, 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 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).

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

802

803 Verification against MIPAS O_3 Verification against HALOE O_3



804

805 **Figure 11** Same as Figure 10 but for verification of ozone MIPAS on the left and

806 ozone HALOE on the right.

807 The reduction in the random error is markebly larger in the lower stratosphere where
808 transport and the vertical gradient of ozone are important. A larger reduction in standard
809 deviation is observed over Antarctica (results not shown). A reduction in the error standard
810 deviation is also observed for CH_4 above 3 hPa. Thus, we see that the presence of AMSU-A
811 temperatures in the assimilation actually degrades the vertical structure, because of the coarse
812 vertical resolution sensitivity of the associated channels, which is apparent in the transport of
813 chemical species in regions of strong vertical concentrations.

814 **6. Weak coupling assimilation due to ozone-radiation interaction**

815 We know (form Part I, section 2.1) that the ozone-radiative interaction time-scale varies from
816 about a week at 1hPa to about a month at 25 hPa, while the ozone photochemical time-scale is a
817 few hours at 1hPa and is on the order of three months at 25 hPa (see Part I, Section 2.1). It is
818 around 10 hPa that these two interactions have comparable time-scales, i.e. about 2 weeks (see
819 Part I, Figure 2). This implies that the assimilation of ozone will have little impact on
820 temperatures above 10 hPa, but the impact, which is radiative in nature, will be noticeable in the
821 lower stratosphere and will build up slowly over time. The assimilation of temperature on the
822 other hand will influence the photochemistry of ozone above 10hPa and will influence ozone
823 transport in the lower stratosphere.

824 To examine these effects in the context of assimilation we will focus on assimilating only limb
825 sounding observations. As stated in Section 5, the assimilation of limb sounding temperatures
826 while excluding stratospheric AMSU-A channels has a stronger impact on both temperature and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2019 doi:10.20944/preprints201910.0241.v1

Peer-reviewed version available at Atmosphere 2019, 10, 798; doi:10.3390/atmos10120798

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 August 17 to September 5, 2003. The global verification of observation-minus-forecast (6 hours) temperatures and ozone is presented in Figure 12.

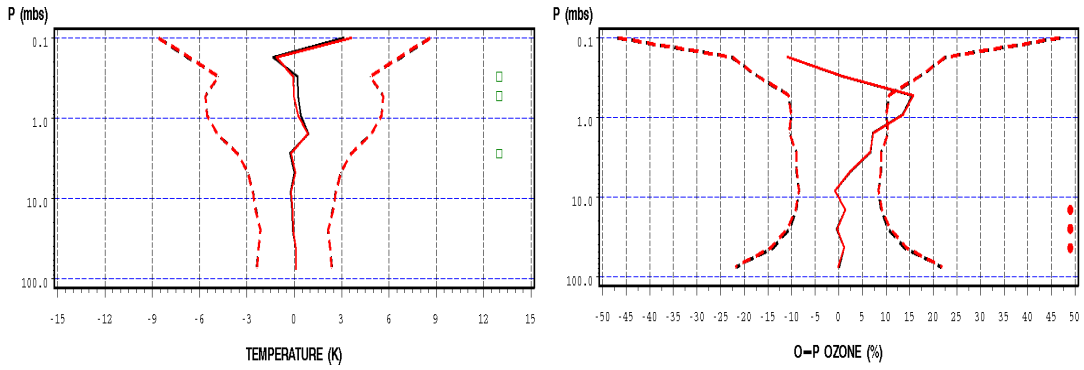


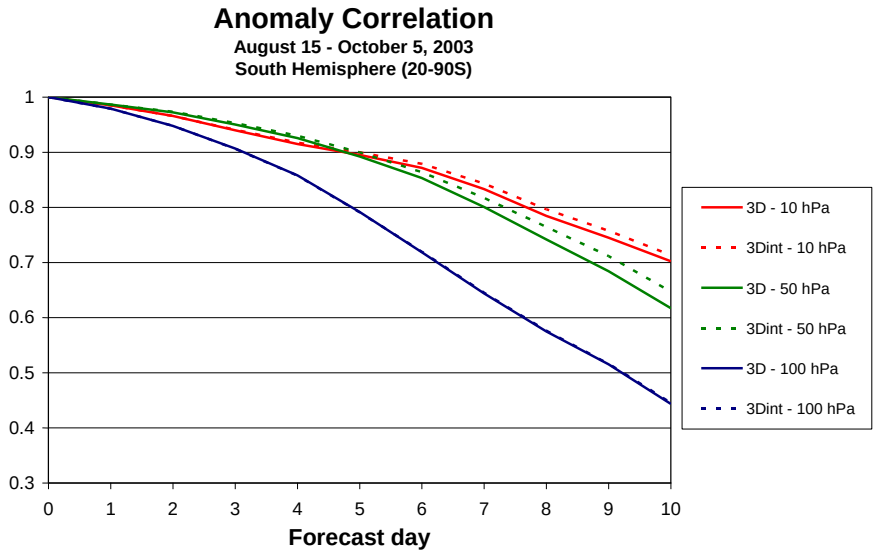
Figure 12 Impact of ozone radiation interaction with the assimilation of MIPAS temperature only and AMSU-A channels 1-8. Left panel shows the global impact on MIPAS observation-minus-forecast temperatures, and the right panel the ozone impact against MIPAS observations. No interaction (black), ozone-radiation interaction (red).

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 hour forecasts) except for small differences in the upper-stratospheric mean temperature and the variance of lower stratospheric ozone.

The small mean difference in temperature between the two experiments around 5hPa 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.

852 In the lower stratosphere, the O-P variance is increased in the case of ozone-radiation
853 interaction. We recall that there is no assimilation of ozone in these experiments, and the impact
854 on ozone can be understood by considering ozone as an unobserved variable as defined in Section
855 2.1. The impact on unobserved variables can be computed from the cross-variable increment,
856 Equation (5), and here in particular, the balance operator **A** between ozone and temperature. The
857 associated background (or model) error covariance is given by Equation (20) and using the
858 operator **A**. We have shown already in Figure 5 that ozone-radiation interaction increases the
859 correlation between temperature and ozone between 10 and 100 hPa (in the northern latitude
860 summer). Consequently, the error-cross covariance and its effect on variance of ozone is
861 increased, and this is what it is observed in the lower panel of Figure 12 .

862 Although the impact of ozone-interaction is nearly absent in analyses (or 6-hour forecasts), it
863 gradually accumulates in forecasts. De Grandpré et al. [39] have reported results of assimilation
864 of temperature and ozone on the temperature predictability using the GEM-BACH with
865 essentially the same experimental setup discussed here. A gradual increase in the anomaly
866 correlation for the period of August 11 to September 5, 2003 was shown reaching nearly half a
867 day in the lower stratosphere as a result of ozone-radiation interaction (either with assimilation
868 of temperatures only or with assimilation of temperature and ozone). Here we show anomaly
869 correlation results in which the assimilation of MIPAS temperature was conducted over a longer
870 time period from August 15 to October 5, 2003 that essentially corroborate the published results.
871 For a description of the calculation of the anomaly correlation (i.e. correlation between forecast
872 and analysis valid at the same time) we refer the reader to de Grandpré et al. [39]



873

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2019 doi:10.20944/preprints201910.0241.v1

Peer-reviewed version available at Atmosphere 2019, 10, 798; doi:10.3390/atmos10120798

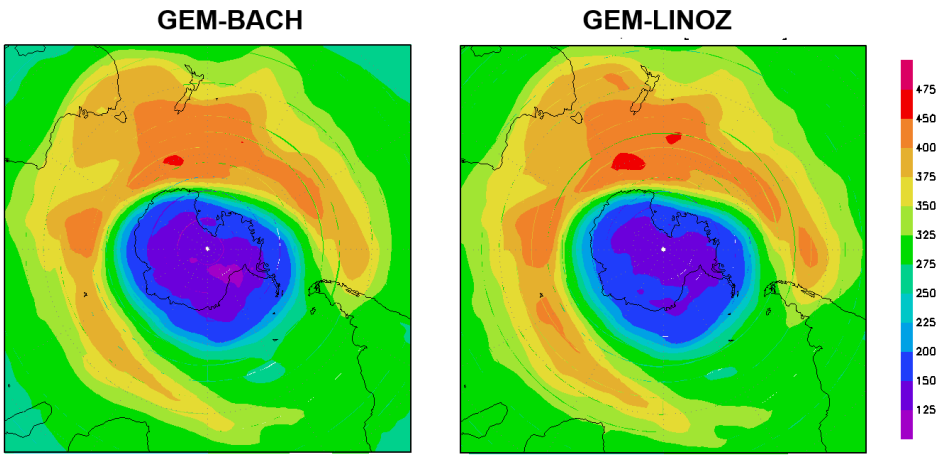
874 **Figure 13** Anomaly correlation at 10 (red), 50 (green) and 100 (purple) hPa in the
875 southern hemisphere (20S-90S) for ozone-radiation interactive (dashed lines) and
876 non-interactive ozone (solid lines) experiments.

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

881 The precise chemistry model used has in fact little impact on these results. To show this we
882 have conducted a similar ozone-radiation interaction assimilation experiment using a linearized
883 chemistry model LINOZ [57] for daily mean values (with semi-Lagrangian transport [58])

884 $$\frac{dO_3}{dt} = c_1 + c_2(O_3 - \bar{O}_3 + c_3(T - \bar{T}) + c_4(O_3^{\uparrow} - \overline{O_3^{\uparrow}}), \quad (21)$$

885 where the coefficients c_1, c_2, c_3, c_4 are determined using a chemical box model, the overbar $\overline{(\,)}$
886 represents climatological values, and $\uparrow$ represents the overhead column. The coefficient
887 $c_2 = 1/\tau_{O_3}$ is related to the photochemical time-scale of ozone (see also Section 2.1, Part I).
888 Figures 14 and 15 show the impact of assimilating temperature and ozone with ozone-radiation
889 interaction activated using the LINOZ ozone model and the BASCOE chemistry model.
890

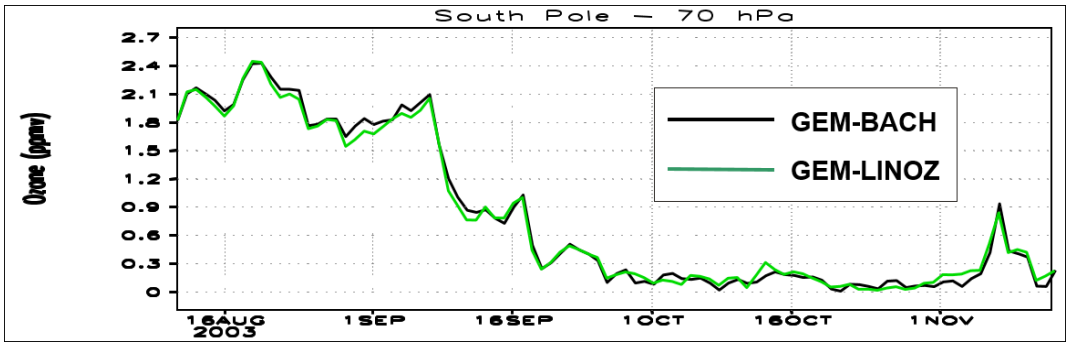


891

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2019 doi:10.20944/preprints201910.0241.v1

Peer-reviewed version available at Atmosphere 2019, 10, 798; doi:10.3390/atmos10120798

892 **Figure 14** Total column ozone (DU) over the South Pole region on October 3, 2003
893 resulting from the assimilation of MIPAS temperature and ozone. Left panel,
894 experiment using the BASCOE chemistry scheme. Right panel, experiment using the
895 LINOZ linearized ozone chemistry scheme.
896



897

898 **Figure 15** Time series of ozone at 70hPa over the South Pole region resulting from
899 the assimilation of MIPAS temperature and ozone. Black curve, experiment using
900 the BASCOE chemistry scheme. Green curve, experiment using the LINOZ
901 linearized ozone chemistry scheme.

902 We conclude from these figures that the analysis and time evolution of ozone over the South
903 Pole region with GEM-BACH ozone-radiation interaction are similar whether we use the
904 comprehensive (BASCOE) chemistry or the linearized (LINOZ) chemistry. Figure 16 shows the
905 temperature forecast, bias and the error standard deviation at 50hPa over the Northern
906 Hemisphere in comparison with MIPAS temperature analyses.
907

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2019 doi:10.20944/preprints201910.0241.v1

Peer-reviewed version available at Atmosphere 2019, 10, 798; doi:10.3390/atmos10120798

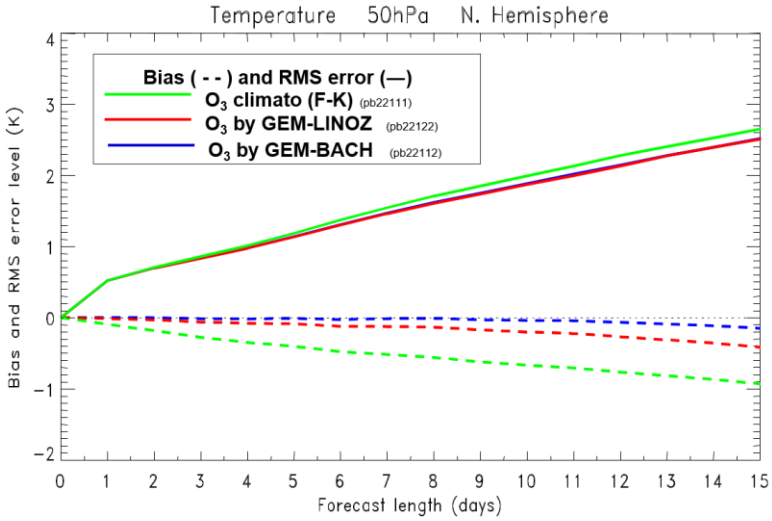


Figure 16 15-day forecast of temperature at 70hPa verified against analyses over the South Pole region resulting from the assimilation of MIPAS temperature and ozone. In black using the BASCOE chemistry and in green using the LINOZ linearized ozone chemistry.

Green curves correspond to a forecast with the climatological ozone in the radiation calculation. Red and blue curves correspond to forecast runs were prognostic ozone in the radiation calculation was used. 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 error, while a faster growth of error is seen when using the ozone climatology. This result is coherent with the anomaly correlation results presented in Figure 13 (and [39]) which indicate greater forecast skill with ozone-radiation interaction than using climatological ozone. The result illustrated in Figure 16, also suggest that an anomaly correlation computed using the LINOZ chemistry should lead to improvement 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 either using the full chemistry or a simplified (linearized) ozone chemistry schemes.

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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2019 doi:10.20944/preprints201910.0241.v1

Peer-reviewed version available at Atmosphere 2019, 10, 798; doi:10.3390/atmos10120798

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 O_3 and AMSU-A temperature (all channels) for a period of 2 weeks from August 17 to September 4, 2003. Figure 17 shows the verification over the globe in the three case: univariate (red), multivariate with the balance operator $\mathbf{A}^{CQC-NMC}$ (green), and multivariate with the LINOZ balance operator $\mathbf{A}^{LINOZ}$ (blue).

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 being for the upper stratosphere temperature where the LINOZ operator reduces slightly the temperature bias (although not significantly), and for ozone an increase of variance for both LINOZ and CQC-NMC operators in the lower stratosphere.

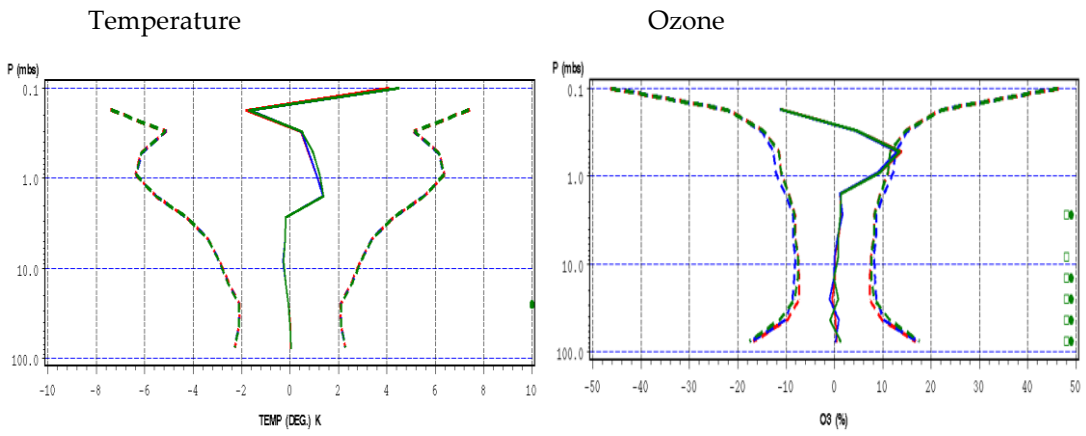


Figure 17 Multivariate temperature-ozone assimilation. Univariate ozone and temperature assimilation (red), multivariate assimilation performed with the LINOZ balance $\mathbf{A}^{LINOZ}$ (blue) and CQC-NMC balance $\mathbf{A}^{CQC-NMC}$ (green). The solid lines denote average differences (biases) and the dashed lines indicate the standard deviations (by $\pm\sigma$). Left panel, temperature O-P (observation minus 6-hour forecast) statistics from comparisons to MIPAS observations. Right panel, ozone O-P statistics. The significance tests (green squares and dots) are between the red and green experiments.

The above 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 6-hour time period.

These results suggest that the modeling assumption of using the temperature instead of the unbalanced temperature, i.e. Equation (7), in the CQC-NMC balance operator, has a detrimental effect. The error correlations below 10 hPa (see Figure 5) are dominated by transport – thus contain the balanced temperature. Although we have not continued this experiment further, it seems necessary to construct balance operators between ozone and unbalanced temperature to truly isolate ozone-radiation from transport in the lower stratosphere.

8. Strongly coupled tracer-meteorology assimilation with 4D-Var

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 [90],[41]. On the other hand, Daley [40] has alluded to the fact that spatial variation of error variance 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

$$\frac{d\,\mathbf{X}}{dt} = \mathbf{V}(x,t) \tag{22}$$

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,

$$u = \frac{\partial \psi}{\partial y} \quad ; \quad v = -\frac{\partial \psi}{\partial x} \tag{23}$$

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 coincide, and could thus use the streamfunction as a proxy for the concentrations. The cross-covariance between the streamfunction and the wind is obtained from

$$\langle u\psi \rangle = \frac{1}{2}\frac{\partial}{\partial y}\langle \psi^2 \rangle \;\; ; \;\; \langle v\psi \rangle = -\frac{1}{2}\frac{\partial}{\partial x}\langle \psi^2 \rangle \,, \tag{24}$$

and thus is clearly depends 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 **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 [60] 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 emphasis here will be on the wind information obtained in 4D-Var from passive tracer information. Note that 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 concentration and the wind trajectories are the same in both cases.

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. The results are shown at 100 hPa, a level where methane induces the most significant wind corrections. The background error was increased near the North Pole, which reduced the wind correction in the region. However, in the Southern Hemisphere, the reduction in observation error caused significant increases in the wind correction. 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.

Old statistics New (HL) statistics

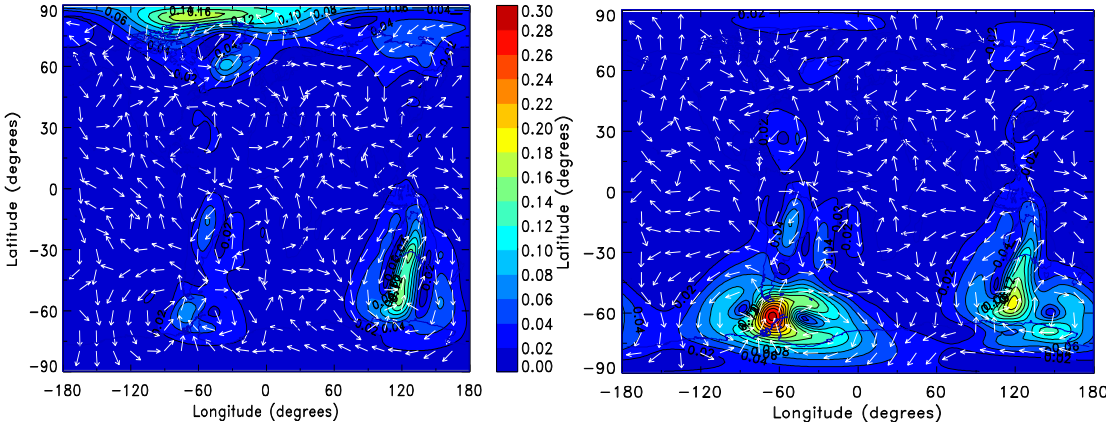


Figure 18 Wind analysis increments in response to MIPAS CH_4 observations obtained with a) the first estimate of background-error statistics for chemical constituents, and b) the new statistics estimated using the Hollingsworth-Lönnberg method. Results are shown here at the 100-hPa level

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 S17-18 (Supplementary Material). The differences in the increments can be explained "mechanistically" by differences in the distribution of the constituents at different levels. Figure S19 (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.

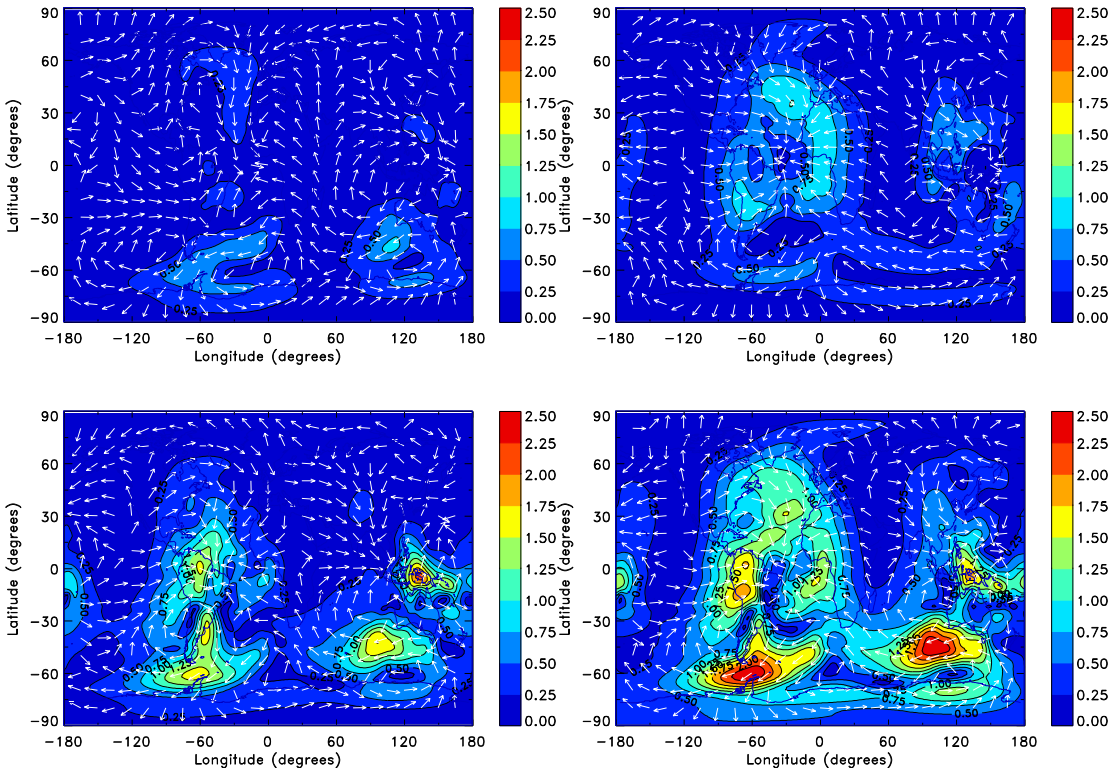


Figure 19 Wind analysis increments at 10 hPa obtained by assimilating CH_4 (top left), O_3 (top right), N_2O (bottom left) and all three species (bottom right).

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. The results are shown in Figure 20 and indicate a reduction in the zonal wind bias all the way to the mid troposphere. The results shown here are based in assimilation cycles covering the period August 15th to October 5th 2003, over which the results were averaged.

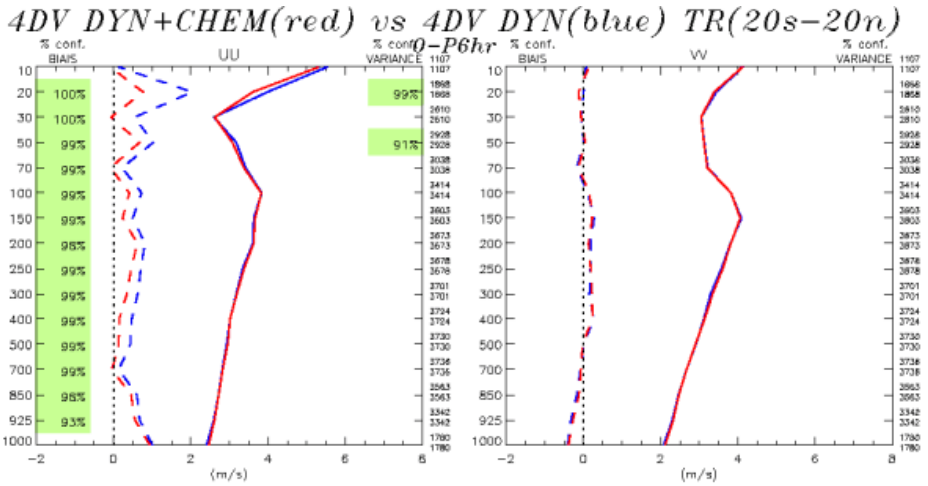


Figure 20 Verification against radiosondes data over the tropical region (20°S – 20°N) for the period August 15th to October 5th 2003. The results in red correspond to a 4D-Var assimilation experiment with assimilation of ozone, methane and nitrous oxide. Results in blue are 4D-Var experiments but without assimilation of the long-lived species. Tables in green on the left and right side of the panel indicate also the statistical significance of the results for the bias (on the left) and for the variance (on the right).

The difference between the 4D-Var wind analyses with and without the assimilation of passive chemical tracers (i.e. no chemistry) is shown in Figure 21. We note that the wind 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 radiosonde data (Figure 20), except near 20 hPa where the difference in 4D-Var is about 3 ms^{-1} while the radiosondes observations indicate a correction of about 1 ms^{-1}. Also, we 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. However, there is a suspicious large wind increment in the Tropics near 2-3 hPa. We have noted in the zonal mean ozone a curious vertical transport as a result of 3D-Var meteorological analysis exactly at the same location, where the model developed a very large vertical diffusion coefficient. The origin of this problem was completely elucidated but seems to be due in part to the vertical discretization (non-staggered) coordinate of the model and to the treatment of the error variance in the tropical region, wich assumes that wind and temperature analysis become univariate close to the equator.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2019 doi:10.20944/preprints201910.0241.v1

Peer-reviewed version available at Atmosphere 2019, 10, 798; doi:10.3390/atmos10120798

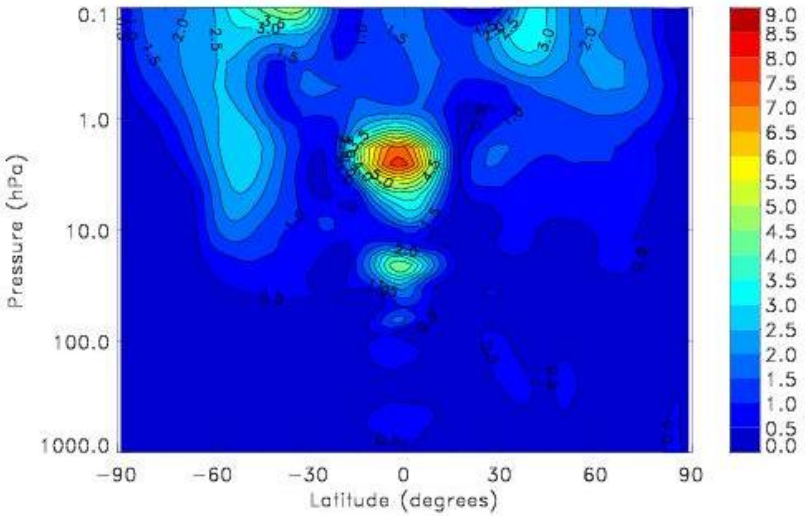


Figure 21 Difference between the wind vector intensity of the analyses obtained from two assimilation cycles executed with and without the assimilation of ozone, methane and nitrous oxide. The results are averaged over the period August 15 to October 5, 2003. The zonal mean average is shown here.

We also observed changes in the temperature that are very small throughout the lower stratosphere below 10 hPa, where constituents are assimilated. Slight differences appear in the Tropics but also in the Northern Hemisphere lower stratosphere which indicates some propagation of the effect into the extra-tropical regions. Figure 22 depicts the O-P temperature time series at 20 hPa in the Northern Hemisphere and indicates a buildup of systematic differences between both assimilation systems throughout the period. For the assimilated species (O_3, CH_4, N_2O) changes induced by the perturbation of the meteorological fields are small throughout the lower stratosphere.

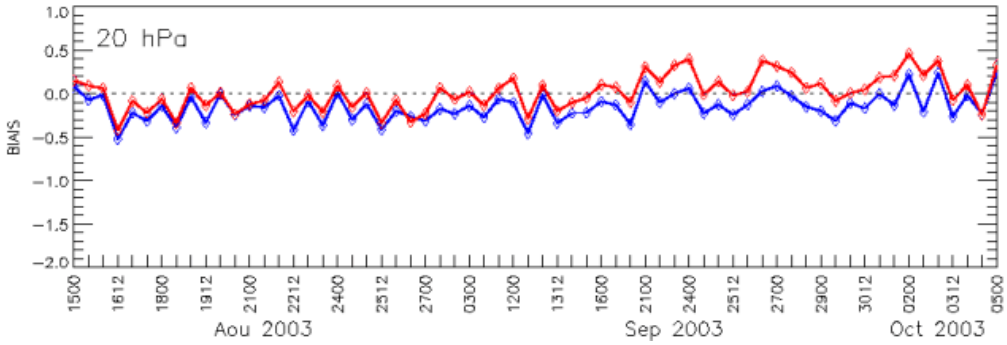


Figure 22 OmP temperature time series between the radiosondes and the 3D-Var (blue) and 4D-var (red) assimilation cycles at 20 hPa in the North Hemisphere.

The impact of the correction on the wind and transport of chemical species is a second-order effect and more difficult to assess, as the 4D-Var assimilated those constituents to produce a wind correction. For ozone, for instance, differences appear mainly in the winter hemisphere (Southern Hemisphere) where dynamical processes are more important. Figure S20 (Supplementary Material) shows the comparison of ozone from the assimilation against MIPAS O_3 observations for the period September 20th to October 5th below 10 hPa in the Southern Hemisphere mid-latitudes and polar regions. Results from 4D-Var show significant improvement in bias at the South Pole but the Southern Hemisphere mid-latitude region appears to have a smaller bias near 100 hPa in the 3D-var system. 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.

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 ?", which in the case of the coupled model considered here (i.e. GEM-BACH) where there is over 61 prognostic variables and in the order of 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 $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_k\}$ can be transformed into a set of uncorrelated random variables $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \ldots, \mathbf{u}_k\}$ 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 [70]. The procedure to construct balance operators is not new, but the way we presented here capture more easily the general nature of the approach which can be translated into any geophysical context.

In order to obtain 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 autocovariance function. We showed that we can extract a spatially correlated and uncorrelated component, from which we assign the observation error variance and the background error variance. Since the NMC method assumes implicitly a wide coverage of observations per analyses, we used instead the Canadian Quick Covariance method (CQC) [60], which consist of using 6-hour differences of the forecasts. The CQC method actually represents the spatial statistics due to advection with physical 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 correlation of most species are in fact separable in the stratosphere except for O_3, HNO_3 and H_2O on large scales (wavenumber 20 and smaller). However, the resulting horizontal correlation length appears to be too small. With the CQC approach, we also computed the cross-covariance between ozone and temperature, and showed 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 unbalanced temperature rather than temperature, but this requires 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 the bias-corrected AMSU radiances can render, 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. Also, that 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 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 unbalanced temperature is partially 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 the distribution of the concentration (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 variance introduce cross-covariances between winds and chemical tracers [40,41,45]. 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 2.5ms^{-1} was also found in the surf-zone above 5hPa but it is unclear if this helped the transport of chemical constituents, possibly due to the fact that chemical tracers that are assimilated in the first place so this increment is 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, but also how delicatly the error covariances need to respect the wide range of time-scales involved. Due to the scope of this study, the wide range of expertise needed, and

accounting for the resources and time limitation such a study requires, it was not possible to answer all remaining questions, however the authors encourage further development in this area.

Appendix A. Derivation of analysis splitting between observed and unobserved variables

Mathematically the problem is posed as follows (the derivation was first publish in a conference proceedings [92]). 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 **B** into square root and invertible matrix **S**,

$$\mathbf{B} = \mathbf{S}\mathbf{S}^T . \tag{A.1}$$

Defining $\boldsymbol{\zeta}$ such that

$$\mathbf{z} = \mathbf{S}\boldsymbol{\zeta} \tag{A.2}$$

the J_b term then simplifies to

$$(\mathbf{z}-\mathbf{z}^f)^T\mathbf{B}^{-1}(\mathbf{z}-\mathbf{z}^f) = (\boldsymbol{\zeta}-\boldsymbol{\zeta}^f)^T(\boldsymbol{\zeta}-\boldsymbol{\zeta}^f) . \tag{A.3}$$

Introducing a representation of observed and unobserved variables in the **B** covariance matrix leads to a decomposition of the form,

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{xx} & \mathbf{B}_{xu} \\ \mathbf{B}_{ux} & \mathbf{B}_{uu} \end{pmatrix} . \tag{A.4}$$

The inverse of **B** is then of the form

$$\mathbf{B}^{-1} = \begin{pmatrix} \mathbf{D} & \mathbf{E} \\ \mathbf{F} & \mathbf{G} \end{pmatrix} \tag{A.5}$$

where

$$\begin{aligned} \mathbf{D} &= \left(\mathbf{B}_{xx} - \mathbf{B}_{xu}\mathbf{B}_{uu}^{-1}\mathbf{B}_{ux}\right)^{-1} \\ \mathbf{E} &= -\mathbf{B}_{xx}^{-1}\mathbf{B}_{xu}\mathbf{G} \\ \mathbf{F} &= -\mathbf{B}_{uu}^{-1}\mathbf{B}_{ux}\mathbf{D} \\ \mathbf{G} &= \left(\mathbf{B}_{uu} - \mathbf{B}_{ux}\mathbf{B}_{xx}^{-1}\mathbf{B}_{xu}\right)^{-1} . \end{aligned} \tag{A.6}$$

To obtain the square root **S**, let it first be represented in the form,

$$\mathbf{S} = \begin{pmatrix} \mathbf{d} & \mathbf{e} \\ \mathbf{f} & \mathbf{g} \end{pmatrix} . \tag{A.7}$$

Then from Equation (A.1) we get

$$\begin{aligned}\mathbf{B}_{xx} &= \mathbf{d}\mathbf{d}^T + \mathbf{e}\mathbf{e}^T \\ \mathbf{B}_{xu} &= \mathbf{f}\mathbf{d}^T + \mathbf{g}\mathbf{e}^T \\ \mathbf{B}_{ux} &= \mathbf{d}\mathbf{f}^T + \mathbf{e}\mathbf{g}^T \\ \mathbf{B}_{uu} &= \mathbf{f}\mathbf{f}^T + \mathbf{g}\mathbf{g}^T \ .\end{aligned} \tag{A.8}$$

There is more than one solution that satisfies these four equations. One of the solutions that leads to a triangular form consists of letting $\mathbf{e} = \mathbf{0}$. We can then easily invert $\mathbf{S}$. So letting $\mathbf{e} = \mathbf{0}$ in Equation (A.8) we first get

$$\mathbf{B}_{xx} = \mathbf{d}\mathbf{d}^{\mathbf{T}} . \tag{A.9}$$

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

$$\mathbf{d} = \mathbf{L} . \tag{A.10}$$

From the second line of Equation (A.8) we get,

$$\mathbf{B}_{ux} = \mathbf{f}\,\mathbf{L}^T \ \Rightarrow \mathbf{f} = \mathbf{B}_{ux}\,\mathbf{L}^{-T}, \tag{A.11}$$

and the third equation is satisfied trivially. Finally the fourth equation of Equation (A.8) takes the form,

$$\mathbf{B}_{uu} = \mathbf{B}_{ux}\,\mathbf{L}^{-T}\,\mathbf{L}^{-1}\,\mathbf{B}_{xu} + \mathbf{g}\mathbf{g}^T \ \Rightarrow \ \mathbf{g}\mathbf{g}^T = \mathbf{B}_{uu} - \mathbf{B}_{ux}\,\mathbf{B}_{xx}^{-1}\,\mathbf{B}_{xu} = \mathbf{G}^{-1} \tag{A.12}$$

where we used the fact that $\mathbf{B}_{xx}^{-1} = \mathbf{L}^{-T}\mathbf{L}^{-1}$. Thus $\mathbf{g}$ is the square-root of the inverse of $\mathbf{G}$

$$\mathbf{g} = \sqrt{\mathbf{G}^{-1}} \ . \tag{A.13}$$

Now let

$$\tilde{\mathbf{z}} = \begin{pmatrix} \tilde{\mathbf{x}} \\ \tilde{\mathbf{u}} \end{pmatrix}, \tag{A.14}$$

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,

$$\tilde{\boldsymbol{\zeta}} = \sqrt{\mathbf{B}^{-1}}\,\tilde{\mathbf{z}}, \tag{A.15}$$

which allows us to write,

$$\tilde{\mathbf{z}}^T\,\mathbf{B}^{-1}\,\tilde{\mathbf{z}} = \tilde{\boldsymbol{\zeta}}^T\tilde{\boldsymbol{\zeta}} \ . \tag{A.16}$$

Specifically, we have,

$$\tilde{\mathbf{z}} = \begin{pmatrix} \tilde{\mathbf{x}} \\ \tilde{\mathbf{u}} \end{pmatrix} = \sqrt{\mathbf{B}}\,\tilde{\boldsymbol{\xi}} = \mathbf{S}\tilde{\boldsymbol{\xi}} = \begin{pmatrix} \mathbf{L} & \mathbf{0} \\ \mathbf{B}_{\mathbf{ux}}\mathbf{L}^{-T} & \sqrt{\mathbf{G}^{-1}} \end{pmatrix} \begin{pmatrix} \tilde{\boldsymbol{\xi}}_1 \\ \tilde{\boldsymbol{\xi}}_2 \end{pmatrix} \tag{A.17}$$

and this system is easily inverted to give,

$$\begin{aligned} \tilde{\boldsymbol{\xi}}_1 &= \mathbf{L}^{-1}\tilde{\mathbf{x}} \\ \tilde{\boldsymbol{\xi}}_2 &= \sqrt{\mathbf{G}}\left(\tilde{\mathbf{u}} - \mathbf{B}_{\mathbf{ux}}\mathbf{B}_{\mathbf{xx}}^{-1}\tilde{\mathbf{x}}\right) \end{aligned} \tag{A.18}$$

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

$$\begin{aligned} 2J(\tilde{\mathbf{x}},\tilde{\mathbf{u}}) &= \begin{pmatrix} \tilde{\mathbf{x}} \\ \tilde{\mathbf{u}} \end{pmatrix}^T \mathbf{B}^{-1} \begin{pmatrix} \tilde{\mathbf{x}} \\ \tilde{\mathbf{u}} \end{pmatrix} + \left(\mathbf{y} - H\,\mathbf{x}\right)^T \mathbf{R}^{-1}\left(\mathbf{y} - H\,\mathbf{x}\right) \\ &= (\boldsymbol{\zeta} - \boldsymbol{\zeta}^f)^T(\boldsymbol{\zeta} - \boldsymbol{\zeta}^f) + \left(\mathbf{y} - H\,\mathbf{L}\boldsymbol{\zeta}_1\right)^T \mathbf{R}^{-1}\left(\mathbf{y} - H\,\mathbf{L}\boldsymbol{\zeta}_1\right) \\ &= (\boldsymbol{\zeta}_1 - \boldsymbol{\zeta}_1^f)^T(\boldsymbol{\zeta}_1 - \boldsymbol{\zeta}_1^f) + (\boldsymbol{\zeta}_2 - \boldsymbol{\zeta}_2^f)^T(\boldsymbol{\zeta}_2 - \boldsymbol{\zeta}_2^f) + \left(\mathbf{y} - H\,\mathbf{L}\boldsymbol{\zeta}_1\right)^T \mathbf{R}^{-1}\left(\mathbf{y} - H\,\mathbf{L}\boldsymbol{\zeta}_1\right) \\ &= 2\,J_1(\boldsymbol{\zeta}_1) + 2J_2(\boldsymbol{\zeta}_2) \end{aligned} \tag{A.19}$$

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

$$J_1(\boldsymbol{\zeta}_1) = \tfrac{1}{2}\left(\boldsymbol{\zeta}_1 - \boldsymbol{\zeta}_1^f\right)^T\left(\boldsymbol{\zeta}_1 - \boldsymbol{\zeta}_1^f\right) + \tfrac{1}{2}\left(\mathbf{y} - H\,\mathbf{L}\boldsymbol{\zeta}_1\right)^T \mathbf{R}^{-1}\left(\mathbf{y} - H\,\mathbf{L}\boldsymbol{\zeta}_1\right), \tag{A.20}$$

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

$$J_2(\boldsymbol{\zeta}_2) = \tfrac{1}{2}\left(\boldsymbol{\zeta}_2 - \boldsymbol{\zeta}_2^f\right)^T\left(\boldsymbol{\zeta}_2 - \boldsymbol{\zeta}_2^f\right). \tag{A.21}$$

The minimization of Equation (A.21) has the trivial solution

$$\boldsymbol{\zeta}_2 - \boldsymbol{\zeta}_2^f = \mathbf{0} \quad . \tag{A.22}$$

Now, assuming that $\sqrt{\mathbf{G}}$ is invertible, the solution (A.22) yields

$$\mathbf{u}^a - \mathbf{u}^f = \mathbf{B}_{\mathbf{ux}}\mathbf{B}_{\mathbf{xx}}^{-1}\left(\mathbf{x}^a - \mathbf{x}^f\right) \quad . \tag{A.23}$$

The minimization of the cost function $J_1(\boldsymbol{\zeta}_1)$ (A.20) is actually identical to the form 3D-Var takes written after preconditioning. Indeed in normal form with the non-transformed variables, (A.20) takes the form,

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^f)^T\mathbf{B}_{\mathbf{xx}}^{-1}(\mathbf{x} - \mathbf{x}^f) + \frac{1}{2}(\mathbf{y} - H\,\mathbf{x})^T\mathbf{R}^{-1}(\mathbf{y} - H\,\mathbf{x}) \quad . \tag{A.24}$$

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 $\mathbf{B}_{xx}$, 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 achived 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 [70] (or see [71] Section 1.2). For random variables (vectors) that have a non-zero mean, the proper definition of an inner product is

$$\langle \mathbf{x}, \mathbf{y} \rangle = E\left[(\mathbf{x} - E(\mathbf{x}))(\mathbf{y} - E(\mathbf{y}))^T \right] = \mathrm{cov}(\mathbf{x}, \mathbf{y}), \tag{B.1}$$

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 (B.1), $\langle \mathbf{x}, \mathbf{y} \rangle$ is a matrix where each entry is non-random. The square of the norm is then the variance, $\|\mathbf{x}\|^2 = \mathrm{var}(\mathbf{x})$, and the correlation matrix $\boldsymbol{\theta}$, between variables $\mathbf{x}$ and $\mathbf{y}$ is obtained as $\cos(\boldsymbol{\theta}) = \langle \mathbf{x}, \mathbf{y} \rangle / (\|\mathbf{x}\| \, \|\mathbf{y}\|)$. Therefore, uncorrelated random variables, $\mathrm{cov}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ are orthogonal, i.e. $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{0}$.

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

$$\begin{aligned}\mathbf{u}_1 &= \mathbf{v}_1\\ \mathbf{u}_2 &= \mathbf{v}_2 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_2)\\ \mathbf{u}_3 &= \mathbf{v}_3 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_3) - \mathrm{proj}_{\mathbf{u}_2}(\mathbf{v}_3)\\ &\dots\end{aligned} \tag{B.2}$$

where the projection (proj) is defined as

$$\mathrm{proj}_{\mathbf{u}}(\mathbf{v}) \triangleq \frac{\langle \mathbf{v},\mathbf{u}\rangle}{\langle \mathbf{u},\mathbf{u}\rangle}\mathbf{u}\,. \tag{B.3}$$

Applying this procedure to random vectors using Equation (B.1) and specifically to (unbiased) model background errors of the streamfunction $\delta\boldsymbol{\psi}$, velocity potential $\delta\boldsymbol{\chi}$, temperature $\delta\mathbf{T}$, and ozone $\delta\mathbf{O_3}$ 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 (21) 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

$$\int_z^\infty n_{\mathrm{O}_3}(z')\,dz' \tag{C.1}$$

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.

$$\mathrm{O}_3 = \frac{n_{\mathrm{O}_3}}{n_A}\,. \tag{C.2}$$

Using the relationship,

$$n_A M_A = \rho_A N_a \tag{C.3}$$

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

$$\int_z^\infty n_{O_3}(z')\,dz' = \frac{N_a}{M_A}\int_z^\infty O_3(z')\rho_A(z')\,dz' = -\frac{N_a}{M_A}\int_p^0 \frac{O_3(p')}{g}\,dp'. \quad \text{(C.4)}$$

Taking perturbations of ozone, $\tilde{O}_3$, and temperature, $\tilde{T}$, in (1) gives the following evolution equation for the perturbations

$$\frac{D\tilde{O}_3}{Dt} = c_2\tilde{O}_3 + c_3\tilde{T} + kc_4\int_p^0 \tilde{O}_3(p')\,dp', \quad \text{(C.5)}$$

where k is a constant that accounts for expressing the overhead column in DU. In Equation (C.5) we have neglected the changes in wind due to temperature perturbation. The last term of Equation (C.5) 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 (C.5), and let's 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

$$\frac{D\tilde{T}}{Dt} = 0 \quad \text{(C.6)}$$

$$\frac{D\tilde{O}_3}{Dt} = c_2\tilde{O}_3 + c_3\tilde{T}. \quad \text{(C.7)}$$

Multiplying (C.6) by $\tilde{O}_3$ and (C.7) by $\tilde{T}$, adding the resulting equations and taking the expectation gives,

$$\frac{D\langle\tilde{O}_3\tilde{T}\rangle}{Dt} = c_2\langle\tilde{O}_3\tilde{T}\rangle + c_3\langle\tilde{T}^2\rangle. \quad \text{(C.8)}$$

In Equation (C.7) the error cross-covariance is between any pressure levels. Carrying out the derivation more explicitly with pressure levels, Equation (C.8) can be rewritten as

$$\frac{D\langle\tilde{O}_3(p)\tilde{T}(p')\rangle}{Dt} = c_2(p)\langle\tilde{O}_3(p)\tilde{T}(p')\rangle + c_3(p)\langle\tilde{T}(p)\tilde{T}(p')\rangle \quad \text{(C.9)}$$

where p and p' are two distinct pressure levels. Multiplying Equation (C.6) by T' also gives,

$$\frac{D\left\langle \tilde{T}(p)\tilde{T}(p')\right\rangle}{Dt} = 0 \quad . \tag{C.10}$$

Similarly a covariance evolution for ozone can be derived as follows. Multiplying Equation (C.7) at pressure level p with $\tilde{O}_3(p')$, and multiplying Equation (C.7) at pressure level p' with $\tilde{O}_3(p)$, adding the equations and taking the expectation gives,

$$\frac{D\left\langle \tilde{O}_3(p)\tilde{O}_3(p')\right\rangle}{Dt} = \left[c_2(p)+c_2(p')\right]\left\langle \tilde{O}_3(p)\tilde{O}_3(p')\right\rangle + c_3(p)\left\langle \tilde{T}(p)\tilde{O}_3(p')\right\rangle + c_3(p')\left\langle \tilde{O}_3(p)\tilde{T}(p')\right\rangle \tag{C.11}$$

In matrix form Equation (C.9) is rewritten as,

$$\frac{D\mathbf{P}_{O_3T}}{Dt} = \mathbf{C}_2\mathbf{P}_{O_3T} + \mathbf{C}_3\mathbf{P}_{TT} \tag{C.12}$$

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

$$\mathbf{C}_2 = \begin{bmatrix} c_2(p_1) & 0 & 0 \\ 0 & \cdots & 0 \\ 0 & 0 & c_2(p_N) \end{bmatrix}, \tag{C.13}$$

and

$$\mathbf{P}_{O_3T} = \begin{bmatrix} \left\langle \tilde{O}_3(p_1)\tilde{T}(p_1)\right\rangle & \cdots & \left\langle \tilde{O}_3(p_1)\tilde{T}(p_N)\right\rangle \\ \cdots & \cdots & \cdots \\ \left\langle \tilde{O}_3(p_N)\tilde{T}(p_1)\right\rangle & \cdots & \left\langle \tilde{O}_3(p_N)\tilde{T}(p_N)\right\rangle \end{bmatrix} \tag{C.14}$$

$$\mathbf{P}_{TT} = \begin{bmatrix} \left\langle \tilde{T}(p_1)\tilde{T}(p_1)\right\rangle & \cdots & \left\langle \tilde{T}(p_1)\tilde{T}(p_N)\right\rangle \\ \cdots & \cdots & \cdots \\ \left\langle \tilde{T}(p_N)\tilde{T}(p_1)\right\rangle & \cdots & \left\langle \tilde{T}(p_N)\tilde{T}(p_N)\right\rangle \end{bmatrix} \quad . \tag{C.15}$$

In matrix form Equation. (B.10) is written as,

$$\frac{D\mathbf{P}_{TT}}{Dt} = 0, \tag{C.16}$$

and Eq. (C.11) can be rewritten as,

$$\frac{D\mathbf{P}_{O_3O_3}}{Dt} = \mathbf{C}_2\mathbf{P}_{O_3O_3} + \mathbf{C}_3\mathbf{P}_{O_3T} + \left(\mathbf{C}_2\mathbf{P}_{O_3O_3} + \mathbf{C}_3\mathbf{P}_{O_3T}\right)^T \tag{C.17}$$

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,

$$\mathbf{C}_2\mathbf{P}_{O_3T} + \mathbf{C}_3\mathbf{P}_{TT} = 0 \tag{C.18}$$

from which we obtain,

$$\mathbf{P}_{O_3T} = -\mathbf{C}_2^{-1}\mathbf{C}_3\mathbf{P}_{TT} \ , \tag{C.19}$$

and the balance operator is then of the form

$$\mathbf{A} = \mathbf{P}_{O_3T}\mathbf{P}_{TT}^{-1} = -\mathbf{C}_2^{-1}\mathbf{C}_3 \tag{C.20}$$

which is a diagonal matrix.

In Figure C.1 we plot the ratio $c_3(p)/c_2(p)$.

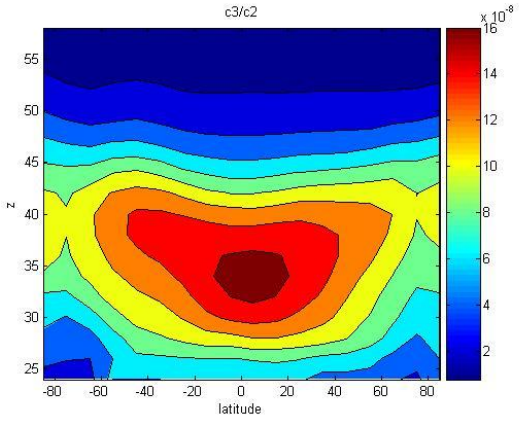


Figure C.1 Ratio $c_3(p)/c_2(p)$ for the month of September.

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 C.2 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.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2019 doi:10.20944/preprints201910.0241.v1

Peer-reviewed version available at Atmosphere 2019, 10, 798; doi:10.3390/atmos10120798

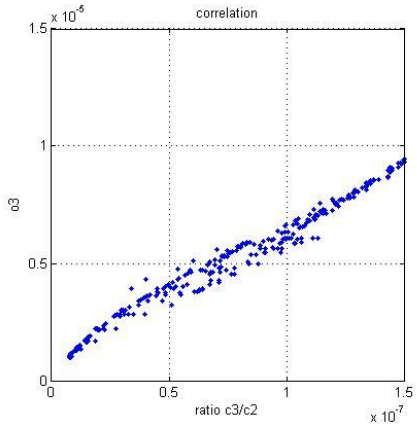


Figure C.2 Point-by-point, (λ, p), scatter of $c_3(p)/c_2(p)$ with $\bar{\mathrm{O}}_3$.

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 (C.5) is now written as,

$$\frac{D\tilde{\mathrm{O}}_3(p)}{\mathrm{D}t} = c_2(p)\tilde{\mathrm{O}}_3(p) + c_3(p)\tilde{T}(p) + kc_4(p)\int_p^0 \tilde{\mathrm{O}}_3(p'')\,dp'' . \qquad \text{(C.20)}$$

From Equation (C.20) and (C.6) we obtain

$$\frac{D\left\langle \tilde{\mathrm{O}}_3(p)\tilde{T}(p')\right\rangle}{Dt} = c_2(p)\left\langle \tilde{\mathrm{O}}_3(p)\tilde{T}(p')\right\rangle + c_3(p)\left\langle \tilde{T}(p)\tilde{T}(p')\right\rangle + k\,c_4(p)\int_p^0 \left\langle \tilde{\mathrm{O}}_3(p'')\tilde{T}(p')\right\rangle d\,p'' \quad . \qquad \text{(C.21)}$$

In discrete form, the last term of the r.h.s. of Equation (C.21) takes the following form:

For $p = p_N$ and $p' = p_j$ $\quad k\,c_4(p_N)\left\langle \tilde{\mathrm{O}}_3(p_N)\tilde{T}(p_j)\right\rangle \Delta p_N$

For $p = p_{N-1}$ and $p' = p_j$ $\quad k\,c_4(p_{N-1})\left\{\left\langle \tilde{\mathrm{O}}_3(p_N)\tilde{T}(p_j)\right\rangle \Delta p_N + \left\langle \tilde{\mathrm{O}}_3(p_{N-1})\tilde{T}(p_j)\right\rangle \Delta p_{N-1}\right\}$

…

For $p = p_i$ and $p' = p_j$ $\quad k\,c_4(p_i)\sum_{k=i}^{N}\left\langle \tilde{\mathrm{O}}_3(p_k)\tilde{T}(p_j)\right\rangle \Delta p_k$

and in matrix form,

$$\boldsymbol{\Sigma}\mathbf{P}_{\mathrm{O}_3 T} \qquad \text{(C.22)}$$

where

$$\mathbf{\Sigma} = \begin{bmatrix} \Delta p_1 & \Delta p_2 & \cdots & \Delta p_N \\ 0 & \Delta p_2 & \cdots & \Delta p_N \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \Delta p_N \end{bmatrix} . \tag{C.23}$$

In matrix form, similarly, to Equation (C.12) we obtain,

$$\frac{D\mathbf{P}_{O_3T}}{Dt} = \mathbf{C}_2\mathbf{P}_{O_3T} + \mathbf{C}_3\mathbf{P}_{TT} + k\mathbf{C}_4\mathbf{\Sigma}\mathbf{P}_{O_3T} = \left(\mathbf{C}_2 + k\mathbf{C}_4\mathbf{\Sigma}\right)\mathbf{P}_{O_3T} + \mathbf{C}_3\mathbf{P}_{TT} \ . \tag{C.24}$$

The balance operator is then of the form

$$\mathbf{A} = -\left(\mathbf{C}_2 + k\mathbf{C}_4\mathbf{\Sigma}\right)^{-1}\mathbf{C}_3 \ . \tag{C.25}$$

To compute the Δp_k term we use a centered formula

$$\Delta p_k = p_{k+\frac{1}{2}} - p_{k-\frac{1}{2}} \ . \tag{C.26}$$

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

$$\frac{dp}{p} = -\frac{g}{RT}dz \tag{C.27}$$

giving, in discrete form,

$$p_k = p_0 \exp\left\{-\frac{g\,\Delta z}{R}\left(\frac{1}{\bar{T}_0} + \frac{1}{\bar{T}_1} + \ldots + \frac{1}{\bar{T}_{k-1}}\right)\right\} \tag{C.28}$$

where $\bar{T}_k$ is the mean layer temperature. The figure below illustrates the vertical stagerring,

$$\begin{array}{ll} p_{k+1} & \text{\_\_\_\_\_} \ z_{k+1} \\ & \quad \bullet \ \bar{T}_k \\ p_k & \text{\_\_\_\_\_} \ z_k \end{array} \ .$$

Figure C.3 Vertical staggering of temperature and height .

Defining

$$p_{k+\frac{1}{2}} = \frac{p_k + p_{k+1}}{2} \qquad p_{k-\frac{1}{2}} = \frac{p_k + p_{k-1}}{2} \ , \tag{C.29}$$

we obtain

$$\Delta p_k = \frac{p_0}{2}\exp\left\{-\frac{g\Delta z}{R}\left(\frac{1}{\bar{T}_0} + \ldots + \frac{1}{\bar{T}_{k-1}}\right)\right\}\left[\exp\left(-\frac{g\Delta z}{R\bar{T}_k}\right) - \exp\left(+\frac{g\Delta z}{R\bar{T}_{k-1}}\right)\right] \ . \tag{C.30}$$

Supplementary Materials: 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_20). Figure S3: Background error variance from 6hr-difference method. Figure S4: Background error correlation spectra 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 O3 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-hr difference method for July 2003. Figure S9: Horizontal coverage of AMSU-A profiles in 6 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: Coefficfient of the LINOZ scheme for September. Figure S15: LINOZ climatology for September. Figure S16: Ratio of unexplained variance to total variance for the balance operator $\mathbf{A}^{CQC-NMC}$. Figure S17: Same as Figure 19 but at 50 hPa. Figure S18: same as Figure 19 but at 100 hPa. Figure S19: Analysis of N_2O and O_3 at 100 hPa on August 11, 2003, 00 UTC. Figure S20: OmP ozone comparison against MIPAS for the 3D-Var assimilation clycle and 4D-Var for the period of Septemeber 20 to October 5, 2003 over the South Pole region and Southern Hemisphere mid-latitudes.

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 anonymous reviewers.

Author Contributions: Conceptualization R.M.; methodology R.M., P.G., Y.R., A.R., J.G.; software Y.R., A.R., J.G., Y.Y., C.C., S.C.; validation P.G., Y.R., A.R., J.G., Y.Y., C.C., S.C.; formal analysis R.M., P.G., Y.R., A.R.,J.G.; investigation R.M., P.G., Y.R., A.R., J.G., writing –original draft preparation R.M.; writing – review and editing R.M., Y.R., A.R., J.G., S.C., visualization Y.R., A.R., J.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).

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

References

1. Lahoz, W. Research satellites. In Data Assimilation: Making Sense of Observations; Lahoz, W.; Khattatov, B.; Ménard, R., Eds.; Springer-Verlag: Berlin Heidelberg, 2010; pp. 301–321. doi:10.1007/978-3-540-74703-1_12.
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, [https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/92JD02828]. doi:10.1029/92JD02828.
3. UARS Science Team; Rood, R.B.; Geller, M.A., Eds.. UARS Data and Scientific Results (Special Issue). *J. Atmos. Sci*., **1994**, Vol. 51, 2781-3105.
4. UARS Science Team; Gille, J.C.; Massie, S.T.; Mankin, W.G., Eds.. Evaluation of the UARS Data (Special Issue). *J. Geophys. Res*. **1996**, D6, 9539-10473.
5. Louet, J. The Envisat Mission and System. http://www.esa.int/esapub/bulletin/bullet106/bul106_1.pdf, 820 2001.
6. Envisat Science Team. Validation Workshop Proceedings, 2002. https://envisat.esa.int/pub/ESA_822_DOC/envisat_val_1202/proceedings/ .

7. MIPAS Science Team. MIPAS Geophysical Validation (Special Issue). *Atmos. Chem. Phys*., **9**, pages = 413-442, 824 2002. http://envisat.esa.int/pub/ESA_DOC/envisat_val_1202/proceedings/ .
8. EOS Aura (website). http://aura.gsfc.nasa.gov/ .
9. EOS Aura Science Team. EOS Aura (Special Issue). *IEEE Transactions of Geoscience and Remote Sensing* **2006**, 44, 827 1063-1379.
10. Dee, D.P. `Coupled DA`. Presentation at the WMO Symposium on Data Assimilation. http://das6.umd.edu/program/Daily/slides/9.4-Dee_Dick.pdf .
11. Zupanski, M. Data assimilation for coupled modeling systems. In *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. III)*, Park, S.E.; Xu. L. (eds), Springer 2017, DOI: 10.1007/978-3-319-43415-5_2.
12. 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**, 26:10,218-10,231.
13. 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, doi:10.1007/s00382-013-1989-0.
14. 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, doi:10.1007/s00382-014-2390-3.
15. Lu, F.; Liu, Z.; Zhang, S.; Liu, Y. Strongly coupled data assimilation using leading averaged coupled covariance (LACC). Part I: Simple model study. *Mon. Wea. Rev*. **2015**, 143:3823-3827, doi:10.1175/MWR-D-14-00322.1
16. 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.
17. Fowler, A.M.; Lawless, A.S. An idealized study of coupled atmosphere-ocean 4D-var in the presence of model error. *Mon. Wea. Rev*. **2016**, 144:4007-4030.
18. 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.
19. Penny, S.G.; et al. Coupled data assimilation for integrated earth system analysis and prediction: Goals, challenges and recommendations. WWRP 2017-3 Report. WMO Geneva 2017 pp50.
20. Penny, S.G.; Hamil, T.M. Coupled data assimilation for integrated earth system analysis and prediction. *Bull. Amer. Meteorol. Soc*. **2017**, DOI:10.1175/BAMS-D-17-0036.1
21. Storto, A.; Martin, M.J.; Deremble, B.; Masina, S. Strongly coupled data assimilation experiments with linearized ocean-atmosphere balance relationship. *Mon. Wea. Rev*. **2018**, 146:1233-1257, doi: 10.1175/MWR-D-17-0222.1.
22. Laloyaux, P.; Frolov. S.; Ménétrier, B.; Bonavita, M. Implcit and explicit cross-correlations in coupled data assimilation. *Q. J. R. Meteorol. Soc.* **2018**, 144:1851-1863.
23. 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. Rem. Sen.* **2012**, 50:4227-4242.
24. de Rosnay, P.; Drusch, M.; Vasiljevic, D.; Balsamo, G.; Albergel, C.; Isaksen, L. A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF. *Q. J. R. Meteor. Soc*. **2013**, 139, 1199–1213, doi:10.1002/qj.2023.
25. 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. Wea. Rev*. **2015**, 143, 4678–4694, doi:10.1175/MWR-D-15-0174.1.
26. Lin, L.-F.; Pu, Z. Characteristics of background error covariance of soil moisture and atmospheric states in strongly coupled land-atmosphere data assimilation. *Mon. Wea. Rev*. **2018**, 57, 2507-2528. Doi:10.1175/JAMC-D-18-0050.1.
27. 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 golbal deterministic prediction system. *Weather and Forecasting*. **2019**, 34, (submitted).

28. Lee, E.; Županski, M.; Županski, D.; Park, S.K. Impact of OMI aerosol optical depth on analysis incrments through coupled meteorology-aerosol data assimilation for an Asian dust storm. *Remote Sensing of Environment* **2017**, 193:38-53. http://dx.doi.org/10.1016/j.rse2017.02.013 .
29. 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**, https://doi.org/10.5194/acp-2018-1249.
30. Zupanski, 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 conditiona in a strongly-coupled data assimilation. *Atmos. Chem. Phys. Discuss*. **2019**, https://doi.org/10.5194/acp-2019-2 .
31. Benedetti, A.; Vitart, F. Can the direct effect of aerosols improve subseasonal predictability ? *Mon. Wea. Rev*. **2018**, 146:3481-3498. DOI:10.1175/MWR-D-17-0282.1
32. Randles, C.A.; da Silva, A.M.; Buchard, V.; Colarco, P.R.; Darmenov, A.; Govindaraju, R.; Smirnov, A.; Holben, B.; Ferrere, R.; Hair, J.; Shinozuka, Y.; Flymm, C.J. The MERRA-2 aerosol reanalysis, 1980 onward. Part I: system description and data assimilation evaluation. *J. Clim*. **2017**, 30:6823-6850, DOI:10.1175/JCLI-D-16-0609.1
33. Bocquet, M., H. Elbern, H. Eskes, M. Hirtl, R. Žabkar, G.R. Carmicheal, J. Flemming, A. Inness, M. Pagaoski, J.L. Pérez Camaño, P.E. Saide, R. San Jose, M. Sofiev, J. Vira, A. Baklanov, C. Carnevale, G. Grell, and C. Seigneur. Data assimilation in atmospheric chemistry models; current status and future prospects for coupled chemistry meteorology models. *Atmos. Chem. Phys*., **2015**, 15 :5325-5358, www.atmos-chem-phys-net/15/5325/2015/ , doi: 10.5194/acp-15-5325-2015.
34. Saide, P.E.; Carmichael, G.R.; Spak, S.N.; Minnis, P.; Ayers, J.K. Improving aerosol distributions below clouds by assimilating satellite-retreived cloud droplet number. *Proc. Natl. Aca. Sci*. **2012**, 109:11939-11943, doi:10.1073/pnas.1205877109 .
35. Park, S.K.; Lim, S.; Zupanski, M. Structure of forecast error covariance in coupled atmosphere-chemistry data assimilation. *Geosci. Model Dev*. **2015**, 8:1315-1320, www.geosci-model-dev.net/8/1315/2-15 , doi:10.5194/gmd-8-1315-2015.
36. Dragani, R.; McNally, A.P. Operational assimilation of ozone-sensitive infrared radiances at ECMWF. *Q. J. R. Meteorol. Soc*. **2013**, 139:2068-2080. https://doi.org/10.1002/qj.2106 .
37. Engelen, R.J.; Bauer, P. The use of variable CO2 in the data assimilation of AIRS and IASIS radiances. *Q. J. R. Metorol. Soc*. **2014**, 140:958-965 .
38. Morcrette, J.-J.: Ozone–radiation interactions in the ECMWF forecast system. ECMWF Tech. Memo 375, 2003, 36 pp.
39. De Grandpré, J.; Ménard, R., Rochon, Y.J.; Charrette, C.; Chabrillat, S.; Robichaud, A. Radiative impact of ozone on temperature predictability in a coupled chemistry-dynamics data assimilation system. *Mon. Wea. Rev*. **2009**, 137:679-692. doi: 10.1175/2008MWR2572.1.
40. 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.
41. Daley, R. Estimating the windfield from chemical constituent observations: experiments with an extended Kalman filter. *Mon. Wea. Rev*. **1995**, 123:181-198.
42. Riishojgaard, L. P. On four-dimensional variational assimilation of ozone data in weather prediction models. *Q. J. R. Meteorol. Soc*. **1996**, 122:1545-1571.
43. 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.
44. 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. www.atmos-chem-phys.net/9/4855/2009/.

45. Milewski, T.; Bourqui, M.S. Assimilation of stratospheric temperature and ozone with an ensemble Kalman filter in a chemistry-climate model. *Mon. Wea. Rev*. **2011**, 139:3389-3404. doi:10.1175/2011MWR3540.1
46. Milewski, T.; Bourqui, M.S. Potential of an ensemble Kalman smoother for stratospheric chemical-dynamical data assimilation. *Tellus* **2013**, 65, 18541, http://dx.doi.org/10.3402/tellusa,v65i0.18541.
47. 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. www.atmos-chem-phys.bnet/13/3501/2013/ , doi:10.5194/acp-13-3501-2013.
48. 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, www.atmos-chem-phys.net/14/3347/2014/ , doi:10.5194/acp-14-3347-2014.
49. 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. www.atmos-chem-phys.net/15/5835/2015 , doi:10.5194/acp-15-5835-2015.
50. Bocquet, M.; Sakov, P. Joint state and parameter estimation with an iterative ensemble Kalman smoother. *Nonlin. Processes Geosphys*. **2013**, 20:803-818, doi:10.5194/npg-20-803-2013.
51. 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. www.geosci-model-dev.net/9/393/2016 , doi:10.5194/gmd-9-393-2016.
52. Ménard, R.; Chabrillat, 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, 4 December 2007, Noorwijk, Netherlands. 486 pp. [Executive summary available from http://esamultimedia.esa.int/docs/gsp/completed/C18560ExS.pdf].
53. 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(2):103-156.
54. Lorenc, A.; Rawlins, F. Why does 4D-Vat beat 3D-Var ?. *Q. J. R. Meterol. Soc*. **2005**, 131:3247-3257. Doi:10.1256/qj.05.85.
55. Courtier, P. Dual formulation of four-dimensional variational assimilation. *Q. J. R. Meteorol. Soc*. **1997**, 123, 2449-2461.
56. Gauthier, P.; Tanguay, M.; Laroche, S.; Pellerin, S. Extension of 3DVAR to 4DVAR : Implementation of 4DVAR at the meteorological service of Canada. *Mon. Wea. Rev.* **2007**, 135, 2339-2354. DOI: 10.1175/MWR3394.1 .
57. 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, [https//agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2000JD900124]. Doi:10.1029/2000JD900124.
58. 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. Wea. Rev*. **2016** ,144:1035-1050.
59. 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. doi:10.5194/acp-12-10015-2012.
60. 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.
61. Derber, J.; Bouttier, F. A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus* **1999**, 51A:195-221.
62. Gilbert, J.C.; Lemaréchal, C. Some numerical experiments with variable-storage quasi-Newton algorithms. *Math. Programm*. **1989**. 45:407-435.
63. Boer, G. Homogenous and isotropic turbulence on the sphere. *J. Atmos. Sci*. **1983**, 40:154-163.
64. Gauthier, P.; Courtier, P.; Moll, P. Assimilation of simulated wind lidar data with a Kalman filter. *Mon. Wea. Rev*. **1992**, 121:1803-1820.

65. 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.
66. Parrish, D.F.; Derber, J.C. The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev*. **1992**, 120:1747-1763.
67. Heckly, W.A.; Courtier, P.; Pailleux, J.; Andersson, E. The ECMWF variational analysis : General formulation and use of background information. *ECMWF Workshop on Variational Assimilation, with special emphasis on three-dimensional aspects*. Reading 9-12 November 1992, p 49-94.
68. 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.
69. 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, U.K., 1998.
70. Ménard, R.; Deshaies-Jacques, J. Evaluation of analysis by cross-validation, PartII: Diagnostic and optimization of analysis error covariance. *Atmosphere* **2018**, 9(2) 70; doi:10.3390/atmos9020070, https://doi.org/10.3390/atmos9020070
71. Caines, P.E. *Linear Stochastic Systems*. John Wiley and Sons, New York, NY, USA, 1998, p 874.
72. Laroche, S.; Gauthier, P. A validation of the incremental formulation of 4D variational data assimilation in a nonlinear barotropic flow. *Tellus*. **1998**, 50A, 557-572.
73. Lawless, A.S.; Nichols, N.K.; Boess, C.; Bunse-Gerstner, A. Using model reduction methods whithin incremental four-dimensional variational data assimilation. *Mon. Wea. Rev*. **2008**, 1511- 1522. DOI: 10.1175/2007MWR2103.1 .
74. Polavarapu, S.; Tanguay, M.; Ménard, R.; Staniforth, A. The tangent linear model for semi-Lagrangian schemes: linearizing the process of interpolation. *Tellus*. **1996**, 48A, 74-95.
75. Tanguay, M.; Polavarapu, S. The adjoint of the semi-Lagrangian treatment of the passive tracer equation. *Mon. Wea. Rev*. **1999**, 127, 551-546.
76. Rutherford, I.D. Data assimilation by statistical interpolation of forecast error fields. *J. Atmos. Sci*. **1972**, 29:809-815
77. Hollingsworth, A., and P. Lönnberg. 1986. The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. Tellus A 38:111–36. doi:10.1111/j.1600-0870.1986.tb00460.x
78. Desroziers, G., L. Berre, B. Chapnik, and P. Poli. Diagnosis of observation-, background-, and analysis-error statistics in observation space. Q. J. Roy. Meteorol. Soc. **2005**, 131, 3385-3396.
79. 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.
80. 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, DOI:10.1002/qj.2650.
81. Bouttier, F. 1994. *Sur la prévision de la qualité des prévisions météorologiques*. Ph.D. dissertation, Université Paul Sabatier, 240 pp. [Available from Université Paul Sabatier, Toulouse, France.]
82. Janjic, T., N. Bormann, M. Bocquet, J.A. Carton, S.E. Cohn, S.L. Dance, S.N. Losa, N.K. Nichols, R. Potthast, J.A. Waller, and P. Weston. On the representation error in data assimilation. *Q. J. R. Meteorol. Soc*., **2017**, DOI:10.1002/qj.3130.
83. Pereira, M.B.; Berre, L. The use of an ensemble approach to study the background error covariance in a global NWP model. *Mon. Wea. Rev*. **2006**, 134, 2466-2489.
84. 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. *Journal of the Air & Waste Management Association* **2016**, 66:874-895 DOI:10.1080/10962247.2016.1177620, http://dx.doi.org/10.1080/10962247.2016.1177620
85. Robichaud, A.; Ménard, R.; Chabrillat, 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, www.atmos-chem-phys.net/10/1739/2010.

86. Riijshøjgaard, L.P., and E, Källén. On the correlation between ozone and potential vorticity for large scale Rossby waves. *J. Geophys. Res*. **1997**, 107, 8793-8804.
87. Li, Y., R. Ménard, L.P. Riishøjgaard, S.E. Cohn, and R.B. Rood. A study on assimilating potential vorticity data. Tellus. **1998**, 50A, 490-506.
88. Allaart, M. A. F., Kelder, H. and Heijboer, L. C. On the relation between ozone and potential vorticity. Geophys. Res. Lett. **1993**, 20, 811–814.
89. Ménard, R. Bias estimation. In *Data Assimilation: Making Sense of Observations* (eds. Lahoz, W., B. Khattatov, and R. Ménard). 2010, Springer, New York, 113-136.
90. Di Tomaso, E., Bormann, N. Assimilation of ATOVS radiances at ECMWF: first year EUMETSAT fellowship report. *EUMETSAT/ECMWF Felloship Programme Research Report 22*, 2011, ECMWF, reading UK., 27 pp.
91. 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.
92. Ménard, R.; Polavarapu, S.; Yang, Y. Model error estimation: its application to chemical data assimilation. *Proceedings of the ECMWF/SPARC Workshop*, Shinfield Park, UK, June 23-26 2003, http://www.ecmwf.int/sites/default/files/elibrary/2004/11098-model-error-estimation-its-application-chemical-data-assimilation.pdf,137-146