CAMS Service Evolution

D2.5 Report on weak-constraint 4DVar, concept

Due date of deliverable	31 May 2024
Submission date	29 May 2024
File Name	CAMEO-D2-5-V1.0
Work Package /Task	2/2.4
Organisation Responsible of Deliverable	ECMWF
Author name(s)	Noureddine Semane
Revision number	V1.0
Status	Issued
Dissemination Level	PUBLIC

Funded by the European Union

The CAMEO project (grant agreement No 101082125) is funded by the European Union.
Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Commission. Neither the European Union nor the granting authority can be held responsible for them.

1 Executive Summary

In WP2 of the CAMEO project task 2.4 aims at improving the CAMS data assimilation (DA) methodology by taking model errors and dynamical constraints of long-lived trace gases into account through the application of weak constraint 4DVar (WC-4DVar).

We successfully implemented WC-4DVAR for stratospheric ozone in the first 17 months of the CAMEO project and extensively tested it in several DA experiments both in the ECMWF CAMS and NWP configurations.

The evaluation of WC-4DVar including ozone shows a positive impact for both CAMS and NWP when the comparison is done against the assimilated observations and a neutral impact when verifying against independent data from ozonesondes.

Table of Contents

2 Introduction

2.1 Background

Variational data assimilation 4DVar (time plus three space dimensions) aims to find a model trajectory that best fits (in a least squared sense) observations over an assimilation time window by adjusting the initial conditions supplied for forward model integration. On the one hand, in the strong constraint 4DVar, it is assumed that the forward model perfectly represents the evolution of the actual atmosphere, and the best fit model trajectory is obtained by adjusting only the initial conditions via minimization of a cost function, subject to the model equations as strong constraint. On the other hand, relaxing the assumption that the model is perfect leads to the weak constraint 4DVar formulation (WC-4DVar), in which the model error is introduced as a correction to the time derivatives of model variables, and the best fit model trajectory is obtained by adjusting simultaneously both model error and initial conditions (Fisher et al. 2005, Trémolet, 2006, 2007).

WP2 of the CAMEO project task 2.4 aims at improving the CAMS DA methodology by taking model errors and dynamical constraints of long-lived trace gases into account through the application of weak constraint 4DVar. Following the efficiency of WC-4DVar at correcting model systematic errors for temperature, divergence, and vorticity in the stratosphere (Laloyaux et al., 2020a, 2020b), this document reports on its extension to include stratospheric ozone in both the Atmospheric Composition (CAMS) and Numerical Weather Prediction (NWP) configurations.

2.2 Scope of this deliverable

2.2.1 Objectives of this deliverable

This deliverable describes the work carried out in the first 17 months of the project.

2.2.2 Work performed in this deliverable

In this deliverable the work as planned in the Description of Action (DoA, WP2 T2.4) was performed.

2.2.3 Deviations and counter measures

No deviations have been encountered.

2.2.4 CAMEO Project Partners:

ECMWF	EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS
Met Norway	METEOROLOGISK INSTITUTT
BSC	BARCELONA SUPERCOMPUTING CENTER-CENTRO NACIONAL DE SUPERCOMPUTACION
KNMI	KONINKLIJK NEDERLANDS METEOROLOGISCH INSTITUUT-KNMi
SMHI	SVERIGES METEOROLOGISKA OCH HYDROLOGISKA INSTITUT

BIRA-IASB	INSTITUT ROYAL D'AERONOMIE SPATIALEDE BELGIQUE
HYGEOS	HYGEOS SARL
FMI	ILMATIETEEN LAITOS
DLR	DEUTSCHES ZENTRUM FUR LUFT - UND RAUMFAHRT EV
ARMINES	ASSOCIATION POUR LA RECHERCHE ET LE DEVELOPPEMENT DES METHODES ET PROCESSUS INDUSTRIELS
CNRS	CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE CNRS
GRASP-SAS	GENERALIZED RETRIEVAL OF ATMOSPHERE AND SURFACE PROPERTIES EN ABREGE GRASP
CU	UNIVERZITA KARLOVA
CEA	COMMISSARIAT A L ENERGIE ATOMIQUE ET AUX ENERGIES ALTERNATIVES
MF	METEO-FRANCE
TNO	NEDERLANDSE ORGANISATIE VOOR TOEGEPAST NATUURWETENSCHAPPELIJK ONDERZOEK TNO
INERIS	INSTITUT NATIONAL DE L ENVIRONNEMENT INDUSTRIEL ET DES RISQUES - INERIS
IOS-PIB	INSTYTUT OCHRONY SRODOWISKA - PANSTWOWY INSTYTUT BADAWCZY
FZJ	FORSCHUNGSZENTRUM JULICH GMBH
AU	AARHUS UNIVERSITET
ENEA	AGENZIA NAZIONALE PER LE NUOVE TECNOLOGIE, L'ENERGIA E LO SVILUPPO ECONOMICO SOSTENIBILE

3 Stratospheric ozone chemistry and 4DVar setup

3.1 Stratospheric ozone chemistry

On the one hand, the ECMWF NWP applies only a linear parameterization (Hybrid Linear Ozone scheme, HLO) to describe stratospheric ozone assuming that the net chemical production or destruction of ozone can be described as a specified function of latitude, height, and time of year, with perturbations that are locally linear functions of temperature and the total column ozone (Cariolle and Déqué 1986, Cariolle and Teyssèdre, 2007). The HLO scheme simulates the impact of chemistry on the tendency of ozone, using a blend of inputs from analyses and chemistry model calculations.
On the other hand, CAMS uses an explicit chemical mechanism to simulate chemistry in the stratosphere based on the chemical module originally developed for the Belgian Assimilation System for Chemical ObsErvations (Errera et al.,2019, BASCOE) to assimilate satellite observations of stratospheric composition. Note that ozone is interactive with radiation in both the ECMWF CAMS and NWP configurations since the ECMWF model cycle CY48R1. The quality of simulated ozone depends on the observations assimilated, the data assimilation method and background errors (i.e. quality of the analysis).

3.2 Weak constraint 4DVar formulation

Let the vector $\boldsymbol{x}_k$ be used to represent the state of the atmosphere at the time k, then its evolution accounting for the model error is written as,

$$\boldsymbol{x}_{k+1} = M(\boldsymbol{x}_k) + \boldsymbol{\eta}$$

where M represents the model and η its error. In this implementation, the model error tendencies are considered constant within the entire assimilation window.

The WC-4DVar cost function is given by:

$$\mathrm{J}(\boldsymbol{x}_0, \boldsymbol{\eta}) = \frac{1}{2}(\boldsymbol{x}_0 - \boldsymbol{x}_0^b)^T \mathbf{B}^{-1}(\boldsymbol{x}_0 - \boldsymbol{x}_0^b) + \frac{1}{2}\sum_{k=0}^{N}(H(\boldsymbol{x}_k) - \boldsymbol{y}_k)^T \mathbf{R}_k^{-1}(H(\boldsymbol{x}_k) - \boldsymbol{y}_k) + \frac{1}{2}(\boldsymbol{\eta} - \boldsymbol{\eta}^b)^T \mathbf{Q}^{-1}(\boldsymbol{\eta} - \boldsymbol{\eta}^b) \quad (2)$$

where $\boldsymbol{\eta}^b$ is the prior estimate of the model error forcing (which are the model error tendencies) estimated in the previous WC-4DVar analysis update and $\mathrm{Q} = \mathrm{E}[\boldsymbol{\eta}\boldsymbol{\eta}^T]$ is the model error covariance matrix, also called the Q matrix, where E represents the expected value. Comparing the strong and weak constraint, in the formulation of the former, it is assumed that $\boldsymbol{\eta} = \boldsymbol{\eta}^b = 0$.

Ozone assimilation relies on prior background constraints (i.e. the short-range forecast of ozone and its error covariance) to control the vertical distribution of ozone information. Note that the largest background errors are located around the level of maximum ozone concentration (Han and McNally, 2010).

3.3 Neural Network derived model error covariance

The assimilated ozone observations in the ECMWF CAMS and NWP configurations are listed in Table 1. The ozone-sensitive IR radiances (assimilated in NWP but not in CAMS) have the largest impact in the lower stratosphere below the level of maximum ozone concentration (Han and McNally, 2010; Dragani and McNally, 2013). Ozone increments are zero-ed in the top 15 model levels (above ~ 1hPa) in the CAMS configuration.

Table 1: Ozone observations

Ozone observations	CAMS	NWP
AURA MLS	Assimilated	Not assimilated
SNPP OMPS-nadir	Assimilated	Not assimilated
METOP-B GOME-2	Assimilated	Assimilated
METOP-C GOME-2	Assimilated	Not assimilated
AURA OMI	Assimilated	Assimilated
S5P TROPOMI	Assimilated	Not assimilated
AIRS/CRIS/IASI ozone-sensitive IR radiances	Not assimilated	Assimilated

In Bonavita and Laloyaux, 2020, it was shown how to set up and train an Artificial Neural Network (ANN) to emulate the model error. This ANN implements a nonlinear regression model that is trained to learn cumulated model errors over a 12-hour assimilation window using analysis increments as predictands and a combination of climatological (lat, lon, time_of_day, month) and state-dependent fields (columns of background fields) as predictors. The predicted cumulated errors are then scaled by the length of the assimilation window (12 h) to provide model error tendencies.

The training data set for the CAMEO work consisted of analysis increments and background forecasts collected over the whole year of 2021 for CAMS and NWP configurations (ECMWF model cycle cy48R1). The trained ANN is used as a generative model of model error, i.e., we use the ANN to generate a representative sample of model errors and compute a Q matrix from this sample. The emulator is acting as a filter, learning only the predictable part of the model error and filtering out the randomness component. Q is computed offline and does not change during the experiment, i.e. is a climatological Q. The computed Q matrix has been localised in the horizontal (with a cosine function tapering the correlations to zero between 4000 and 6000 km) to remove spurious hemispheric-wide correlations; and in the vertical, with a quadratic function of the distance from the diagonal to control sampling noise (Bonavita and Laloyaux, 2022).

Figure 1 shows vertical profiles of the standard deviation of ANN-derived model error tendencies for CAMS (left panel) and NWP (right panel). The diagnosed model error standard deviations are significantly different for CAMS and NWP configurations. The difference is due to the ozone-analysis dependence on the stratospheric chemistry scheme, the observations assimilated, and background errors.

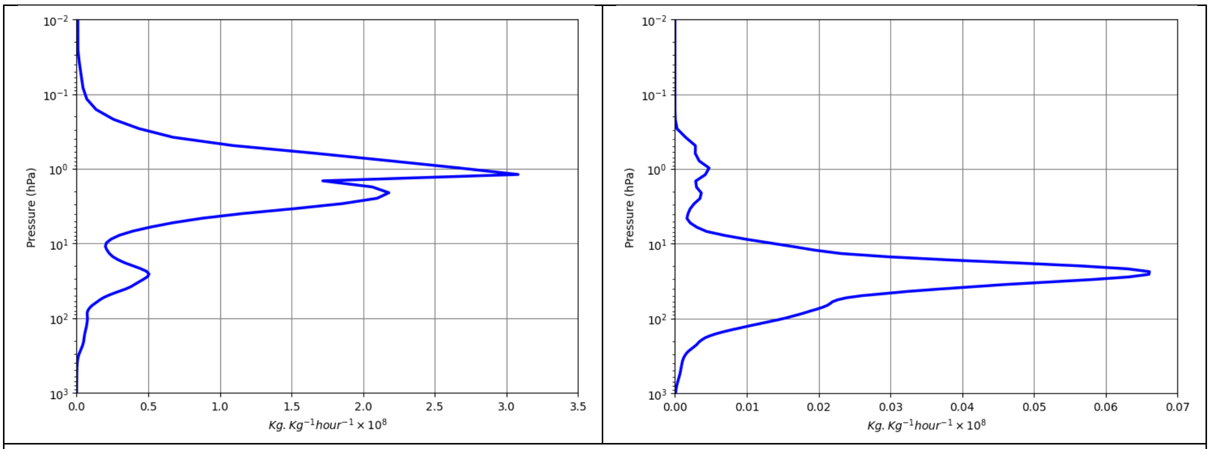

Figure 1: Vertical profiles of the standard deviation of ANN-derived model error tendencies for the CAMS (left panel) and NWP (right panel).

The structure of vertical correlations of model error is a crucial aspect of the Q matrix, as it determines how model error information is spread in the atmospheric column. As depicted in Figure 2, CAMS and NWP show substantial differences in upper levels due to different stratospheric chemistry schemes.

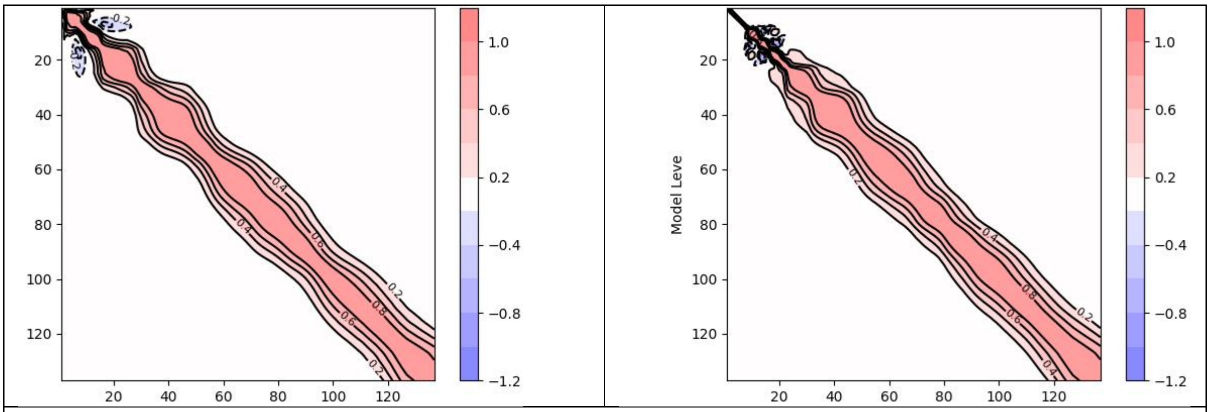

Figure 2: Vertical correlations of ANN-derived model error tendencies, before applying localisation, for CAMS (left panel) and NWP (right panel). Model level 1 is the top of the atmosphere, model level 137 corresponds to the surface.

As shown in Figure 3, the model error horizontal correlation structure functions (red curves) are broader than those of the background error (black curves).

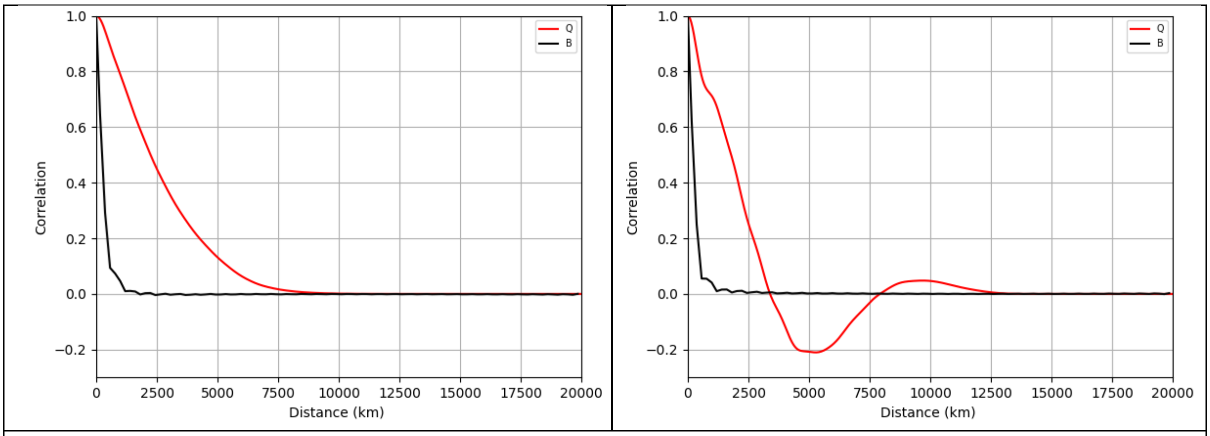

Figure 3: ANN-derived model error (red) and background error (black) horizontal correlation structure functions, before applying localisation, at model level 36 (~20 hPa) for CAMS (left panel) and NWP (right panel).

4 WC-4DVar results

To examine the impact of weak constraint on the data assimilation system, summer (JJA 2022) and winter (DJF 2022-2023) season experiments were run in ECMWF model cycle cy49R1 with and without weak constraint for stratospheric ozone for both CAMS and NWP configurations. WC-4DVar was run with a reduction of the standard deviation of the model error by a factor 20 for CAMS and a factor 4 for NWP over the 5-50hPa vertical range. The standard deviation of the model error is set to zero outside the 5-50hPa vertical range. The latter restriction is chosen as we want, as a first step, to focus only on the ozone maximum level (~20 hPa) where the background errors are the largest.

4.1 DA departure statistics

Analysis and background (i.e. the short-range forecast of ozone) departures statistics (shown below) from a cycling data assimilation experiment are one of the main tools to verify the effectiveness of any upgrade to the data assimilation system.

Figure 4 shows a significant reduction in analysis and background observation departures for MLS around 20 hPa if WC-4DVar is used for ozone in the CAMS configuration.

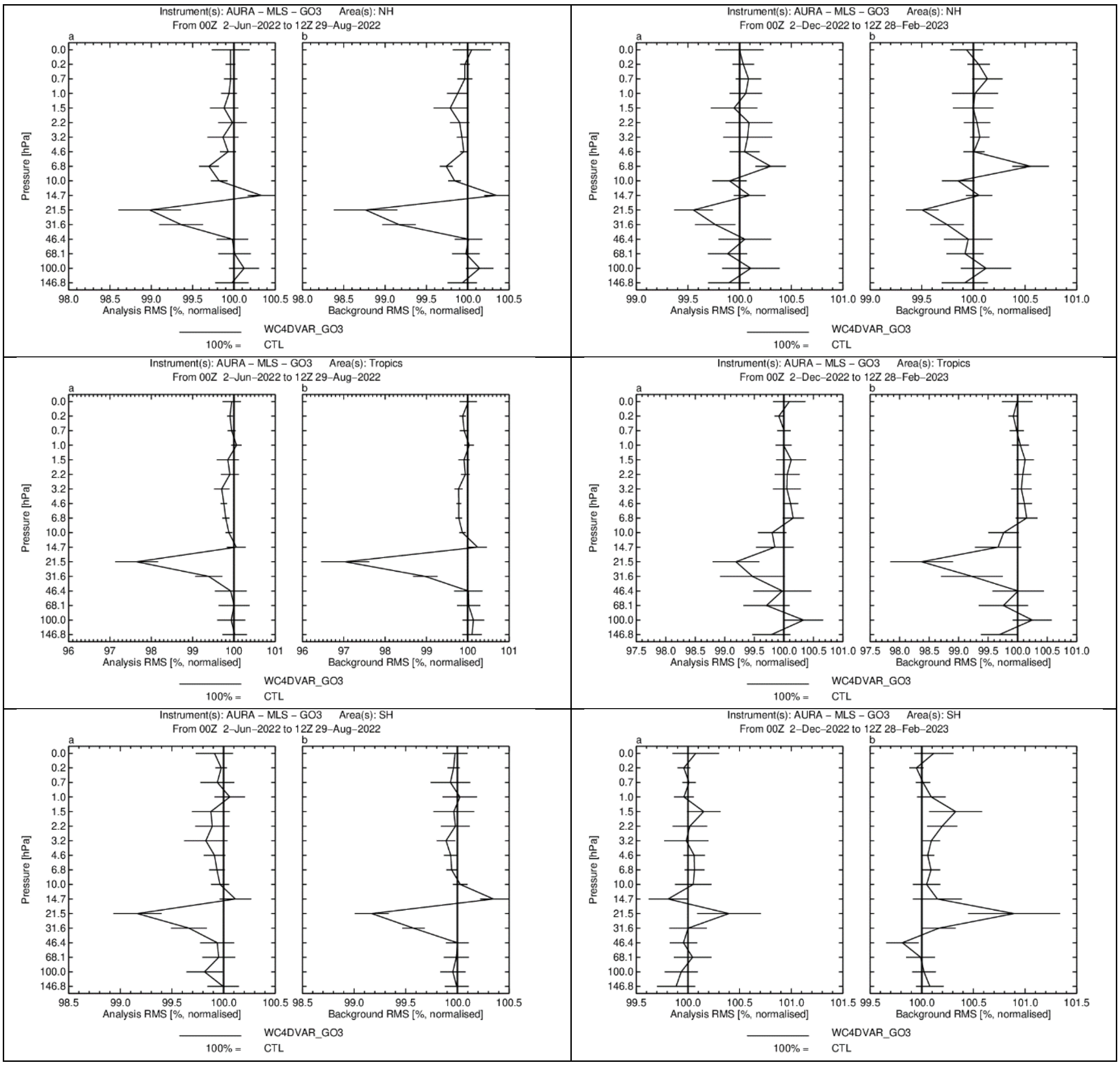

Figure 4: Relative change in the rms analysis and background fits for MLS ozone profiles for CAMS configuration in JJA 2022 (right panels) and DJF 2022-2023 (left panels): Northern Hemisphere (top panels); Tropics (middle panels); Southern Hemisphere (bottom panels). Values lower than 100% indicate that the WC-4DVar experiment has smaller analysis/background departures than the control 4DVar.

Figure 5 shows a significant reduction in analysis and background observation departures for IASI ozone-sensitive channels between wavenumber 1000 and 1100 cm^{-1} if WC-4DVar is used for ozone in the NWP configuration.

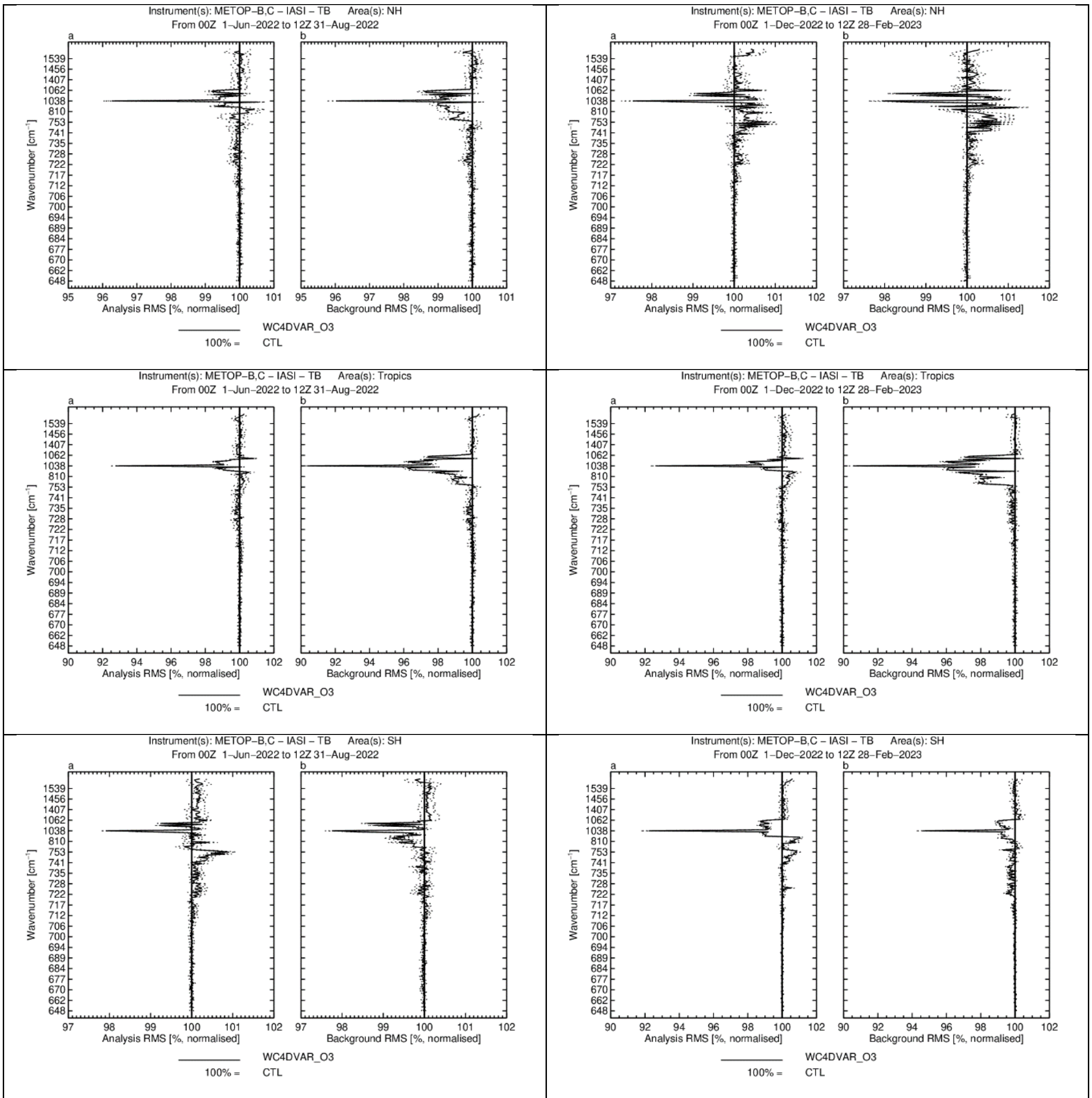

Figure 5: Relative change in the rms analysis and background fits for IASI hyperspectral Infrared sensor in JJA 2022 (right panels) and DJF 2022-2023 (left panels) for NWP configuration: Northern Hemisphere (top panels); Tropics (middle panels); Southern Hemisphere (bottom panels). Values lower than 100% indicate that the WC-4DVar experiment has smaller analysis/background departures than the control 4DVar. IASI ozone-sensitive channels are between wavenumber 1000 and 1100 cm^{-1} .

4.1 Verification against ozone sondes

The comparison against ozone sondes shows, as depicted in Figures 6 and 7, a neutral impact for both CAMS and NWP configurations.

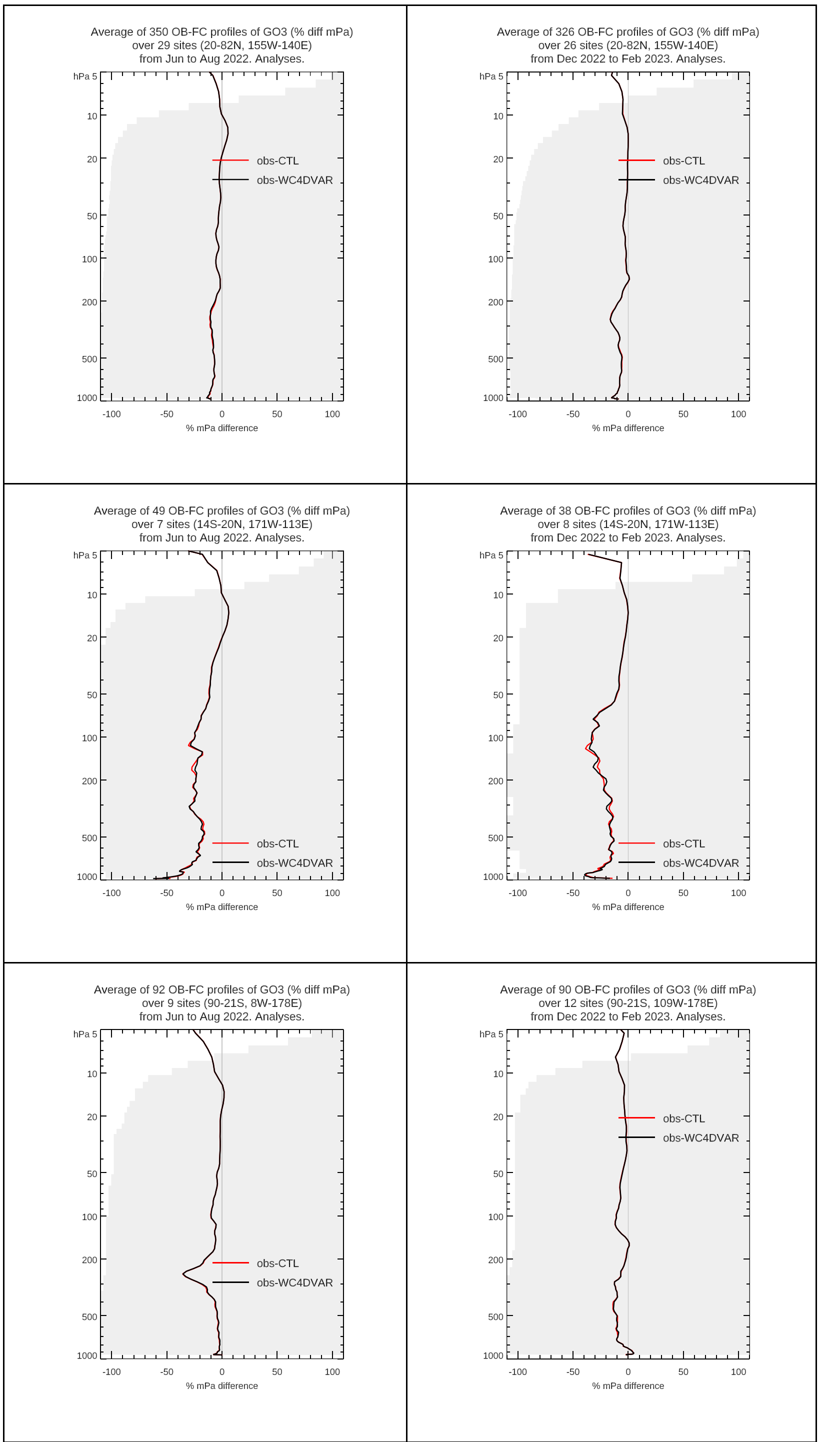

Figure 6: Mean Background fit (normalised by observation value) to ozonesondes for JJA 2022 (right panels) and DJF 2022-2223 (left panels) for CAMS: Northern Hemisphere (top panels); Tropics (middle panels); Southern Hemisphere (bottom panels).

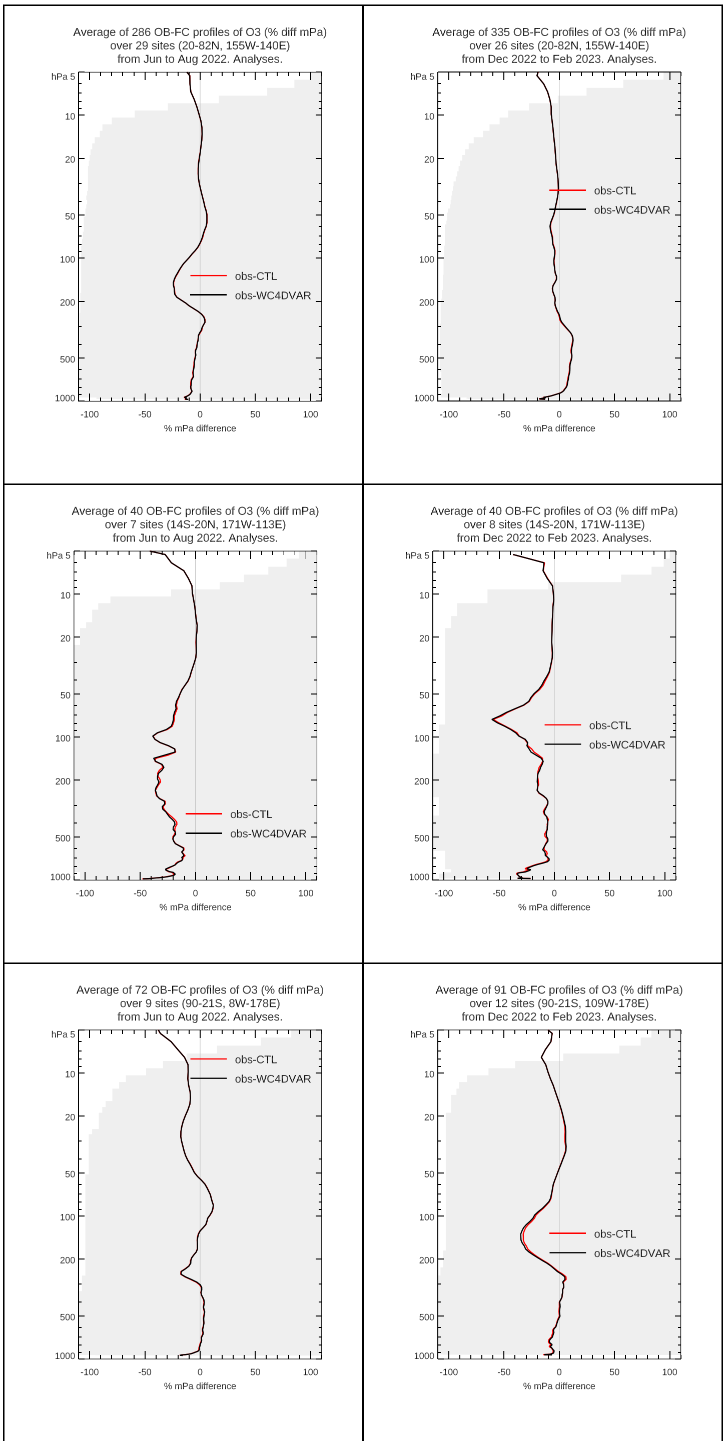

Figure 7: Mean Background fit (normalised by observation value) to ozonesondes for JJA 2022 (right panels) and DJF 2022-2223 (left panels) for NWP: Northern Hemisphere (top panels); Tropics (middle panels); Southern Hemisphere (bottom panels).

5 Conclusion

In WP2 of the CAMEO project task 2.4 aims at improving the CAMS DA methodology by taking model errors and dynamical constraints of long-lived trace gases into account through the application of weak constraint 4DVar. Following the efficiency of WC-4DVar at correcting model systematic errors for temperature, divergence, and vorticity in the stratosphere (Laloyaux et al., 2020a, 2020b), this document reports on its extension to include stratospheric ozone in both the Atmospheric Composition (CAMS) and Numerical Weather Prediction (NWP) configurations.

We successfully implemented WC-4DVAR for stratospheric ozone in the first 17 months of the CAMEO project and tested it in several data assimilation experiments both in the ECMWF CAMS and NWP configurations. Model bias correction of stratospheric ozone in the 5-50hPa vertical range shows that the corrected first-guess trajectory better fits the infrared channels that are sensitive to ozone (for NWP) and the MLS ozone profiles (for CAMS). While the WC-4DVAR for stratospheric ozone is now technically working, further work, including longer assimilation experiments and a detailed study of the initial condition increments and the increments of the strong constraint 4D-Var, is needed to assess why there is no improvement in the fit to independent data from ozonesondes. Further verification is ongoing before potential implementation in the next ECMWF model cycle CY50R1.

6 References

Bonavita, M., and Laloyaux, P., 2022: Estimating Model Error Covariances with Artificial Neural Networks, 2209.11510, arXiv.

Bonavita, M. and Laloyaux, P., 2020: Machine learning for model error inference and correction. Journal of Advances in Modeling Earth Systems, 12, e2020MS002232. https://doi.org/10.1029/2020MS002232

Cariolle, D. and Déqué, M. (1986). Southern hemisphere medium-scale waves and total ozone disturbances in a spectral general circulation model. Journal of Geophysical Research: Atmospheres, 91(D10), 10825-10846.

Cariolle, D. and Teyssedre, H., 2007: A revised linear ozone photochemistry parametrization for use in transport and general circulation models: multi-annual simulations. Atmospheric Chemistry and Physics, 7(9), 2183-2196.

Dragani, R., and McNally, A., 2013: Operational assimilation of ozone-sensitive infrared radiances at ECMWF, Quarterly Journal of the Royal Meteorological Society, 139, 2068–2080.

Errera, Q., Chabrillat, S., Christophe, Y., et al, 2019: Technical note: Reanalysis of AURA MLS chemical observations. Atmospheric Chemistry and Physics, 19(21), 13647–13679, doi:10.5194/acp-19-13647-2019,https://acp.copernicus.org/articles/19/13647/2019/.

Fisher, M, Leutbecher, M, and Kelly, G., 2005: On the equivalence between Kalman smoothing and weak-constraint four-dimensional variational data assimilation, Q. J. R. Meteorol. Soc., 131, 3235–3246.

Han, W. and McNally, A., 2010: The 4D-Var assimilation of ozone-sensitive infrared radiances measured by IASI, Q. J. R. Meteorol. Soc., 136, 2025–2037, 2010.

Laloyaux, P, Bonavita, M, Dahoui, M, et al., 2020a: Towards an unbiased stratospheric analysis, Q J R Meteorol Soc, 146: 2392– 2409. https://doi.org/10.1002/qj.3798

Laloyaux, P., Bonavita, M., Chrust, M., and Gürol, S., 2020b: Exploring the potential and limitations of weak- constraint 4D-Var, Q J R Meteorol Soc., 146: 4067– 4082. https://doi.org/10.1002/qj.3891

Trémolet, Y., 2006: Accounting for an imperfect model in 4D-Var, Q. J. R. Meteorol. Soc., 132, 2483–2504.

Trémolet, Y., 2007: Model error estimation in 4D-Var, Q. J. R. Meteorol. Soc., 133, 1267–1280.

Document History

Version	Author(s)	Date	Changes
0.1	Noureddine Semane	30 April 2024	Initial version
1.0	Noureddine Semane	29 May 2024	First issued version

Internal Review History

Internal Reviewers	Date	Comments
Oriol Jorba (BSC) and Arjo Segers (TNO)	May	Internal review comments and recommendations