TECHNICAL MEMORANDUM

848

Emission optimisations first attempt based on the Ensemble of Data Assimilation for atmospheric composition

Jérôme Barré, Sebastien Massart, Melanie Ades, Luke Jones, Richard Engelen and the CAMS team

Copernicus & Research Departments

June 7, 2019

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen terme

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/en/research/publications

Contact: library@ecmwf.int

European Centre for Medium-Range Weather Forecasts
Shinfield Park, Reading, RG2 9AX, England

Abstract

The main motivation of running ECMWF analyses using the IFS Ensemble of Data Assimilation (EDA) is to provide the best possible initial conditions for forecast initialisation. Atmospheric composition is as much a boundary condition problem as an initial condition problem, hence there is a need to infer surface fluxes in addition to the 3D representation of atmospheric variables. A first attempt to use the ECMWF's IFS Ensemble of Data Assimilation (EDA) information to constrain the Copernicus Atmosphere Monitoring Service (CAMS) surface emissions is presented. With a focus on carbon monoxide, the methodology presented here aims to use the ensemble information and its flow dependency to derive a space and time varying relationship between assimilation increments on the atmospheric composition fields and surface emissions.

This study focused on carbon monoxide (CO) as it is well observed, well modelled and is a primary pollutant mainly originating from anthropogenic and fire sources. The methodology is described and details how the emission ensemble perturbation are done. Using a statistical approach, the ensemble information is then used to derive Jacobians that provide a sensitivity between the increments to the 3D concentration fields and the surface emission or fluxes. We show how the Jacobians are computed and filtered to provide an update on the emission that is carried forward in time to the next assimilation window.

Results show impacts and comparisons of various experiments with assimilation solely (i.e. only the initial conditions are modified by the observations) and assimilation plus inversions. In addition, different prior assumptions for the ensemble emission perturbations and different perturbation persistence times are investigated. Significant changes on the ensemble mean and spread are observed on the surface CO emissions but also in the near surface 3D CO concentrations.

Validation shows that this approach is sensitive to the design of the prior error used for the emissions and how the information is propagated forward in time. Improvement is clearly seen from this approach on the bias scores but not so much in the root mean square errors (RMSE). Increasing the prior error perturbations and persistence time leads to significant RMSE degradation. Discussion about priorities for the next possible implementation phases of such an inversion capability for CAMS are also emphasised.

Contents

1 Introduction **3**

2 Methodology **4**

2.1 Model Setup 4

2.2 Perturbations of surface fluxes 4

2.3 Outer-loop Emission Coupling 5

2.3.1 Jacobians computation and filtering 5

2.3.2 Surface flux optimisation within the outer-loop 6

2.3.3 Surface flux increments forward propagation in time 8

3 Results **9**

3.1 Experiments description 9

3.2 Emission optimisation/inversion results 9

3.3 Evaluation against Airbase and Airnow surface measurement networks 11

4 Discussions and further works **13**

1 Introduction

The European Unions Copernicus Atmosphere Monitoring Service (CAMS) operationally provides daily forecasts of global atmospheric composition. It uses the ECMWF Integrated Forecasting System (IFS) with 4-Dimensional Variational (4D-Var) data assimilation capability, which includes meteorological and atmospheric composition variables and observations, such as reactive gases and aerosols (Inness et al. (2019a)) and greenhouse gases (e.g. Massart et al. (2014),Massart et al. (2016)), for its global forecasts and reanalyses. With atmospheric composition being as much a boundary condition problem as an initial condition problem, there is a strong need to constrain surface fluxes in addition to the 3D representation of atmospheric variables. This is called surface fluxes inversion or emission inversion. Two main motivations drive the long-term development of an operational inversion capability for CAMS: to improve the composition forecasts close to the surface and to inform on potential deficiencies and biases on the emission inventories.

As a first case study for emission inversion within the IFS we have focused on carbon monoxide (CO) for the main following reasons:

- Carbon monoxide is a primary pollutant, serves as a tracer of pollution emission and transport, with both natural and anthropogenic sources. It is directly emitted as a product of incomplete combustion from industrial and urban fossil/biofuel burning as well as largescale biomass burning. It is also directly emitted by plants and oceans, although to a much lesser extent.
- Carbon monoxide is well observed from space and current data assimilation within the CAMS IFS system includes level 2 product (retrievals) of the Infrared Atmospheric Sounding Interferometer (IASI), the Measurement Of Pollution In The Troposphere (MOPITT) and soon the new TROPOspheric Monitoring Instrument (TROPOMI). TropOMI provides a daily global coverage of the troposphere (Inness et al. (2019b)) and opens exiting perspectives regarding the ability to constrain and inform on surface fluxes in a near-real time fashion.
- The current CAMS implementation assumes that the chemical integrator is identity in the tangent linear and adjoint formulation which can cause problems in the analyses results. However, carbon monoxide has an average lifetime of about 1-2 months and therefore can be approximated reasonably well as a non reactive tracer within the current 12 hour IFS assimilation window. This means the tangent-linear and adjoint equations for the chemistry, necessary to correctly employ the 4D-Var minimisation technique, can be simplified to the identity matrix without too great a loss in accuracy. Moreover CO chemistry can be linearised with accuracy (Claeyman et al. (2010)) reducing significantly the numerical costs that are crucial for running ensembles of simulations.

The Ensemble of Data Assimilations (EDA,Bonavita et al. (2012)) is currently implemented at ECMWF to estimate flow dependent background error statistics and provide uncertainty analysis. It is not currently implemented operationally within the CAMS framework but only as research mode Massart and Bonavita (2016). On-going research within CAMS, such as the present study, explore the potential of such EDA methodology for improving atmospheric composition analyses and forecasts. The ensemble information extracted from the EDA is also helpful to derive covariate terms not only in space for the same variable but also between variables. In the case of this paper, we are using the ensemble covariances to link the 3D atmospheric composition fields to the corresponding 2D surface fluxes.

This paper explores the capability of such an EDA system to infer the surface emission or fluxes for pollutants using the ensemble information and a statistical approach. We present and describe the pro-

posed methodology in section 2, followed in section 3 by the evaluation of the results and we end with a discussion of the suggested directions that could be taken in section 4.

2 Methodology

2.1 Model Setup

In this document we use the IFS CY45r1 version in the composition mode. We are using two outer loop minimization at T95 and T159 resolution and 5 days forecasts at T511 with 60 levels and the EDA has 25 members. We use the linear CO scheme implementation for its economical numerical cost instead of the full tropospheric and stratospheric chemical scheme operationally running for CAMS. Our work relies partly on the works of Massart and Bonavita (2016) (thereafter M2016) for setting up an EDA for atmospheric composition. As in the M2016 EDA for composition setup, we kept the same perturbations as for the numerical weather prediction (NWP) EDA with capability to perturb the atmospheric tracer tendencies and the surface fluxes. In this work we bypassed the perturbations of the tracer tendencies and only activated the perturbations of the surface fluxes since the focus of the study is the surface fluxes to 3D concentration sensitivity. Perturbations of the tracer tendencies is another topic of research that needs to be explored carefully but in the context of the linear CO scheme within a 12 hour assimilation window such perturbation effects can be considered as minor. The perturbations of the observations are left activated for composition and consist of adding a random noise on the observation values sampled from a zero-mean Gaussian with variance equal to the prescribed observation variance.

2.2 Perturbations of surface fluxes

The anthropogenic, biogenic and ocean fluxes are provided by inventories whereas the biomass burning and wildfire emissions are provided by the CAMS (Global Fire Assimilation System) GFAS (Kaiser et al. (2012)). The perturbation scheme used for the surface fluxes is derived from the CH_4 perturbation scheme of M2016. We tested and extended this capability to the reactive gases within the CAMS IFS composition configuration, with the ability to separately perturb emissions from fires and the emissions from the anthropogenic,biogenic and ocean components. The perturbations are two-dimensional pseudo-random fields η generated on a regular latitude longitude grid with a zero mean and a standard deviation of one. We recall from M2016 the relationship that is used to perturb the surface fluxes Φ:

$$\Phi_m(\zeta) = \sum_s \Phi(\zeta,s)\big(1+\alpha_s.\eta_m(\zeta,s)\big) \tag{1}$$

Where $\Phi(\zeta,s)$ and $\Phi_m(\zeta)$ are the original emissions and the perturbed emission totals per member m and per sector s at a geographical location ζ. The method uses a global amplitude factor per sector α_s to adjust the a-priori pseudo random perturbations $\eta_m(\zeta,s)$ depending on the sector (see section 3, table 1). In this study we have generated two types of perturbation where the fire emission from GFAS are separated from the emission totals compiled from the inventories (see figure 1). This will allow de-correlation between the two types of emissions. The perturbation correlation length scales have been chosen wider for the inventory as the error is expected to be correlated at the scale of the country and biome, which can extent in the 10^3 km order of magnitude (i.e. 10 grid points on a regular $0.45^o \times 0.45^o$ grid). Fires are sporadic and localised emission events are expected to potentially have narrower error correlation length scale. Fires can also be dependent on the vegetation type (biome) and

seasonal situations such as droughts. Therefore, we choose a correlation length scale in the order of 10^2 km (i.e. 3 grid points on a regular $0.45^o \times 0.45^o$ grid). We look at the impact of such perturbations on the ensemble spread in section 3. Note that independently perturbing emission sectors allows us to retrieve independent signal in the ensemble spread in 3D atmospheric concentrations. The ability to separate fires from the rest of emissions is currently due to the way emission pre-processing is done. Fire emission from GFAS are processed independently from the emissions coming from inventories (e.g. anthropogenic,biogenic,oceans,etc.) and are then merged as a single emission variable at the script level in IFS. Independent perturbations of different emission sectors could be envisioned if the emission pre-processing was done differently.

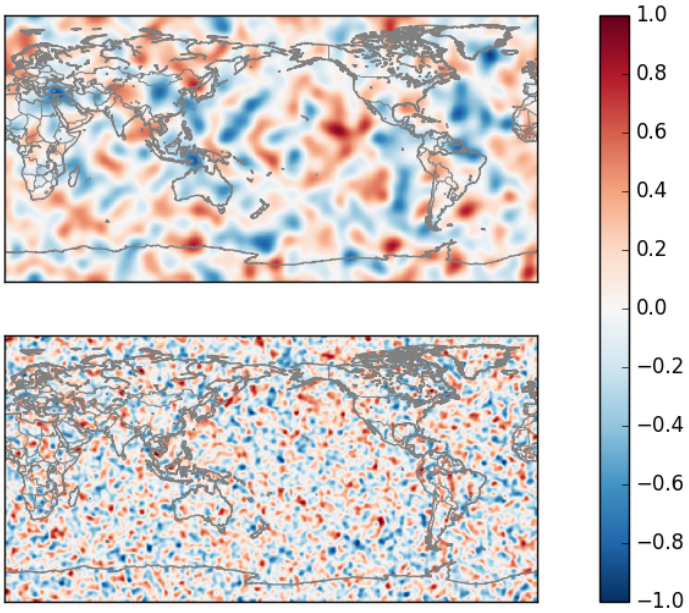

Figure 1: Perturbation structure $\eta_i(\zeta, s)$ for one member used in the set of experiments in this paper. The top panel shows the perturbation used for the merged sectors inventory (anthropogenic, biogenic and ocean). The bottom panel shows the perturbation used on the GFAS fire emissions outputs.

2.3 Outer-loop Emission Coupling

2.3.1 Jacobians computation and filtering

We relate the change in concentrations due to the 4D-Var minimisations (increments) to the surface fluxes by calculating Jacobians using the ensemble forecasts information that contains the full physics and chemistry. In our case the forecasts are run at T511 which is much higher than the two minimizations at T95 and T159. For each model level l and at each longitude latitude ζ we compute the Jacobians that relate the emissions e to the concentrations c, $J_{ec}(\zeta, l) = \partial e / \partial c$ using the linear fit relationship in the

least-squares sense using the ensemble, such as:

$$J_{ec}(\zeta,l) = \frac{\sigma_{e,c}}{\sigma_{c,c}} = r\frac{\sigma_e}{\sigma_c} \tag{2}$$

where r is the correlation coefficient between the ensemble distribution of the concentrations $c(\zeta,l)$ and the surface fluxes $e(\zeta)$, and σ_c, σ_e,$\sigma_{e,c}$ and $\sigma_{c,c}$ their respective standard deviations, covariances and variances computed using the 25 EDA members. The Jacobians are then calculated at analysis time and are 3D fields of dimension ζ,l. As the ensemble size is significantly small (i.e. 25), filtering is required. We then decided to filter using a correlation threshold function $L(r)$, such as:

$$J_{ec}(\zeta,l)_{filtered} = J_{ec}(\zeta,l)L(r) \tag{3}$$

with

$$L(r) = \begin{cases} 0.0 & \text{if } r < r_{min}, \\ (r-r_{min})/(r_{max}-r_{min}), & \\ 1.0 & \text{if } r > r_{max}. \end{cases} \tag{4}$$

we choose r_{min}=0.5 which is the critical correlation value for a given degrees of freedom $(m-2)$ and p-value of 0.01, r_{max}=0.75 which is for the moment arbitrarily chosen to insure a smooth filtering. In the case of CO emissions it is important to remove the anti-correlations as they are not physical, e.g. an increase of CO should not decrease the CO surface fluxes. The filtering then does not allow significant negative correlation values. The figures 2 and 3 show the geographical variation of the magnitude of the Jacobians and the filtering at 1000hPa and 600hPa. Figure 4 shows the time evolution of the Jacobians and filtering and its flow dependent nature given by the ensemble at a given location.

2.3.2 Surface flux optimisation within the outer-loop

After each outer-loop iteration i of the minimisation, increments on the 3D atmospheric concentrations δc_i are produced at an increasing resolution. We then use the filtered Jacobians $J_{ec,i}$ thereafter, interpolated at the current i resolutions (in this case T95 and T159). The increments over the surface fluxes δe_i are then computed as follows,

$$\delta e_i(\zeta) = \sum_l (J_{ec,i}(\zeta,l)\delta c_i(\zeta,l)) \tag{5}$$

which is the sum over the vertical model levels of the product between the increments and the Jacobians. In order to avoid doubling the effects of emission changes and atmospheric concentrations, the share of the concentration increments that has been transferred to the surfaces fluxes needs to be removed using the ”reverse” Jacobians to keep consistency in units, which is expressed as,

$$J_{ce,i} = r\frac{\sigma_c}{\sigma_e}. \tag{6}$$

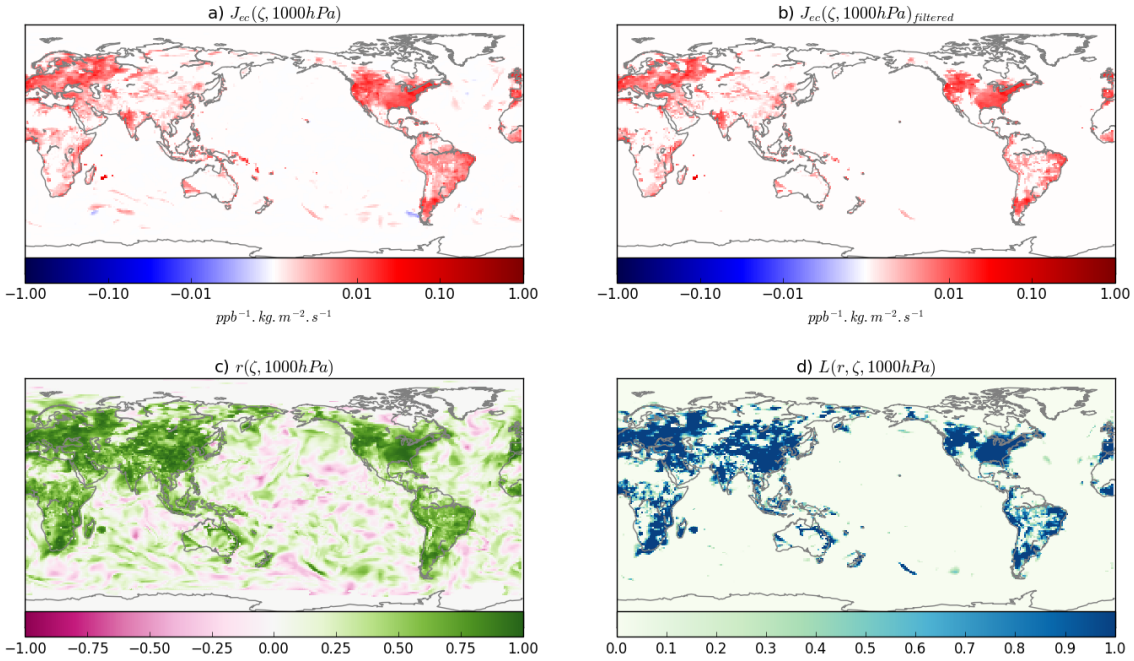

Figure 2: Example of Jacobian filtering at 1000hPa a) Unfiltered Jacobian b) Filtered Jacobian c) Correlation between the ensemble atmospheric CO fields and surface CO fluxes d) Filtering function

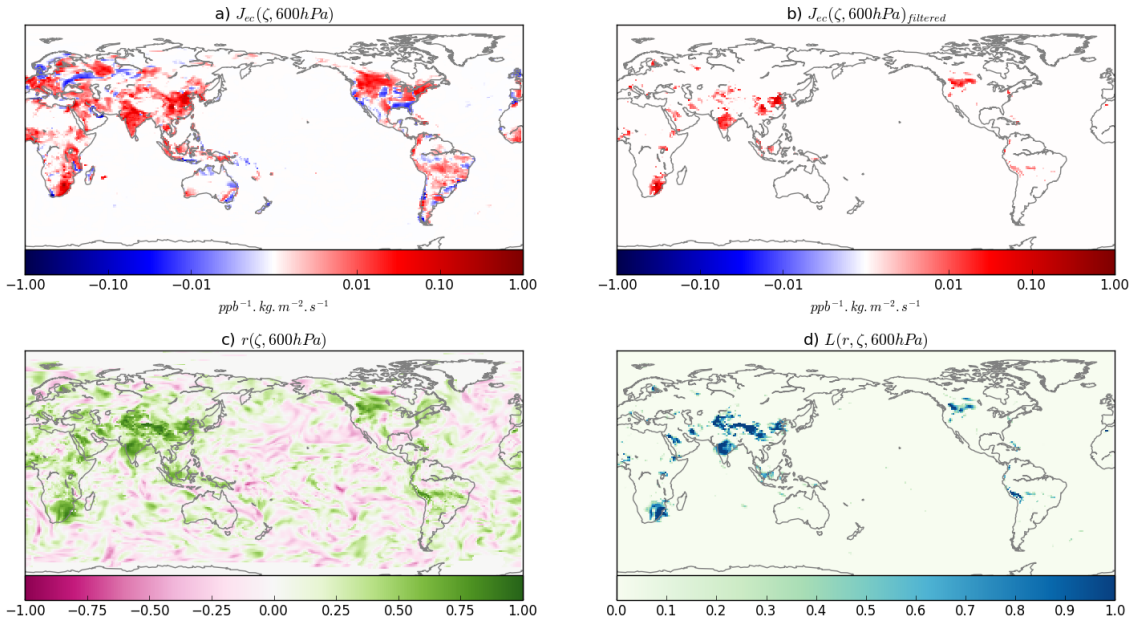

Figure 3: Same as figure 2 but at 600hPa

The adjusted increments are then,

$$\delta c_{i,adjusted}(\zeta,l) = \delta c_i(\zeta,l) - J_{ce,i}(\zeta,l)\delta e_i(\zeta) \tag{7}$$

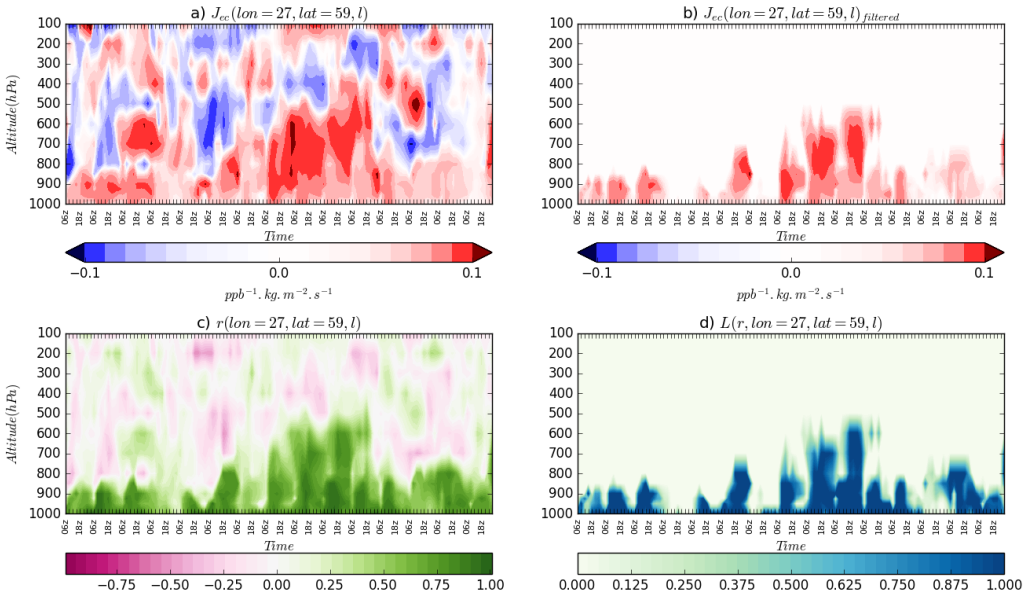

Figure 4: Hovmuler diagram showing the timely variations of the Jacobians with altitude at a given location a) Unfiltered Jacobians b) Filtered Jacobians c) Correlation between the ensemble atmospheric CO fields and surface CO fluxes d) Filtering function

This operation is applied on every member m of the EDA, hence an ensemble of emission increments are produced.

2.3.3 *Surface flux increments forward propagation in time*

After the adjustements on emissions are performed and carried through each outer-loop step, the final ensemble emission increments from the last minimization $\delta e_m(\zeta)$ are interpolated on the forecasts resolution (in this case T511) and then used to adjust the emission perturbations onto the next forecast. In order to make the changes on emissions persistent in time we keep adjusting the perturbation using the emission increments from previous assimilation windows $\delta e'_m(\zeta)$ but damped over time using a decay or persistence factor $\beta \in [0,1]$. The equation 1 then becomes,

$$\Phi_m(\zeta) = \sum_s \Big(\Phi(\zeta,s)\big(1+\alpha_s.\eta_m(\zeta,s)\big)\Big) + \delta e_m(\zeta) + \beta\,\delta e'_m(\zeta) \tag{8}$$

Currently the perturbation adjustment is done on the emission totals only. The methodology presented here allows flexibility with reasonable developments to compute Jacobians per sectors (i.e. allocating grib codes per sectors) and therefore derive emission increments per sector $\delta e_{m,s}(\zeta)$. The equation 8 would then be of the form,

$$\Phi_m(\zeta) = \sum_s \Big(\Phi(\zeta,s)\big(1+\alpha_s.\eta_m(\zeta,s)\big) + \delta e_{m,s}(\zeta) + \beta_s\,\delta e'_{m,s}(\zeta)\Big) \tag{9}$$

Experiment	CO assimilation	CO inversion	$\sigma_{emissioninventories}$	σ_{fires}	$\beta(t_{1/2})$
CR	No	No	20%	25%	—
AR	Yes	No	20%	25%	—
IAR1	Yes	Yes	20%	25%	0.9 (∼3.5 days)
IAR2	Yes	Yes	40%	50%	0.975 (∼14 days)

Table 1: List of experiments

with the persistence factors β_s for each sector s. This would allow shorter persistence for fire emissions and longer persistence for anthropogenic emissions. Which would make the constraint on emissions more flexible and realistic, but also account for the accuracy of the temporal profiles in the emissions prior.

3 Results

3.1 Experiments description

To show the impact of the presented methodology we have carried out the several EDA experiments with 25 members. The Lin-CO scheme is used so only CO fields are modelized and only CO observations are used (IASI and MOPITT retrievals, see Inness et al. (2015),Inness et al. (2019a) for details). The experiments are :

- Control Run (CR): where all CO observations are not assimilated but surface fluxes are perturbed
- Assimilation Run (AR): where CO observations are assimilated
- Inversion & Assimilation Run 1 (IAR1): where CO observations are assimilated and CO surface fluxes are corrected. Initial prior errors on surface fluxes are similar to the control and assimilation run.
- Inversion & Assimilation Run 2 (IAR2): same as IAR1 but initial prior errors on surface fluxes are doubled and with a higher persistence factor β.

Each experiments started on May 1^{st} 2017 and continued until the end of August of the same year.

3.2 Emission optimisation/inversion results

Figure 5 shows the impact of the two different inversion experiments relative to the CR on the ensemble emission totals (from fires and inventories) mean and spread averaged over August 2017. For the ensemble mean both experiment are showing the same structure of positive and negative corrections. Increases of emissions over the European Mediterranean basin and black sea, China's Hebei province and surroundings, middle east and western US. Decreases of emissions over India and surroundings, Eastern US, African Lesotho and surroundings and Northern Europe. The ensemble spread sees different adjustment depending on whether the emission prior errors are doubled or the persistence is increased. The IAR1 experiment shows small positive and negative adjustments on the spread whereas the IAR2

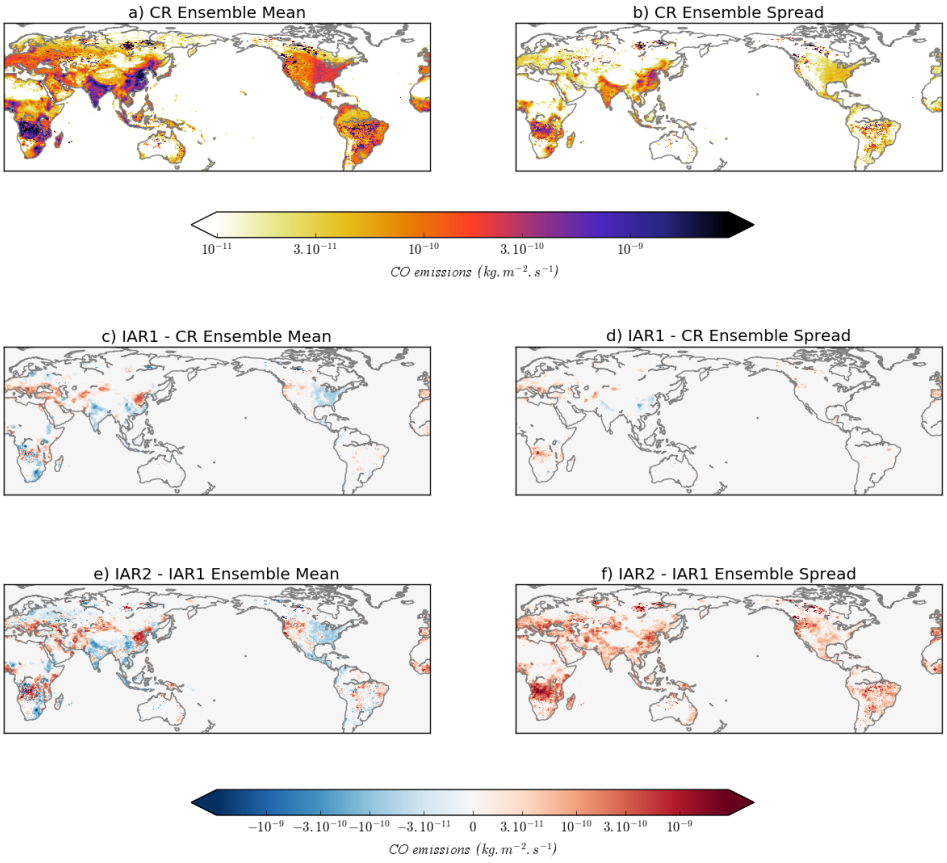

Figure 5: Control Run ensemble mean emissions a) and spread b). From c) to f), differences in IAR1 and IAR2 on CO surface emissions relative to the Control Run (see table 1) for August 2017.

experiment shows a strong increase due to the increased a-priori perturbation amplitudes specified (see table 1).

Figures 6 and 7 show the the impact on CO concentrations of each experiments relative to the CR at 1000hPa and 700hPa respectively. The difference AR-CR shows the impact of the assimilation of CO only on the ensemble 3D fields mean and spread. The assimilation correction patterns show what has been already shown in Inness et al. (2015), a general increase of CO in the extra tropics and near the main emission regions and a decrease of CO in the remote (i.e. far from the sources) regions. A slight decrease in the spread is to be noted in the AR-CR difference, close to the strong CO sources: northern India, China and also central and southern African regions. This suggests that the ensemble spread in the CR resulting from emission and NWP perturbations is somehow overestimated. Conversely, an increase of the spread would suggest an underestimation of the original spread seen in the CR.

The differences IAR1-CR and IAR2-CR show the impact of the dual inversion-assimilation with lower or higher prior emission spread and shorter or longer persistence respectively. At 1000hPa the corrections on the ensemble mean operated by the AR are enhanced by the IAR1 and IAR2, but larger in IAR2. The

corrections near the surface when inversion is additionally performed are enhanced because the inversions keep persistence of the systematic corrections that the assimilation is applying above potentially emission biased areas. Changes in the ensemble spread follow a similar behaviour for IAR1, but with significant spread increases around Middle East and around the Mediterranean basin, potentially indicating that the a-priori emission spread specified in IAR1 was too weak for those regions. For IAR2, the spread cannot be directly compared as the imposed prior spread from perturbation is larger, resulting in an overall spread increase. However some regions are showing a spread decrease due to the inversion and assimilation. At 700hPa enhancements are less obvious as the assimilation of the 3D fields dominate. Very little to no differences could be noticed on the ensemble mean changes between AR, IAR1 and IRA2. Only IAR2 shows some spread increase, most noticeable for example around the British Columbia area where strong wild-fires occur with injection and transport of such fire emissions up to the middle troposphere.

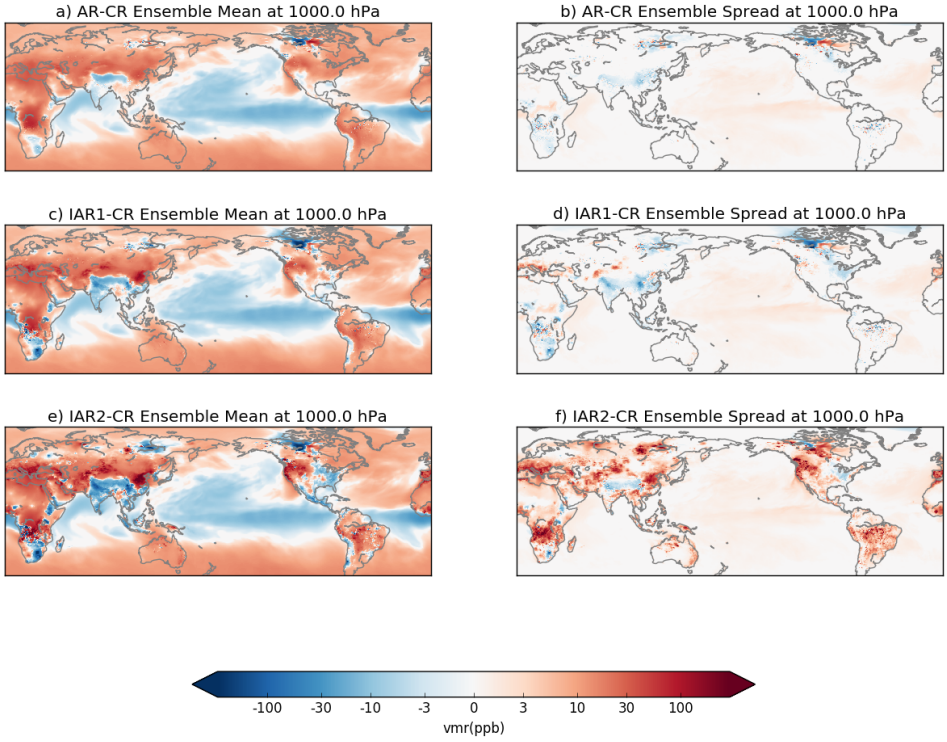

Figure 6: Differences in AR, IAR1 and IAR2 on CO fields at 1000hPa relative to the Control Run (see table 1).

3.3 Evaluation against Airbase and Airnow surface measurement networks

In this section we have performed evaluation against surface sites where emission changes should be reflected the most. We have used the European Airbase and US Airnow networks over the 5-31 August 2017 period (we have discarded the 5 first days of August due to a failure in the IAR2 runs, since the EDA is very costly to run a 26 days period will suffice to make our point in this first validation). We have selected the appropriate site representativeness (suburban and rural) to the model forecast resolution T511 ($\sim$40km).

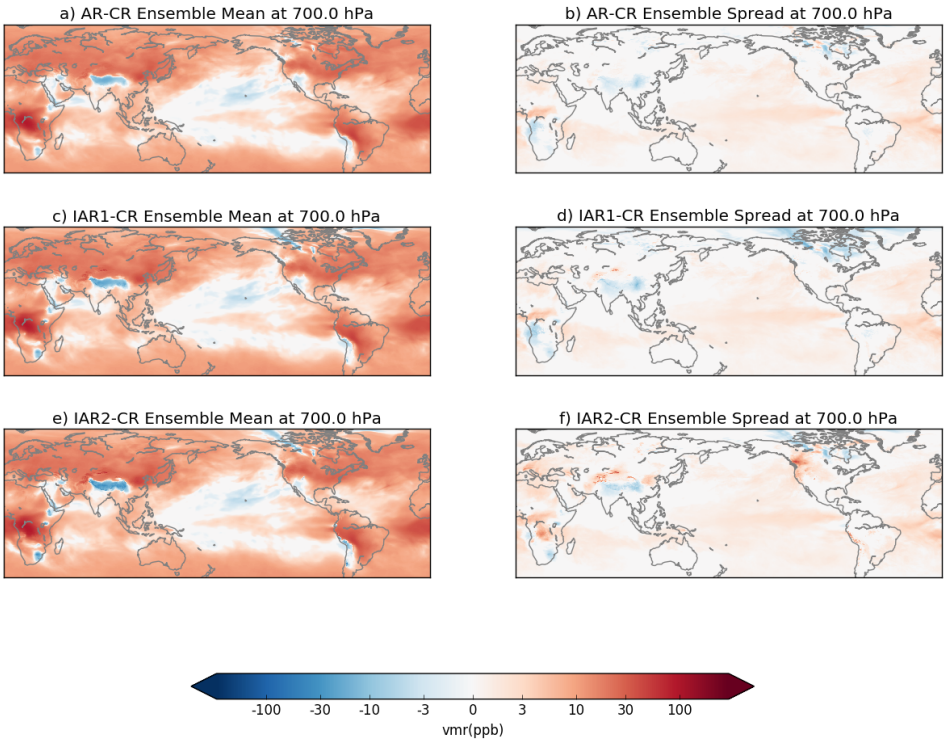

Figure 7: Differences in AR, IAR1 and IAR2 on CO fields at 700hPa relative to the Control Run (see table 1).

Figure 8 shows the evaluation with the overall bias and root mean square error (RMSE) for all the selected stations over Europe and the US for the short forecasts at T+0 (analysis time) and T+6. The strongest effect is seen in IAR2 as expected from section 3.2, with important increases in the overall CO concentration that lead to an apparent bias improvement but at the cost of strongly degrading the RMSE scores. No noticeable RMSE degradation is seen in IAR1, although the bias is significantly improved. Stronger effects are seen over Europe than over the US, which reflect the effects on surface or near surface concentrations in section 3.2 and figure 5. Inversion experiments show systematic increase of emissions and near surface concentrations over Europe, whereas the US sees milder positive and negative corrections.

Finally, figure 9 shows the evaluation with the overall bias and RMSE for all the stations over Europe and US at forecasts lead times ranging from T+0 to T+120. The same behaviour of IAR2 show an overall improvement of the bias but at the cost of a strong degradation of the RMSE. The IAR1 shows bias improvement over Europe only with no degradation of the RMSE in both Europe and the US. As mentioned above changes in emission and concentrations seen over the US are much smaller than changes over the Europe and this reflects in the evaluation.

The current evaluation shows then that the methodology proposed is sensitive to the a priori error inputed into the system and also to the persistence of the information carried forward. Further testing will be needed to assess whether the persistence or the a-priori perturbation on the emissions need to be adjusted. The example of the IAR2 experiment, designed with larger emission spread and longer persistence, shows that such a configuration is more likely to propagate errors on the surface and hence over time.

It also shows that the corrections enforced on the emissions are helping to improve the overall bias on the surface concentrations of pollutants but do not improve the RMSE score. This could be explained by the fact that the optimisation is done at the lower resolution of T159 than the actual forecast resolution of T511. Also due to the polar-orbiting geometry of the instruments used (MOPITT and IASI) with a maximum of two overpasses a day, a constraint on the emissions diurnal cycle is not possible. Lastly, the update on emissions using IFS cannot be done at a higher frequency than the current assimilation window length (12 hours) making it even more complicated to constrain the emissions diurnal cycle.

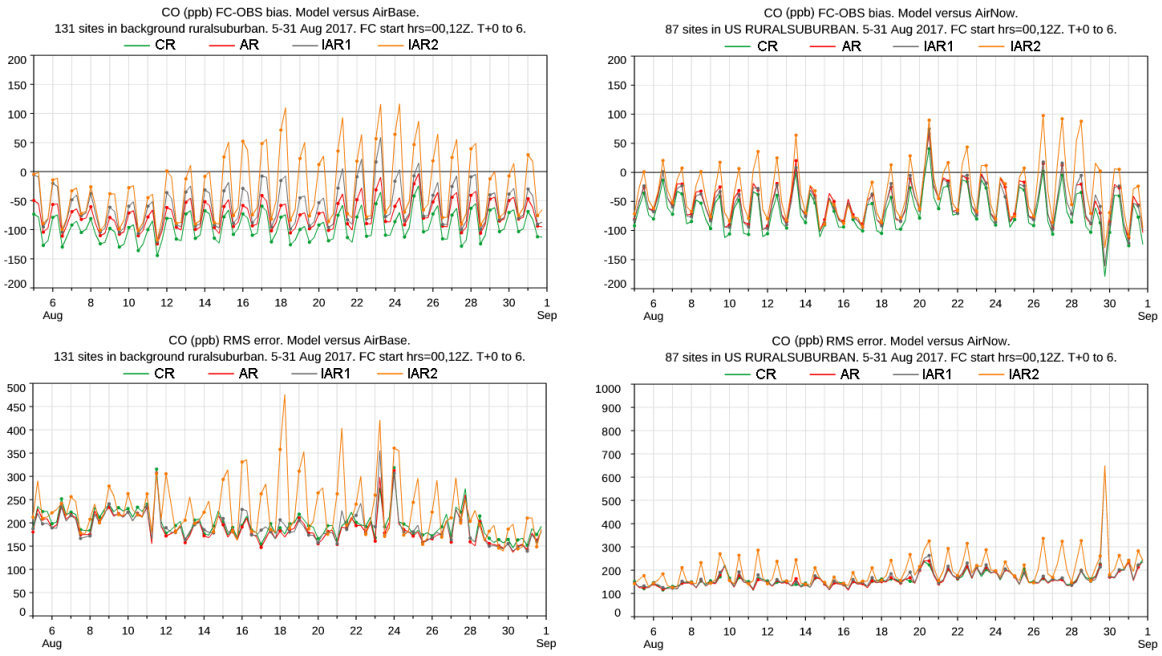

Figure 8: Evaluation against the Airbase european (left panels) and the Airnow US (right panels) surface CO measurement networks. Bias (top panels) and and RMSE (bottom panels) are computed using the rural and suburban classification sites only (i.e. urban excluded).

4 Discussions and further works

In this paper we have designed a first methodology that uses the ensemble information to constrain the surface sources of atmospheric pollutants using the Jacobians derived from the IFS EDA. Focusing on CO, which is well observed in the troposphere from space and which is also a primary pollutant emitted mainly from fires and anthropogenic pollution, we have shown the potential of the methodology. The impact of dual assimilation-inversion should be even larger with shorter lifetime species such as NO_2, where its variability mainly controlled by local sources. Developments are currently carried out within IFS towards NO_2 inversions.

The methodology presented here offers the advantage of an easy implementation (only modifying and adding few IFS scripts) and offer the possibility to relate the increment on a given parameter to another parameter using the statistical information that the EDA provides. In our case we have applied this methodology to relate atmospheric pollutant increments to surface fluxes. Such a statistical approach

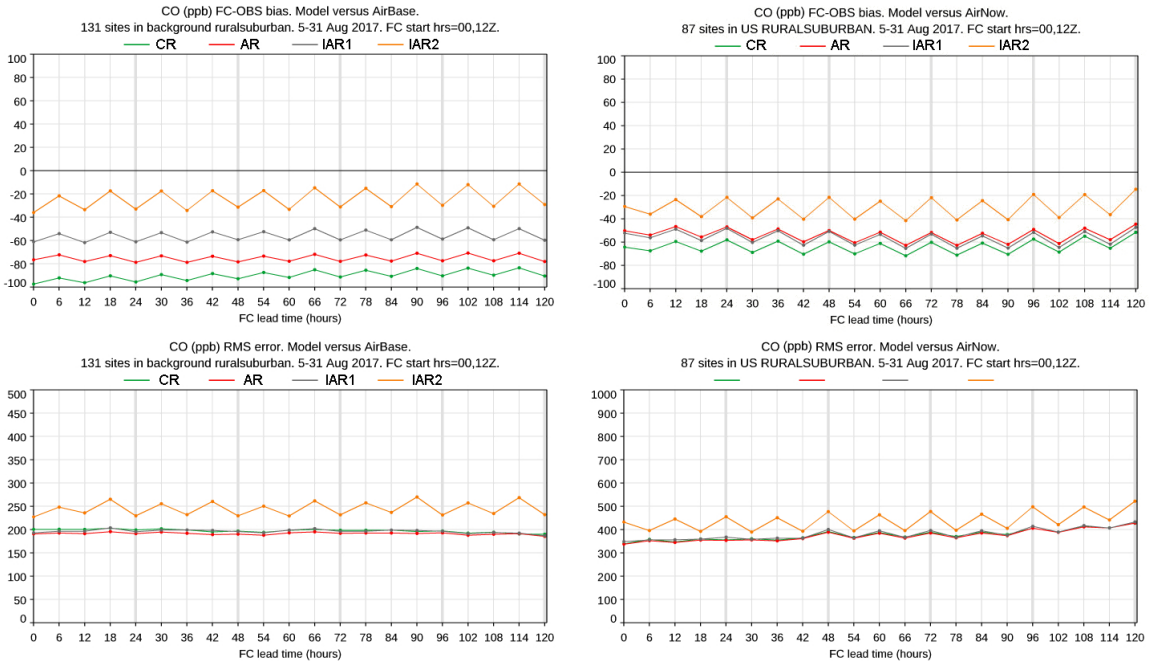

Figure 9: Evaluation against the Airbase european (left panels) and the Airnow US (right panels) surface CO measurement networks at forecast lead times between T+0 and T+120 hours. Bias (top panels) and and RMSE (bottom panels) are computed using the rural and suburban classification sites only (i.e. urban excluded).

using the ensemble information could also be used to relate co-variates. One example would be to relate observed atmospheric composition tracers and NWP variables (e.g. winds). Of more importance for atmospheric composition would be the ability to link poorly or unobserved pollutants (i.e. volatile organic compounds and particulate matter) or even use the co-emitted relationship between species to help informing on the very challenging surface fluxes such as CO_2.

In this paper we have showed that for this methodology to be optimal, it would require very careful tuning of the prior error distribution of the surface emission perturbation inputed into the EDA. Given the results of sections 3.2 and 3.3, such error should be varying geographically, in time but also between different sectors. Inversion and assimilation results point out that certain regions are lacking in spread where others are showing a clear over spread. This poses a challenge for implementation within the IFS as the system does not have currently an inflation capability on the surface fluxes. It is also a challenge in the general inversion community as current state-of-the-art inversion frameworks do not fully integrate a model evolution for fire and anthropogenic emissions. Having models for anthropogenic and fire emissions will allow to represent an evolving emission prior error distribution at each assimilation window.

Finally, such method do not pretend to replace the implementation of a tangent linear and adjoint (TL/AD) to retrieve a sensitivity between observation departures and surface fluxes. This TL/AD work is also currently carried forward and should be assessed jointly with this method. The TL/AD would offer less flexibility within the first layer of implementation (such as inference of cross-correlated or co-variate components). However, the TL/AD would be computationally cheaper (than running an EDA) and directly fit the current CAMS IFS configuration, which is not running in EDA mode but only as a single 4D-Var realisation (within an outer-loop). Optimally the benefit of both TL/AD and EDA based approach should be compared and combined if CAMS can afford to operate an EDA in the future. Within the cur-

rent simplifications of the CAMS system for fast reactive chemistry in the TL/AD, great care should be made when using increments variationally derived over a 12-hourly assimilation window to impact emissions. The same considerations should be accounted when implementing solely the TL/AD terms that permit sensitivity to the surface fluxes (e.g. diffusion and deposition). In this study, the satistically derived Jacobians using an ensemble of forecasts with the full physics and chemistry schemes implicitly accounts for emission processes implicitly, i.e. diffusion, deposition, chemistry and transport. The challenge remains on how to insert optimally the ensemble information into the IFS. This area of research should be carefully considered for optimal emission or surface fluxes inversions in the future.

References

Bonavita, M., Isaksen, L., and Hlm, E. (2012). On the use of eda background error variances in the ecmwf 4d-var. *Quarterly Journal of the Royal Meteorological Society*, 138(667):1540–1559.

Claeyman, M., Attié, J.-L., El Amraoui, L., Cariolle, D., Peuch, V.-H., Teyssèdre, H., Josse, B., Ricaud, P., Massart, S., Piacentini, A., Cammas, J.-P., Livesey, N. J., Pumphrey, H. C., and Edwards, D. P. (2010). A linear co chemistry parameterization in a chemistry-transport model: evaluation and application to data assimilation. *Atmospheric Chemistry and Physics*, 10(13):6097–6115.

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., and Suttie, M. (2019a). The cams reanalysis of atmospheric composition. *Atmospheric Chemistry and Physics*, 19(6):3515–3556.

Inness, A., Alben, I., Agusti-Panareda, A., Borsdoff, T., Flemming, J., Landgraf, J., and Ribas, R. (2019b). Monitoring and assimilation of early tropomi total column carbon monoxide data in the cams system. *ECMWF Technical Memoranda*, (838).

Inness, A., Blechschmidt, A.-M., Bouarar, I., Chabrillat, S., Crepulja, M., Engelen, R. J., Eskes, H., Flemming, J., Gaudel, A., Hendrick, F., Huijnen, V., Jones, L., Kapsomenakis, J., Katragkou, E., Keppens, A., Langerock, B., de Mazière, M., Melas, D., Parrington, M., Peuch, V. H., Razinger, M., Richter, A., Schultz, M. G., Suttie, M., Thouret, V., Vrekoussis, M., Wagner, A., and Zerefos, C. (2015). Data assimilation of satellite-retrieved ozone, carbon monoxide and nitrogen dioxide with ecmwf's composition-ifs. *Atmospheric Chemistry and Physics*, 15(9):5275–5303.

Kaiser, J. W., Heil, A., Andreae, M. O., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.-J., Razinger, M., Schultz, M. G., Suttie, M., and van der Werf, G. R. (2012). Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power. *Biogeosciences*, 9(1):527–554.

Massart, S., Agusti-Panareda, A., Aben, I., Butz, A., Chevallier, F., Crevoisier, C., Engelen, R., Frankenberg, C., and Hasekamp, O. (2014). Assimilation of atmospheric methane products into the macc-ii system: from sciamachy to tanso and iasi. *Atmospheric Chemistry and Physics*, 14(12):6139–6158.

Massart, S., Agustí-Panareda, A., Heymann, J., Buchwitz, M., Chevallier, F., Reuter, M., Hilker, M., Burrows, J. P., Deutscher, N. M., Feist, D. G., Hase, F., Sussmann, R., Desmet, F., Dubey, M. K., Griffith, D. W. T., Kivi, R., Petri, C., Schneider, M., and Velazco, V. A. (2016). Ability of the 4-d-var analysis of the gosat besd xco_2 retrievals to characterize atmospheric co_2 at large and synoptic scales. *Atmospheric Chemistry and Physics*, 16(3):1653–1671.

Massart, S. and Bonavita, M. (2016). Ensemble of data assimilations applied to the cams' greenhouse gases analysis. *ECMWF Technical Memoranda*, (780).