TECHNICAL MEMORANDUM

780

Ensemble of Data Assimilations applied to the CAMS' greenhouse gases analysis

Sébastien Massart
and Massimo Bonavita

Copernicus Department
& Research Department

June 2016

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 Ensemble of Data Assimilations (EDA) is currently the method used at ECMWF to estimate the background error statistics for the meteorological analysis. It particularly allows to have flow-dependent background errors which is beneficial for ECMWF's operational deterministic meteorological analysis.

The Copernicus Atmosphere Monitoring Service (CAMS) provides an atmospheric tracers analysis based on Composition-IFS (C-IFS). Distinctively from ECMWF's operational deterministic meteorological analysis, CAMS analysis uses climatological estimates of background error. We want to investigate whether using EDA-based flow-dependent background errors for atmospheric tracers could be beneficial for the CAMS analysis as it is for ECMWF's operational analysis.

The first step before using EDA-based background errors in the CAMS analysis is to design the EDA for atmospheric tracers. This document describes the Composition-EDA (C-EDA), a modified version of ECMWF's EDA that accounts for the sources of uncertainties associated with atmospheric tracers. The focus is on two species of the anthropogenic long-lived greenhouse gas family: carbon dioxide (CO_2) and methane (CH_4).

A default C-EDA experiment with only perturbations of the meteorological parameters showed that the spread of the ensemble is larger at the surface than in the troposphere for both CO_2 and CH_4. Moreover the ensemble standard deviation has a large variability in time at the surface probably associated with the time variation of the boundary layer. It also has strong spatial variations at the surface and in the troposphere probably associated with the spatial distribution of the surface fluxes. This provides further motivation for implementing fully flow-dependent background errors in the analysis of atmospheric tracers.

The spread of the ensemble is estimated to be too low when comparing to the ensemble mean background error. We demonstrate the adding directly a perturbation on the surface fluxes helps to increase the ensemble spread. Nevertheless, it is found that the amplitude of the surface fluxes perturbation used in this document could be too large according to a diagnostic based on the ensemble mean background error. This emphasizes that the perturbation of the surface fluxes has to be implemented in the C-EDA experiments, but this requires further developments.

1 Introduction

To produce an analysis of the meteorological situation, the European Centre for Medium-Range Weather Forecasts (ECMWF) relies on its forecast model and observations combined together by a 4D-Var assimilation system. The assimilation process balances the weight given to the observations and that given to the background (short-term forecast). The background weight is determined by the background error statistics which are currently estimated using an Ensemble of Data Assimilations (EDA, Bonavita *et al.*, 2012) in the ECMWF's Integrated Forecasting System (IFS).

The EDA is designed to represent the main sources of background error. It is composed of an ensemble of low resolution 4D-Var analysis cycles. Each member of the ensemble makes use of perturbed observations, perturbed sea-surface temperature (SST) fields and perturbed model physical tendencies. The EDA is running in parallel to the operational deterministic 4D-Var analysis to represent the evolution of its errors. Estimating the evolution of the background errors is one strength of the EDA as it has been demonstrated that using flow-dependent background errors is beneficial for ECMWF's operational deterministic meteorological analysis (Bonavita *et al.*, 2015).

As part of the previous EU-funded Global and regional Earth-system (Atmosphere) Monitoring using Satellite and in-situ data (GEMS) and Monitoring Atmospheric Composition and Climate research and development projects (MACC to MACC-III) and of the current EU-funded Copernicus Atmosphere Monitoring Service (CAMS), ECMWF has devoted significant efforts to integrating a detailed representation of atmospheric composition and associated processes into IFS. The resulting system, Composition-IFS (C-IFS, Peuch *et al.*, 2015) uses the same 4D-Var system as the operational ECMWF meteorological analysis to produce analyses of atmospheric tracers (aerosols, reactive gases and greenhouse gases). The CAMS analysis currently uses static climatological background error statistics and not statistics derived from an EDA.

We want to investigate whether using EDA-based flow-dependent background errors for atmospheric tracers would also be beneficial for the CAMS analysis. Weather conditions can have a strong influence on atmospheric composition. Accounting for the meteorological situation in the analysis of atmospheric tracers through the background error statistics should lead to improved quality of the analysis. The first step towards using EDA-based background errors in the CAMS analysis is to design the EDA for atmospheric tracers. This document describes some modifications brought to the current ECMWF's EDA in order to account for the sources of uncertainties associated with atmospheric tracers. For instance, the surface fluxes and the chemical reactions are some of the error sources of background errors.

The atmospheric tracers in C-IFS are grouped into three categories: aerosols, reactive gases and greenhouse gases. In this report we focus on two species of the anthropogenic long-lived greenhouse gas family: carbon dioxide (CO_2) and methane (CH_4). The greenhouse gas module is less computationally expensive to run in C-IFS than the aerosols module or the reactive gases module. This makes CO_2 and CH_4 suitable candidates to test different EDA configurations. These two gases are relatively well-mixed in the atmosphere and are strongly related to surfaces fluxes. The atmospheric concentration of CH_4 also has a chemical dependence.

This report details in Sec. 2 the modifications and extensions of the cycle CY41R1 EDA we implemented to account for the atmospheric tracers' sources of uncertainties. Section 3 presents the configuration of the EDA experiments we ran. Section. 4 presents the results of these experiments. Finally the last section (Sec. 5) summarizes the results and brings some information on future developments required to further improve the atmospheric composition EDA.

2 Practical implementation of the EDA for atmospheric composition

The EDA is composed of a given number M of members. Each member is a succession of analysis and short-range forecast cycles at a lower resolution than the operational deterministic analysis/forecast cycle. The spread of the ensemble should represent the uncertainty in the background (short-range forecast). For that purpose, each member of the ensemble assimilates perturbed observations, uses perturbed SST and perturbed model physical tendencies (Bonavita *et al.*, 2012).

To design the atmospheric composition EDA (or C-EDA) we kept the same perturbations as for the EDA plus we added the possibility to perturb the atmospheric tracer tendencies (Sec. 2.1) and to perturb the surface fluxes (Sec. 2.2). Due to the strong spatial variations in the distribution of surface fluxes, the CO_2 and CH_4 ensemble spread was found to be much noisier than for the meteorological parameter close to the surface. Section 2.3 presents the filtering of the ensemble spread. Section 2.4 presents the used calibration of the CO_2 and CH_4 ensemble spread based on its comparison with the ensemble mean background error.

2.1 Tendencies perturbations

The perturbation of the model physical tendencies relies on a noise generated with the Stochastically Perturbed Parametrization Tendencies method (SPPT, Palmer *et al.*, 2009). This method is used to generate a series of two-dimensional horizontal random fields with a zero mean and specified standard deviation, horizontal correlation length scale and time correlation length scale. We used the default configuration which consists of a horizontal correlation length scale of 500 km, a correlation time of 6 h and a standard deviation of 0.52 (Fig. 1). The resulting two-dimensional field is added to the model physical tendencies for each model time step and for all model levels but the lowest ones. The physical tendencies of the model levels located in the boundary layer are indeed not perturbed in order to avoid numerical instabilities.

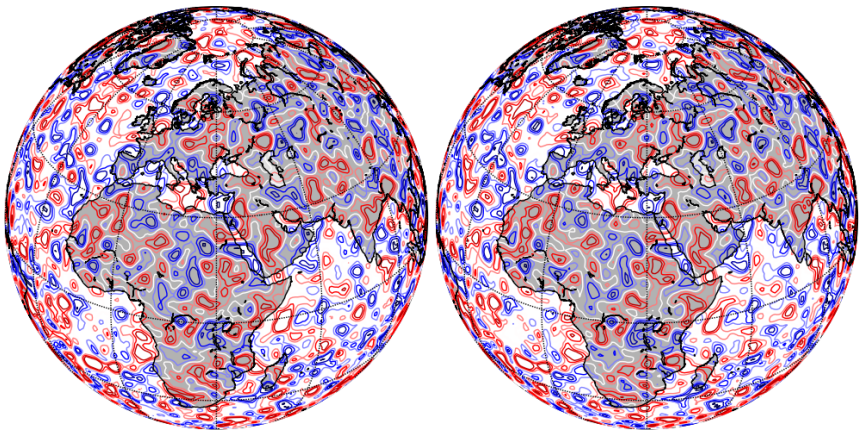


Figure 1: Example of the SPPT perturbation fields for two consecutive 3-hours steps; contour interval 0.25; red (blue) contours correspond to positive(negative) values.

The perturbation of model physical tendencies provides a different transport of the chemical tracers for each member. This accounts for the uncertainty of the transport. In a first configuration, we decided not to add any other perturbation of the chemical tracers tendencies. Nevertheless, the perturbation of the

model physical tendencies has an indirect impact on the CO_2 surface fluxes as they are coupled with the meteorological parameters. As a consequence, in this first configuration, CO_2 surface fluxes are already perturbed.

The chemical loss and production of the chemical tracers are other sources of uncertainty. To account for them, we decided to add the possibility to perturb the chemical tracers tendencies the same way as for the model physical tendencies. The only difference is that we apply the perturbation for all model levels.

In C-IFS, surface fluxes are transformed into chemical tracers tendencies in the boundary layer through turbulent diffusion. Perturbing the tracers tendencies in the boundary layer adds indirect perturbation on the surface fluxes. As there is no chemistry in C-IFS for CO_2, the perturbation of CO_2 tendencies implies an added indirect perturbation on the CO_2 surface fluxes. The perturbation of CH_4 tendencies further adds a perturbation of the CH_4 chemistry.

2.2 Surface fluxes perturbations

Surface fluxes are among the main drivers of the greenhouse gases concentration in C-IFS. Ocean and anthropogenic CO_2 surface fluxes are prescribed in C-IFS using inventories. Fire emissions are derived from CAMS GFAS product. Terrestrial biogenic carbon fluxes are modelled by the carbon module of the land surface model CTESSEL (Boussetta *et al.*, 2013). A more comprehensive description of the CO_2 surface fluxes can be found in Agustí-Panareda *et al.* (2014). Methane fluxes are prescribed in C-IFS using inventory and climatological data sets. A more comprehensive description of the CH_4 surface fluxes can be found in Massart *et al.* (2014).

Surface fluxes are associated with large uncertainties in terms of amplitude and sometime in terms of distribution too. We therefore introduced the capability to perturb the surface fluxes of each member to account for their uncertainty. We selected two different approaches to perturb the surface fluxes for CO_2 and CH_4. For CO_2, we decided to perturb the scaling factors of the biogenic flux adjustment scheme (BFAS, Agustí-Panareda *et al.*, 2016). For CH_4, we decided to modify the surface fluxes using two-dimensional perturbed fields.

2.2.1 Building CO_2 perturbed surface fluxes

The biogenic CO_2 surface fluxes, associated with organic soil and vegetation, are the CO_2 surface fluxes with the highest uncertainties on the global scale. We decided to perturb only this component of the CO_2 surface fluxes as a first step. In practice, we perturbed the vegetation gross primary production (GPP) and the heterotrophic respiration (R_{eco}) as computed by CTESSEL. We then used the Net Ecosystem Exchange (NEE) as biogenic surface fluxes, which is the difference between GPP and R_{eco}.

The computed values of GPP and R_{eco} depend on the values of some of the meteorological parameters. As a consequence, the values of GPP and R_{eco} are different for each member due to the perturbation of the physical tendencies in the EDA. Preliminary results showed that the spread of the resulting NEE between the members was too weak compared to the expected spread estimated with the ensemble mean background error (see section 2.4). To increase the spread, we decided to perturb even more directly the NEE. We based our perturbation strategy on the recent developments of BFAS by which GPP and R_{eco} rescaling factor are computed for various regions and various vegetation types. If $\gamma^G(v,t)$ is the GPP rescaling factor for the vegetation type v and the time t as computed by BFAS, we multiplied this rescaling factor by $1+\alpha\,\eta_m^G(v,t)$ where α is the amplitude of the perturbation and $\eta_m^G(v,t)$ a random number following a normal distribution. The perturbed GPP rescaling factor $\tilde{\gamma}_m^G(v,t)$ for the member m,

the vegetation type v and the time t is then

$$\tilde{\gamma}_m^G(v,t) = \gamma^G(v,t)\left[1. + \alpha\,\eta_m^G((v,t)\right] . \tag{1}$$

For a given vegetation type v, we furthermore imposed a time correlation σ_t between the random numbers $\eta_m^G(v,t)$. A similar perturbation is used for the R_{eco} rescaling factors computed by BFAS.

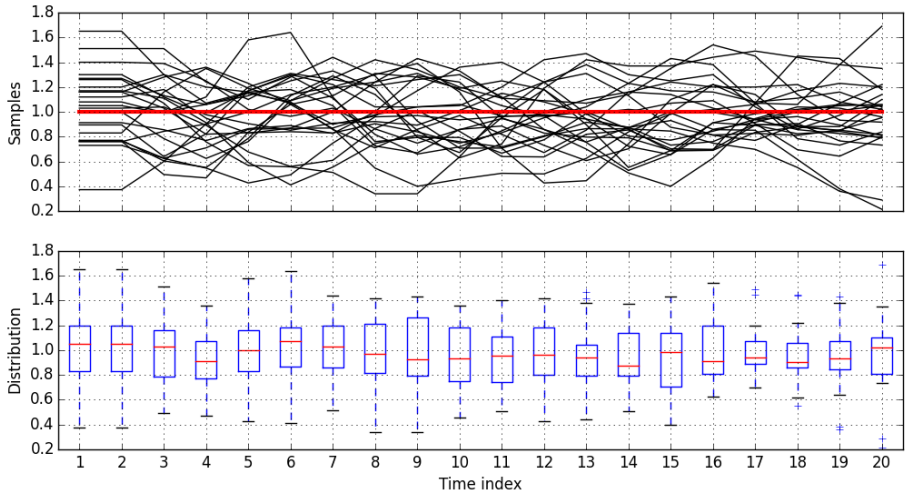


Figure 2: Example of time series of perturbation of the gross primary production rescaling factor for the crops over Europe. Top: perturbed rescaling factor for the 25 members (black lines) and reference rescaling factor (red line). Bottom: distribution of the perturbation over the 25 members.

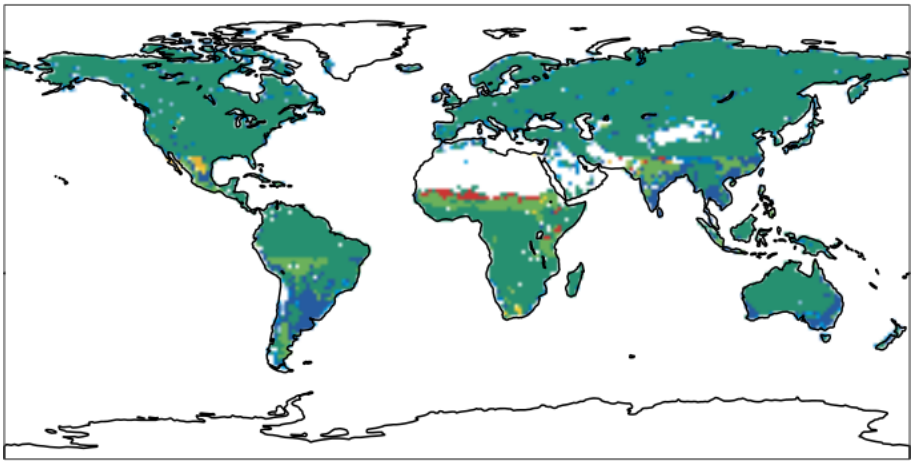
Figure 2 illustrates the time series of $\tilde{\gamma}_m^G(v,t)$ for 25 members and for the crops over Europe. In this example, the time decorrelation length is set to 3 time steps and the amplitude of the perturbation is $\alpha = 25\%$.

From the GPP and R_{eco} rescaling factors, BFAS produces re-scaling maps of the biogenic fluxes. Following this approach, we also produced rescaling maps for each member. Figure 3 illustrates the GPP re-scaling map obtained after the perturbation of the rescaling factors of each vegetation type when the amplitude of the perturbation is $\alpha = 25\%$.

The BFAS aims at reducing the large scale atmospheric CO_2 biases by adjusting the surface fluxes. Perturbing the parameters of BFAS results in adding large scale bias to the atmospheric CO_2 field. This also changes the CO_2 budget of each member without ensuring that the mean ensemble budget is equal to the budget of the control. These issues will have to be addressed in the future, but for this first try of perturbing the parameters of the CO_2 surface flux, we assume that our strategy is satisfactory enough.

The rescaling factors from BFAS are constant during the short-term forecast in C-IFS as they are meant to correct the 10-day regional budget. Similarly, we kept the perturbed rescaling factors constant during the short-term forecast.

(a) Reference



(b) First member

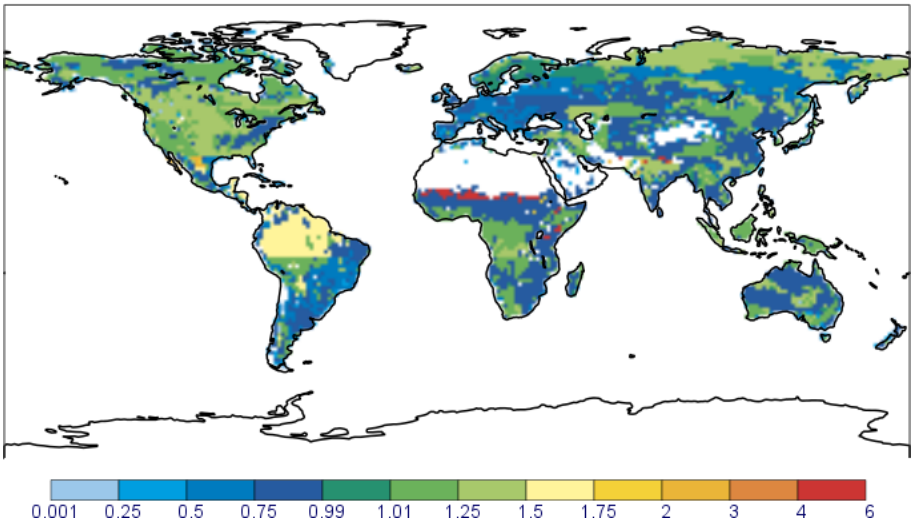


Figure 3: Example of gross primary production rescaling factor maps: (a) reference map (without perturbation) ; (b) map for the first member after perturbation of the gross primary production rescaling factors for each vegetation type.

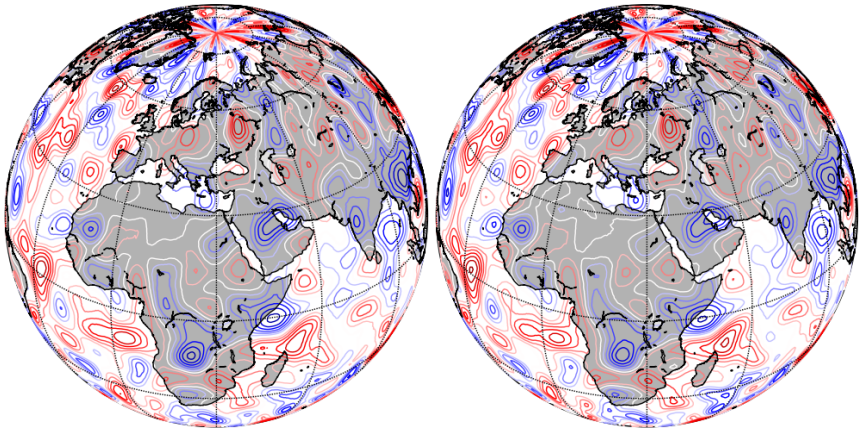
2.2.2 *Building* CH_4 *perturbed surface fluxes*

The methane surface flux $\Phi(\boldsymbol{\xi},t)$ is a two dimensional field with $\boldsymbol{\xi}$ the spatial coordinate and t the time coordinate. To perturb CH_4 surface fluxes of each member m, we modified this reference surface fluxes $\Phi(\boldsymbol{\xi},t)$ using a two dimensional random field $\eta_m(\boldsymbol{\xi},t)$ with a zero mean, a standard deviation of one, a spatial correlation length r_x and a temporal correlation length r_t. The surface flux $\Phi_m(\boldsymbol{\xi},t)$ of each member m is computed with

$$\Phi_m(\boldsymbol{\xi},t) = \Phi(\boldsymbol{\xi},t)\,(1. + \alpha\,\eta_m(\boldsymbol{\xi},t))\,, \tag{2}$$

where α is a parameter to adjust the amplitude of the perturbation.

(a) Two successive perturbation fields; contour interval 0.5; red (blue) contours correspond to positive(negative) values



(b) Perturbation time serie

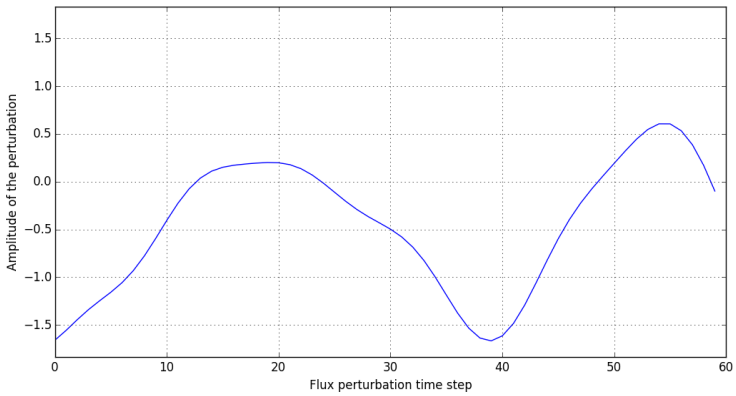


Figure 4: Example of perturbation of the fluxes: (a) two-dimensional field of the perturbation η for two successive times and a given member ; (b) time series of the perturbation η in a given grid point for a given member.

To generate the random fields $\eta_m(\boldsymbol{\xi},t)$, we had two approaches. The first approach consists at using the two-dimensional random field generated by the SPPT method and used to perturb the model physical tendencies. The advantage of this approach is that the implementation is easy as the random field is already available in the code. The weakness is that we have to use the same horizontal correlation length scale and time correlation length scale as for the perturbation of the model physical tendencies.

To allow more flexibility we generated n_t random fields on a regular $nlat \times nlon$ grid with a zero mean, a standard deviation of one and a time correlation length of σ_t as a second approach. The introduction of spatial correlation is based on the convolution of each field with a Gaussian kernel, generated with the Gaussian window

$$w(k) = e^{-\frac{1}{2}\left(\frac{k}{\sigma_w}\right)^2}, \tag{3}$$

with k the number of points in the output window and σ_w the standard deviation. We furthermore imposed the constrains $k = 5\sigma_w$, so that the number of points is large enough compared to the standard deviation but it not too large in order to save computational time. Figure 4a illustrates the field $\eta_m(\boldsymbol{\xi}, t)$ obtained for a 401×800 grid with $\sigma_w = 10$ (grid point). Figure 4b illustrates the time variation of $\eta_m(\boldsymbol{\xi}, t_i)$ at a given point when $\sigma_t = 6$ (time step).

One main difference between the two-dimensional field generated by the SPPT method and the one generated by the above method is the specification of the horizontal correlation length scale. By construction, the length scale is specified in distance (kilometres for instance) for the SPPT method, while the length scale is specified in in number of grid points for the the second method. The construction of the perturbation field could nevertheless be changed in the future to account for different correlation length scales if needed. In addition, a more elaborated perturbation field could have the correlation length scales accounting for the surface type (land or sea), the surface vegetation or the emission category (anthropogenic, biogenic, ...).

Currently, CH_4 fluxes are constant in C-IFS during the forecast as for CO_2. For practical reasons, we kept this behaviour for the perturbation. The perturbation nevertheless changes from one forecast to the next with a time correlation of σ_t. We generated the perturbations on a regular 401×800 grid and we interpolated the perturbations towards the resolution of the model.

2.3 Filtering the background error variances

With 25 members, the sample size is not large enough to expect an accurate error estimate. Ad hoc filtering is required on the raw ensemble estimate as highlighted by Bonavita *et al.* (2012). A filter based on two 50-member EDA that were run for 45 days is used for the meteorological parameters. Being at a stage where we are designing the C-EDA configuration, we do not have yet a selected configuration from which we could run similar C-EDA experiments to compute the filter for the atmospheric tracers. Instead, we tried to use the same filter as it is currently used for the U component of wind for convenience.

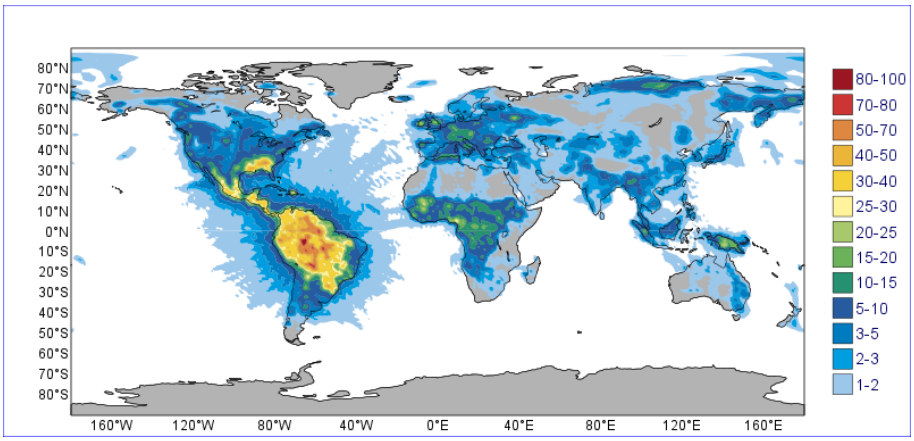
In C-IFS, CO_2 or CH_4 variance fields are grid-point fields. The raw ensemble estimate of the CO_2 or CH_4 variance fields presents sharp gradients at the surface. As a consequence, when the fields are transformed in the spectral space to apply the filter and then transformed back in the physical space, the Gibbs phenomenon appears as illustrated by Figs. 5a and 5b. We therefore decided to apply the following additional spectral smoother while projecting the raw ensemble estimate of the CO_2 or CH_4 variances in the spectral space:

$$\rho(n) = 1 + (\alpha - 1)\left[\frac{n(n+1)}{N_{trunc}(N_{trunc}+1)}\right]^{\beta} \tag{4}$$

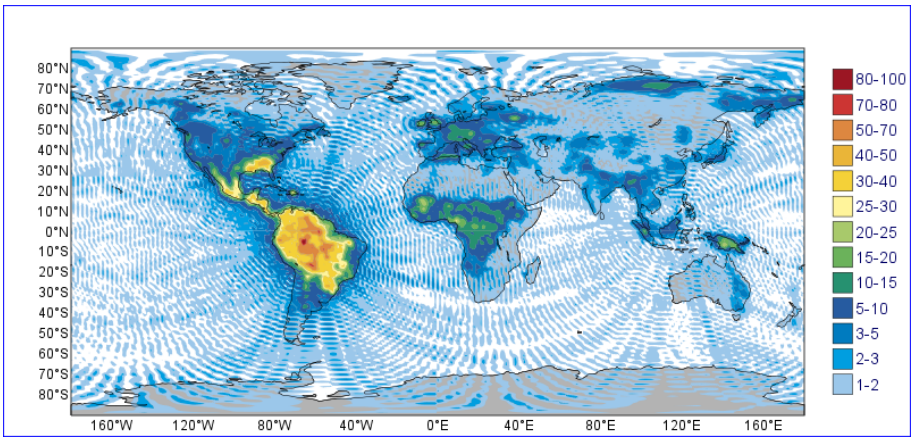
where n is the spectral wave number, N_{trunc} is the truncation wave number, α is the reduction factor and β the exponent for the spectral smoother. After several tests, we decided to use the following values for this smoother: $\alpha = 100$ and $\beta = 2$, as the noise was removed using these values (Fig. 5c).

The effect of the supplementary smoother and the filter on the power spectra of the CO_2 and CH_4 variance is illustrated Fig. 6. The filtering using the filter for the U component of wind reduces the energy in the

(a) Standard deviation from the ensemble



(b) Same as (a) but after after additional filtering



(c) Same as (b) but smoothing

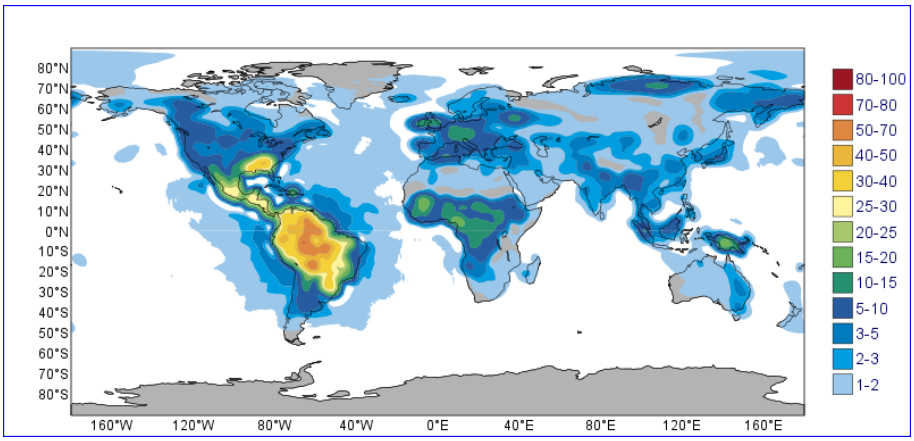
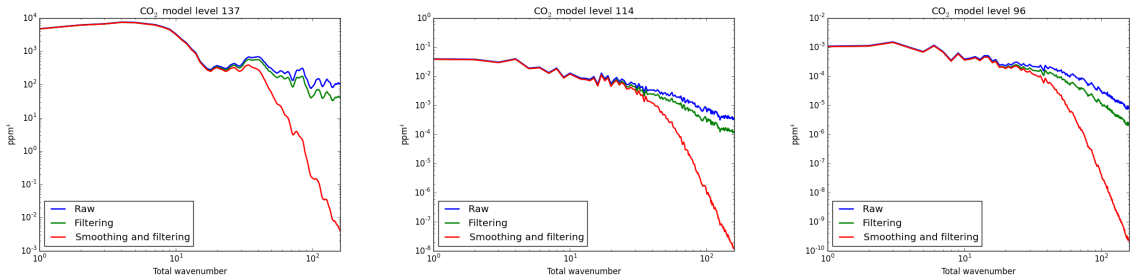


Figure 5: CO_2 *standard deviation at the surface (in* ppm*) from the EDA on 1 October, 2014 at 9 UTC.*

(a) CO_2 *Level 137 (surface)* *(b)* CO_2 *Level 114 (≈ 850* hPa*)* *(c)* CO_2 *Level 96 (≈ 500* hPa*)*



(d) CH_4 *Level 137 (surface)* *(e)* CH_4*Level 114 (≈ 850* hPa*)* *(f)* CH_4*Level 96 (≈ 500* hPa*)*

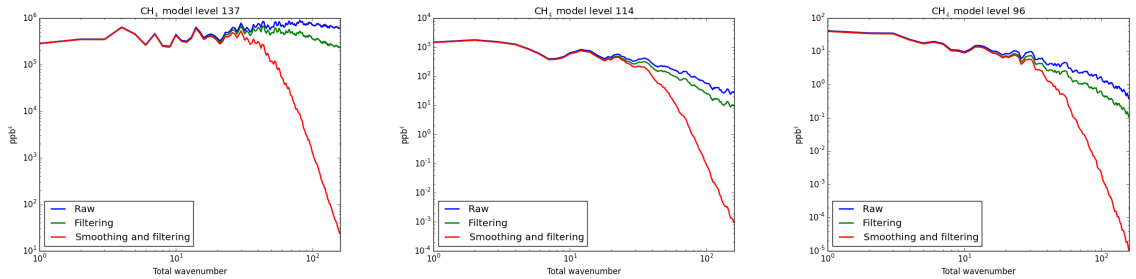


Figure 6: Power spectra (in ppm^4*) of the ensemble short-range (t+9h)* CO_2 *forecast variance (top panels) and* CH_4 *forecast variance (bottom panels) on 1 October, 2014 for 3 model levels. The blue line is the estimation from 25 members. The green line is the estimation after smoothing. The red line is the estimation after smoothing and filtering.*

smaller scales, but not to a large extent. This indicates that our hypothesis that the dominant source of atmospheric tracers variability comes from meteorological variability is not quite right and atmospheric tracers variability is much higher than the variability of the U component of wind. The smoother reduces much further the energy in the smaller scales, starting from the wavenumber $n = 30$ (for a truncation $N_{trunc} = 159$).

2.4 Background error calibration

It is well-known that the spread in the EDA members is not large enough compared to the expected background error (Bonavita *et al.*, 2012). To deal with this issue, the approach of Bonavita *et al.* (2012) consists of calibrating the EDA spread to enforce statistical consistency between the EDA variance estimates and the diagnosed ensemble mean background error. We followed this approach for the C-EDA where the spread Σ is computed using M CO_2 or CH_4 model fields $\boldsymbol{x}_m$ from 3-hr short-range forecasts starting from the M previous analyses,

$$\Sigma = \frac{1}{M-1} \sum_{m=1}^{M} (\boldsymbol{x}_m - \overline{\boldsymbol{x}})^2 \,, \tag{5}$$

where $\bar{\boldsymbol{x}}$ is the average of $\boldsymbol{x}_m$ over the M members. The ensemble mean background error Ω is diagnosed with

$$\Omega = \frac{M}{M+1}\left(\frac{1}{M}\sum_{m=1}^{M}\boldsymbol{x}_m - \boldsymbol{y}\right)^2, \tag{6}$$

where $\boldsymbol{y}$ should be the true state. In operations, the high resolution deterministic forecast is used as a proxy for the true state $\boldsymbol{y}$ of the meteorological fields. Here, we do not have a high resolution deterministic forecast. We could use the control (member of the C-EDA without perturbation), but the difference between the members and the control does not account for all the sources of background error. We therefore use the forecast from an experiment similar to the control. The only difference between them is the fact that the background error statistics of the meteorological analyses variables come from an EDA for the reference experiment used as $\boldsymbol{y}$, while the background error statistics are static in the C-EDA control. This helps to account for more sources of background error but it is probably not enough.

The C-EDA spread Σ is split into bins and latitude band (Northern Hemisphere, Tropics, Southern Hemisphere) for each variable. For each subset, the corresponding C-EDA spread Σ is rescaled to ensure the equality with the corresponding ensemble mean background error Ω over a period of several days (see Bonavita *et al.* (2012) for more details on the methodology).

3 Experiments configuration

We ran three C-EDA experiments in various configurations with CO_2 and CH_4 as part of the C-IFS atmospheric tracers (see details of the configurations below). The resolution of the experiments is TL255L137. Every experiment is run using 25 members and starts on 1 October, 2014 for a period of 20 days. During this period, we are assimilating column-averaged dry-air mole fractions of CO_2 (XCO_2) from the Greenhouse gases Observing Satellite (GOSAT). More details on the configuration of the assimilation of XCO_2 GOSAT data can be found in Massart *et al.* (2015). In brief, XCO_2 GOSAT data are sparse in space (about 50 datapoints per assimilation cycle), only available over land and during daylight, and sensitive to the lower troposphere. We also assimilate column-averaged dry-air mole fractions of CH_4 (XCH_4) from the Infrared Atmospheric Sounding Interferometer (IASI). More details on the configuration of the assimilation of XCH_4 IASI data can be found in Massart *et al.* (2014). In brief, XCH_4 IASI data are dense in space, over ocean and land and during day and night time, and sensitive to the middle troposphere.

The details of the three experiments we ran are summarized Table 1 and described below:

1. BASIC : experiment similar to the classic EDA, but with the addition of CO_2 and CH_4. Note that CO_2 surface fluxes are indirectly perturbed due to the perturbation of the model physical tendencies.

2. FLUX : experiment with perturbation of CO_2 and CH_4 surface fluxes as described previously. For both CO_2 and CH_4, the amplitude of the perturbation is α=25% and the time decorrelation length is $\sigma_w = 3$ time steps. For CH_4, the spatial correlation length scale is about 150 km at the equator ($\sigma_w = 3$ grid points for a perturbation field build on a regular $0.45^\circ \times 0.45^\circ$ grid).

3. ALL : same experiment as FLUX , but with perturbations on CO_2 and CH_4 tendencies.

Table 1: List of C-EDA experiments used in this report. For each experiment, Exp. *is the name used in the report to refer to this experiment,* Flux pert. *indicates if the fluxes where directly perturbed,* Tendency pert. *indicates if the* CO_2 *and* CH_4 *physical tendencies where perturbed,* Exp. Id. *is the experiment id, and* Colour code *is the colour used in the figures to refer to this experiment.*

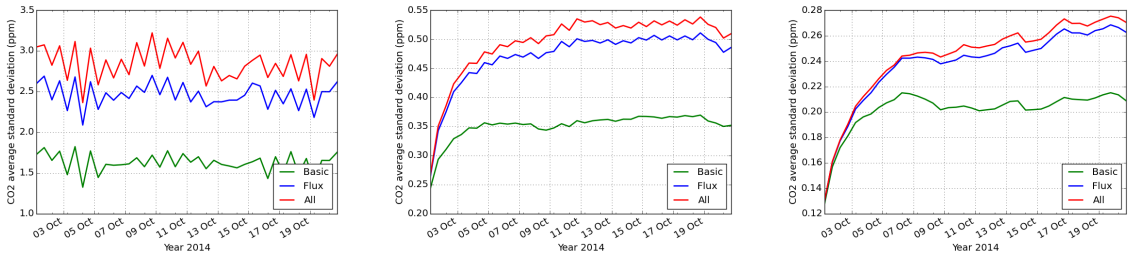
Exp.	Flux pert.	Tendency pert.	Exp. Id.	Colour code
BASIC	no	no	gdz5	
FLUX	yes	no	gehi	
ALL	yes	yes	gehx	

4 Results

4.1 Time evolution of the standard deviation

For each C-EDA experiment, each analysis time and each level we computed the ensemble standard deviation of the CO_2 and CH_4 fields from the 25 ensemble of 3-hr forecasts. We then computed a global average of the ensemble standard deviation in order to have the time evolution of the global average. Figure 7 presents the time series of the global averages for 3 specific levels: the surface, a level corresponding to about 850 hPa (lower troposphere) and a level corresponding to about 500 hPa (middle troposphere).

(a) CO_2 *Level 137 (surface)* *(b)* CO_2 *Level 114 (≈ 850* hPa*)* *(c)* CO_2 *Level 96 (≈ 500* hPa*)*



(d) CH_4 *Level 137 (surface)* *(e)* CH_4 *Level 114 (≈ 850* hPa*)* *(f)* CH_4 *Level 96 (≈ 500* hPa*)*

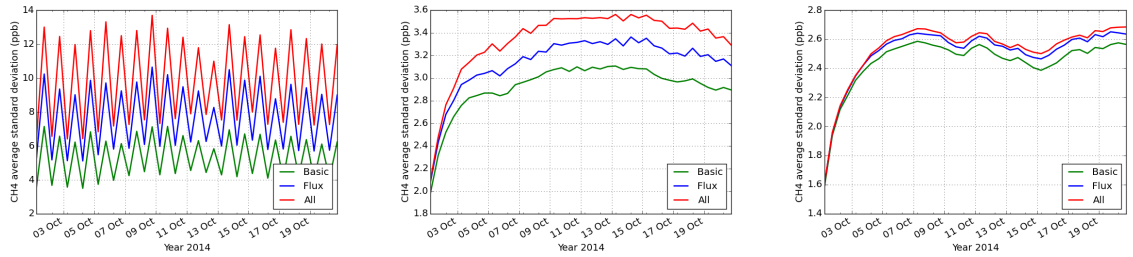


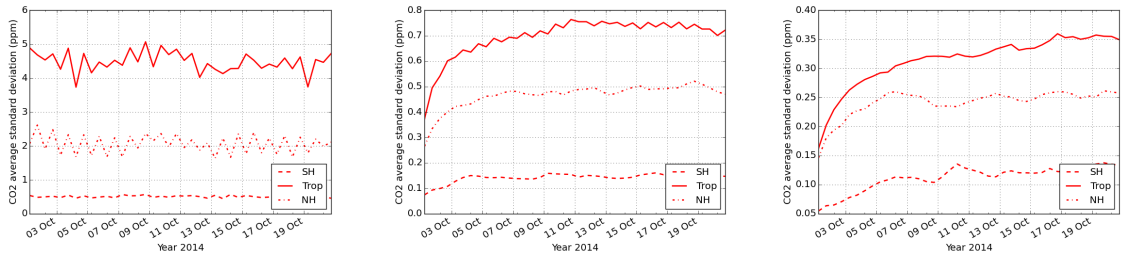
Figure 7: Time evolution of the global average of the ensemble spread for CO_2*: (a) to (c) and for* CH_4*: (d) to (f) for 3 model levels. The different colours are for different experiments (see Tab. 1 for details).*

4.1.1 *Surface evolution*

At the surface, the global averages of CO_2 and CH_4 ensemble standard deviation have significantly different values for the various experiments (Figs. 7a and d). For CO_2, with only the perturbation of the meteorological parameter, the BASIC experiment has an ensemble standard deviation of about 1.6 ppm on average. With the perturbation of the fluxes, the ensemble standard deviation from the FLUX experiment is about 2.5 ppm on average. Adding the physical tendency perturbation further increases the ensemble standard deviation to a value of about 3 ppm on average for the ALL experiment. This means that introducing more perturbation allowed to almost double the CO_2 spread between the members at the surface. Similar ratios between the experiments are found for CH_4.

A strong diurnal cycle is found in the ensemble standard deviations for CO_2 and CH_4 for the three C-EDA experiments. The time variations of CO_2 and CH_4 fields at the surface are mainly caused by the changes of the fluxes, the transport (with the vertical mixing in particular) and the boundary layer height. It is nevertheless unexpected to find such a strong diurnal cycle in global average as the diurnal cycle is different for each point of the globe. The reason behind this behaviour is that global averages of CO_2 and CH_4 ensemble standard deviations are largely influence by the values over some regions where the ensemble standard deviation is much higher than elsewhere (see Sec 4.2).

(a) CO_2 *Level 137 (surface)* *(b)* CO_2 *Level 114 (≈ 850* hPa*)* *(c)* CO_2 *Level 96 (≈ 500* hPa*)*



(d) CH_4 *Level 137 (surface)* *(e)* CH_4 *Level 114 (≈ 850* hPa*)* *(f)* CH_4 *Level 96 (≈ 500* hPa*)*

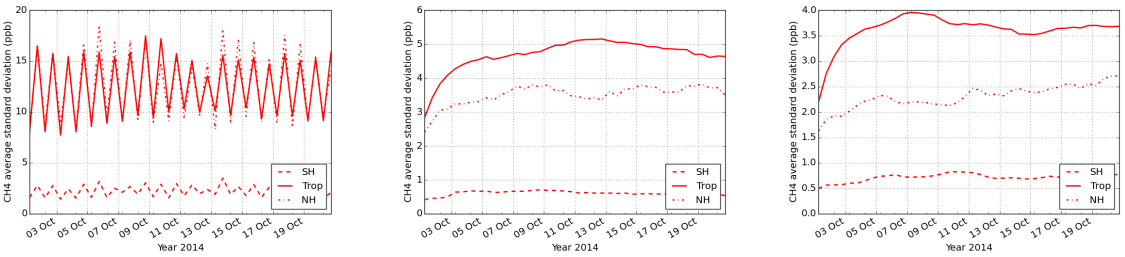


Figure 8: Time evolution of the regional averages of the ensemble spread for CO_2*: (a) to (c) and for* CH_4*: (d) to (f) for 3 model levels. The different line styles are for different regions: dash line for Southern Hemisphere (SH, latitude* $\leq 20°S$ *), full line for Tropics (Trop, latitude* $\geq 20°S$ *and* $\leq 20°N$*) and dotted line Northern Hemisphere (NH, latitude* $\geq 20°N$*).*

For CH_4 the surface fluxes are constant during the day and they vary slowly from day to day. The diurnal cycle of the CH_4 ensemble standard deviation is then most likely to come from the variability of the transport. For CO_2 the biogenic surface fluxes vary with the meteorological forcing. This could explain

why we have more day-to-day differences in the CO_2 ensemble standard deviation than in the CH_4 one. As an example, around 14 October, the CO_2 ensemble standard deviation at the surface does not vary as much as for the rest of the period (Fig. 7a).

For the ALL experiment, we decomposed the global average ensemble standard deviations by regional averages: Southern Hemisphere (SH, latitude $\leq 20°S$), Tropics (Trop, latitude $\geq 20°S$ and $\leq 20°N$) and Northern Hemisphere (NH, latitude $\geq 20°N$). For CO_2, the ensemble standard deviation is higher in the Tropics, with values more than twice the values of the NH and ten times the values of the SH (Fig. 8a). This difference between the regions could be explained firstly by the fact that the CO_2 surface fluxes are stronger over the Tropics especially in October when the NH senescence season started. Because the perturbation of the surface fluxes is multiplicative, the stronger the fluxes the larger the perturbation of the fluxes and the larger the spread between the members for the CO_2 concentration at the surface. The enhancement of the CO_2 ensemble standard deviation over the Tropics could also be a consequence of the response of the biogenic CO_2 surface fluxes with the meteorological forcing which is higher in this region.

The ensemble standard deviation of CH_4 at the surface is similar for the Tropics and NH with values about 6 times larger than for SH (Fig. 8d). CH_4 surfaces fluxes are much higher over land than over ocean. Due to the multiplicative perturbation of surface fluxes the spread of the surface fluxes between the ensemble members is larger over land than over ocean. As a consequence, the spread and therefore the ensemble standard deviation are larger at the surface for the CH_4 concentration over land. SH being mostly covered by water, the CH_4 ensemble standard deviation is lower than the two other regions. NH and Tropics have similar spread of CH_4 at the surface unlike CO_2 because of the non interaction between the meteorological forcing and CH_4 surface fluxes.

4.1.2 Middle troposphere evolution

The global averages of CO_2 and CH_4 ensemble standard deviation at ≈850 hPa and 500 hPa are growing during the first days before reaching a plateau for almost all experiments (Fig. 7). The spin up period is between 5 and 10 days, but the plateau is not completely reached after 20 days for the FLUX and ALL experiments for CO_2 at ≈500 hPa (Fig. 7c).

The initial conditions are the same for all members of the ensemble for both the meteorological fields and the greenhouse gases fields. The spin up period is the time during which the initial conditions are forgotten and after which each member becomes more independent to the others.

At these levels (≈850 hPa and 500 hPa), the spread between the member of the ensemble for CO_2 and CH_4 comes directly from the assimilation of perturbed observations, directly from the spread of the meteorological forcing (different transport between the members) and indirectly from the CO_2 and CH_4 spread at the surface that is slowly transported vertically. The comparison between the BASIC and the FLUX experiments indicates how each process (direct and indirect) impacts the spread of the ensemble.

For CH_4, the ALL experiment has slightly higher values of the global average ensemble spread than the FLUX experiment and the FLUX experiment has higher values than the BASIC one at ≈850 hPa (Fig. 7e). At ≈500 hPa, they all have similar global average ensemble spread (Fig. 7f). This means that the perturbation of the CH_4 fluxes has less impact in the CH_4 ensemble standard deviation in the middle atmosphere. The impact of the perturbation of the greenhouse gases tendencies is also weak at ≈850 hPa and almost null at ≈500 hPa.

For CO_2, the perturbation of the surface fluxes and the perturbation of the tendencies have more impact

on the global averages of the ensemble standard deviation (Figs. 7b and c). At ≈850 hPa, the perturbation of the surface fluxes increases by about 50% the standard deviation. This perturbation is also the reason why the spin up period is longer at ≈500 hPa. Moreover, the diurnal signal found at the surface is still present at higher altitudes, but it is strongly reduced.

4.2 Calibration

As described in Section 2.4, we computed the spread Σ for three regions (NH, Tropics and SH). For each Σ, we computed the equivalent ensemble mean background error Ω and arranges each Σ/Ω pair into 10 bins according to their values. For each bin, we computed the average of each pair over a period covering the last 10 days of the period under study, so we removed the first 10 days in order to account for the spin up period. The result is the skill/spread plots of Fig. 9 for CO_2 and Fig. 10 for CH_4. However, it is worth noticing that the discussion hereafter relies on the computation of the ensemble mean background error Ω and the choice of the experiment from where we get the true state $\boldsymbol{y}$ of Eq. (6).

The skill/spread plots at the surface show that the BASIC experiment does not have enough spread in general for CO_2 (Figs. 9a, d and g) and for CH_4 (Figs. 10a, d and g). It is particularly the case for the whole range of Σ values for NH, and for the largest values of Σ for SH, which are found over land (not shown). The highest values of Σ for the Tropics and for CO_2 are the only ones for which Σ and Ω are similar. The highest values of Σ for the Tropics are most likely associated with the biogenic CO_2 surface fluxes (that are influenced by the meteorological situation). The perturbation of the meteorological forcing over the Tropics seems then to bring enough variability to the the biogenic CO_2 surface fluxes.

Adding further perturbation to the surface fluxes makes the FLUX experiment have a better Σ/Ω ratio than the BASIC experiment for CO_2, except for the high values of Σ for CO_2 and for the whole range of Σ values for CH_4. This means that first the CO_2 surface flux perturbation helps to improve the CO_2 ensemble spread at the surface. The amplitude of the perturbation of α=25% seems nevertheless to be too much, or, if the highest values of Σ are associated with the highest values of the biogenic fluxes, the choice of a multiplicative perturbation could not be the best option. For CH_4, the amplitude of the perturbation of α=25% is for sure too much, but the choice of the multiplicative perturbation seems relevant as all the Σ bins have a similar Σ/Ω ratio. Adding the perturbation of the tendencies further increases the spread as described section 4.1 which makes the ALL experiment having too much spread at the surface for both CO_2 and CH_4.

Higher up in the troposphere, the BASIC experiment already has a spread Σ similar or higher than the mean background error Ω for both CO_2 (Figs. 9b, e, h for ≈850 hPa and Figs. 9c, f, i for ≈ 500 hPa) and CH_4 (Figs. 10b, e and h for ≈850 hPa and Figs. 10c, f and i for ≈ 500 hPa). We showed in section 4.1 that the spread of the ensemble is increased at these model levels by adding the perturbation of the surface fluxes and the perturbation of CO_2 and CH_4 tendencies. As a consequence, the spread is too large for CO_2 and CH_4 at these model levels for the FLUX and ALL experiments.

We finally computed the time average of the global mean of the raw ensemble standard deviation and the scaled ensemble standard deviation over the last ten days of the simulations and for each model level. Figure 11 presents the ratio between the averages for CO_2 and CH_4 and for the three C-EDA experiments as a function of the model level. It confirms the previous findings, namely (i) a too low spread in the BASIC experiment in the lower troposphere and down to the surface and (ii) a too large spread in the two other C-EDA experiments. Moreover, it shows that the introduction of direct surface flux perturbation of CO_2 impacts the ensemble spread on the whole troposphere (up to about model level 60) and the stratosphere as well (up to about model level 20). This is also true for CH_4 but to a less extend.

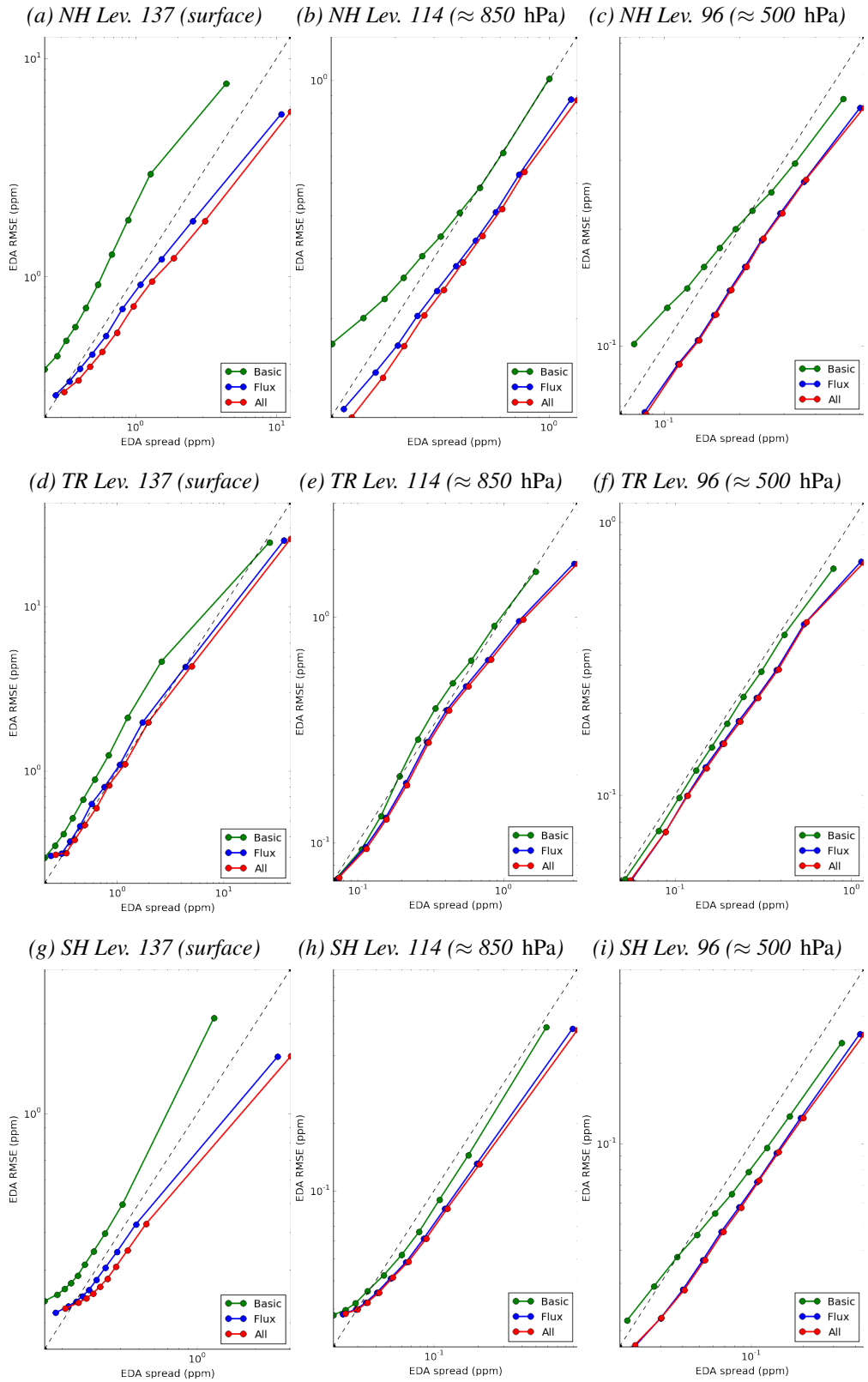


Figure 9: Spread-error diagram for CO_2 for three model levels (left to right) and three regions (top to bottom), averaged over the last 10 days of the period under study.

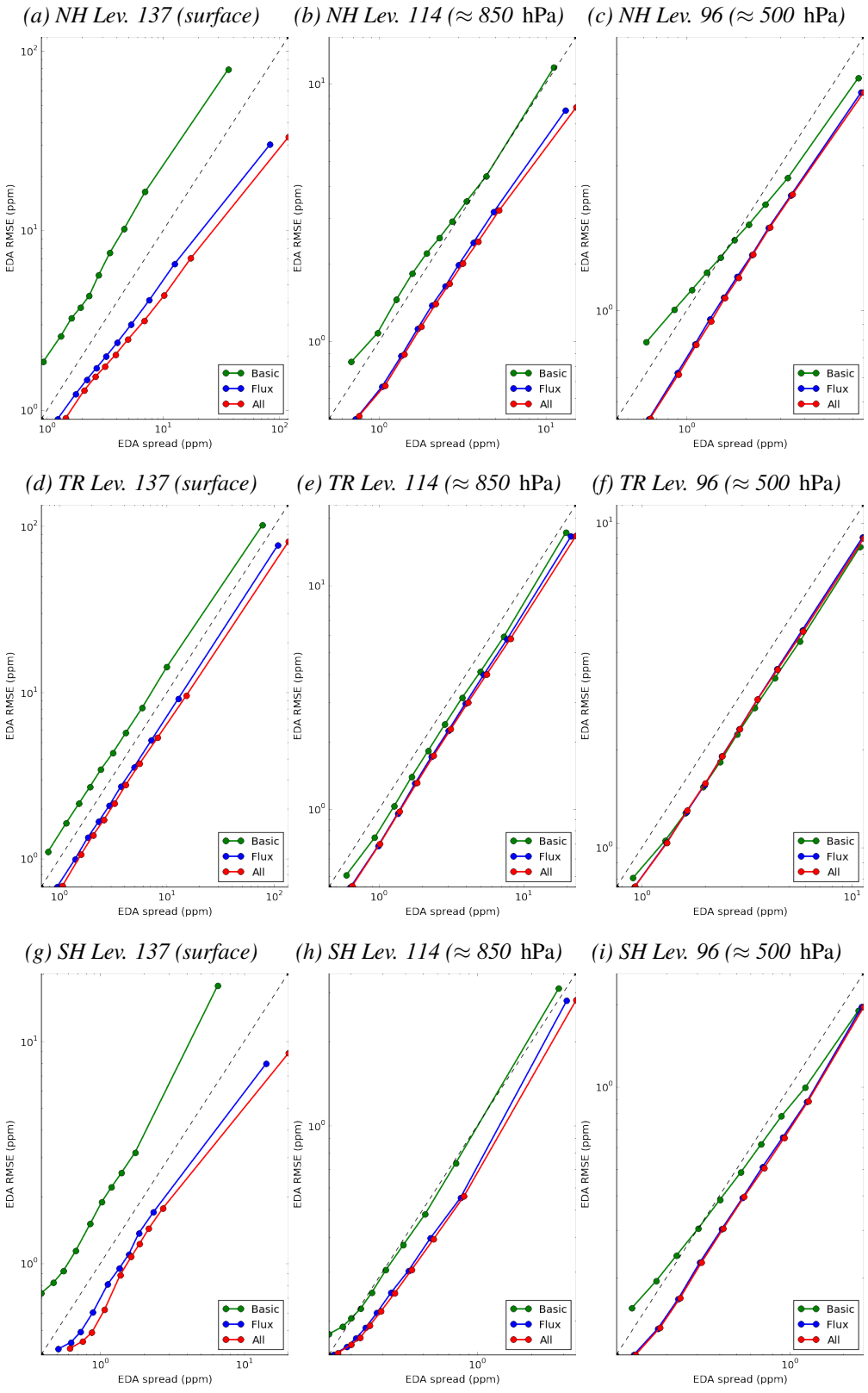


Figure 10: Same as Fig.9 but for CH_4.

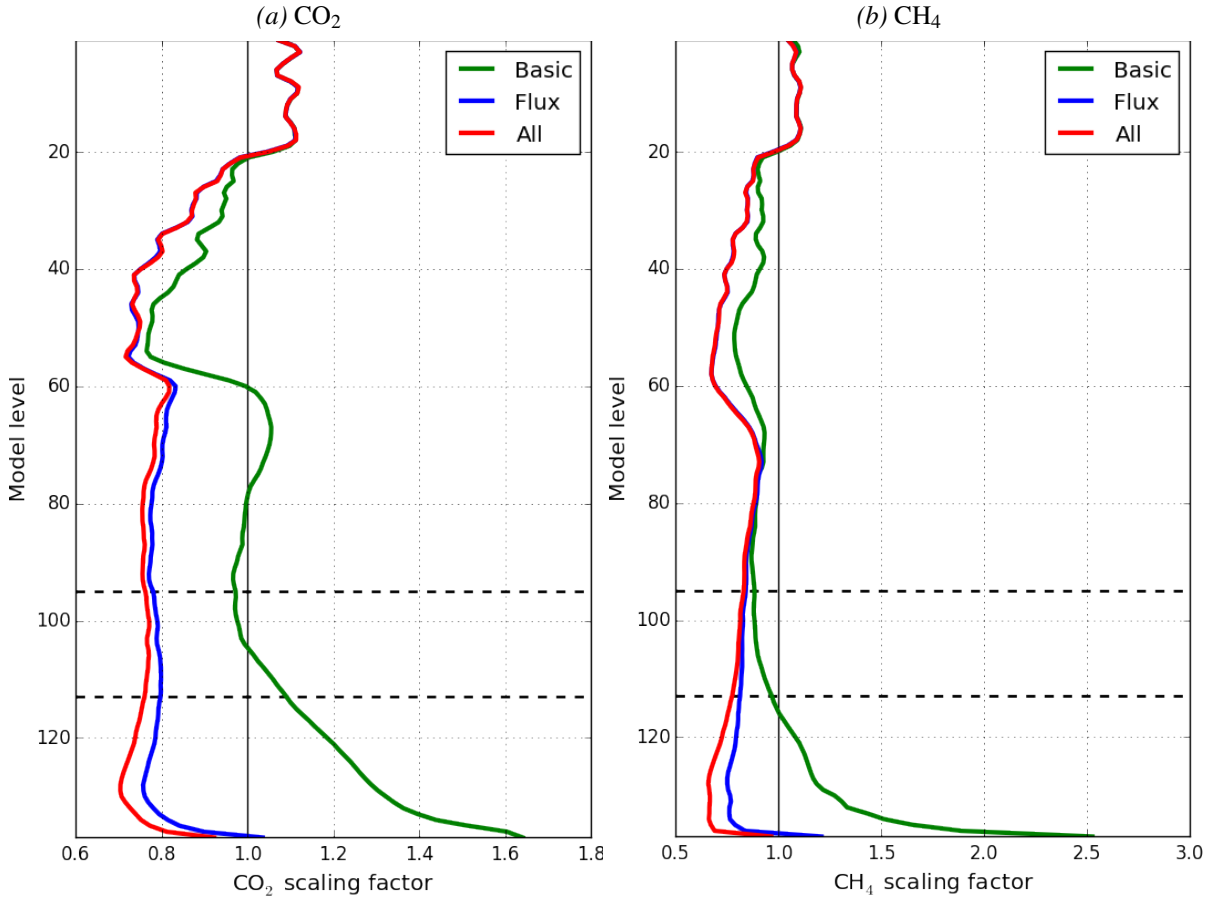


Figure 11: Ratio between the time and space averages of the scaled ensemble standard deviation and the raw ensemble standard deviation for (a) CO_2 *and (b)* CH_4 *as a function of model levels. Results for the period 10–20 October. Dotted vertical lines for model level 114 and 96.*

4.3 Scaled ensemble standard deviation

In this section, we discuss the difference between the scaled ensemble standard deviation of CO_2 and CH_4 of the three experiments. We are focusing on the same three model levels as previously and on the ensemble standard deviation computed for the 0900Z analysis. We also present the scaled ensemble standard deviation for the 2100Z analysis for the ALL experiment and the three selected model levels.

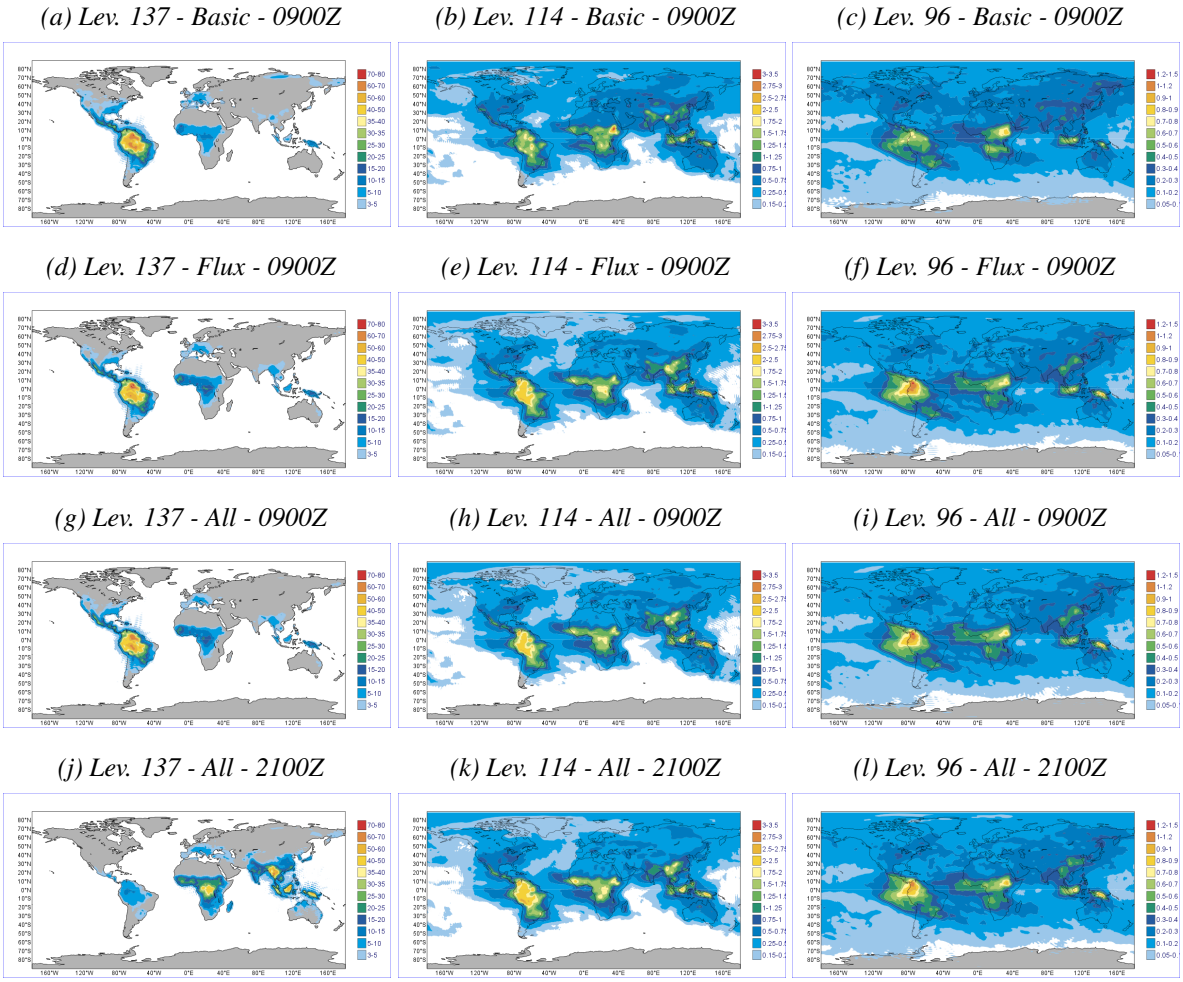


Figure 12: Scaled standard deviation of CO_2 *(in ppm) averaged over a period between 10 and 20 October for three C-EDA experiments (top to bottom) and for three model levels: from left to right, surface (level 137),* ≈ 850 hPa *(level 114) and* ≈ 500 hPa *(level 96). The scaled standard deviation is averaged for the 0900Z analysis apart for the last experiment where the 2100Z analysis is also presented.*

Figures 12 and 13 show that for each level the large scale structures in the CO_2 and CH_4 scaled ensemble standard deviation are similar for the three experiments. This suggests that the calibration method allows to compensate the lack or excess of spread depending on the experiment. There are nevertheless some differences. For example the CO_2 scaled ensemble standard deviation has lower values over Central and West Africa at the surface for the BASIC experiment than for the two other experiments. The background value is on the other hand larger for the BASIC experiment than for the others from model levels 114 to 96. For CH_4, the scaled ensemble standard deviation has larger values over Asia at the surface for the

BASIC experiment, and the background value is also larger, especially at high latitudes north.

The diurnal cycle discussed in section 4.1 is also well illustrated in the scaled ensemble standard deviation maps at the surface for CO_2 (Figs 12 g and j) and for CH_4 (Figs 13 g and j). The spread is larger over Amazon than elsewhere for the 0900Z analysis, and larger over the Tropical Africa (for CO_2) and in Asia (for CO_2 and CH_4) for the 2100Z analysis. This highlights the effect of the boundary layer height. Higher up in the troposphere, the scaled ensemble standard deviation is much less sensitive to the diurnal cycle.

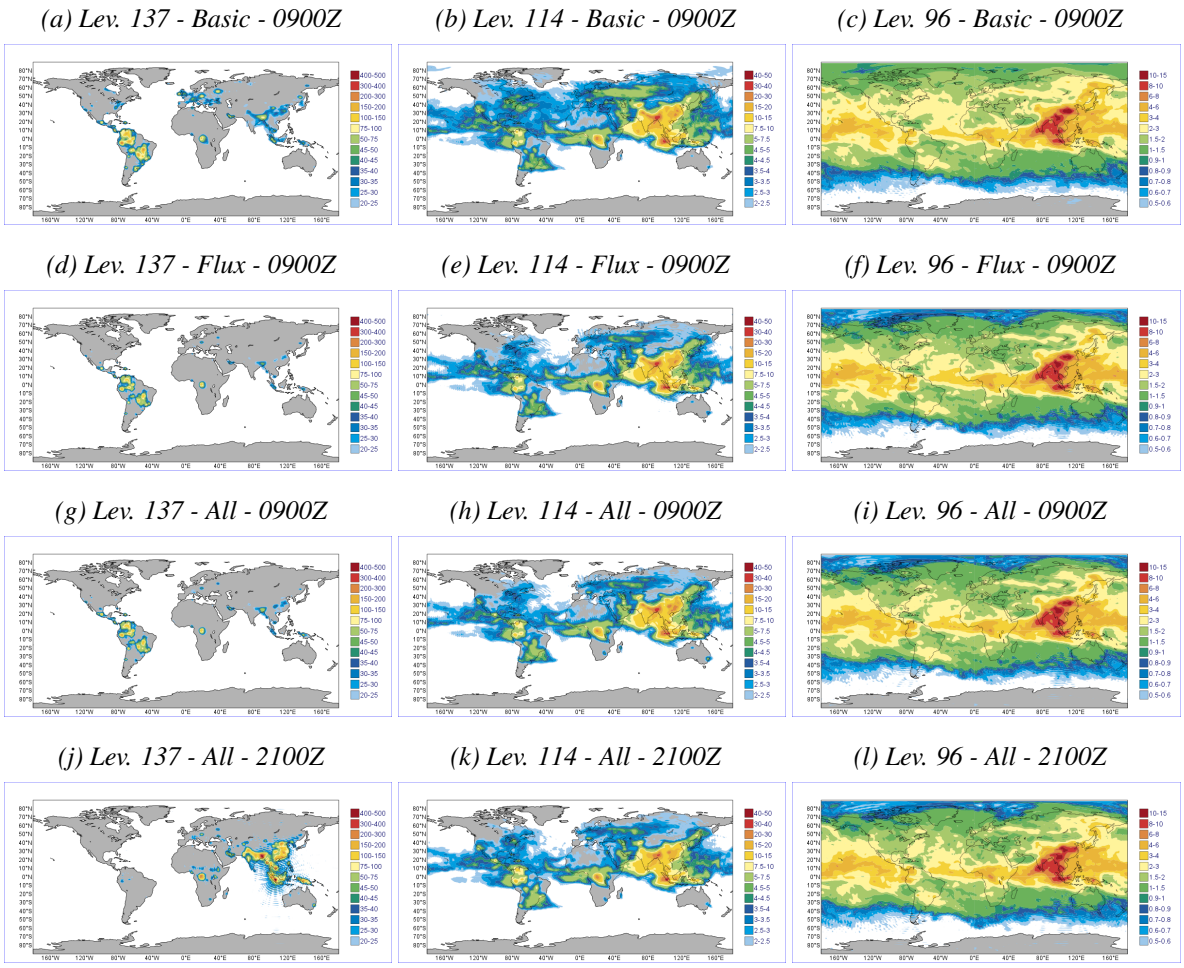


Figure 13: Same as Fig. 12 but for CH_4 *(in ppb)*

4.4 Spatial correlation

We finally investigated the spatial correlation of the background error. For that purpose, we computed the spatial correlation between the ensemble error of a given point of the grid and the ensemble error of the surrounding points. We estimated the ensemble error by removing the value of the control to each of the 25 members of the ensemble. We also concatenated the ensemble errors computed over the last 10 days of the period under study. This allows to increase the sample and to have more significance in the computed correlation.

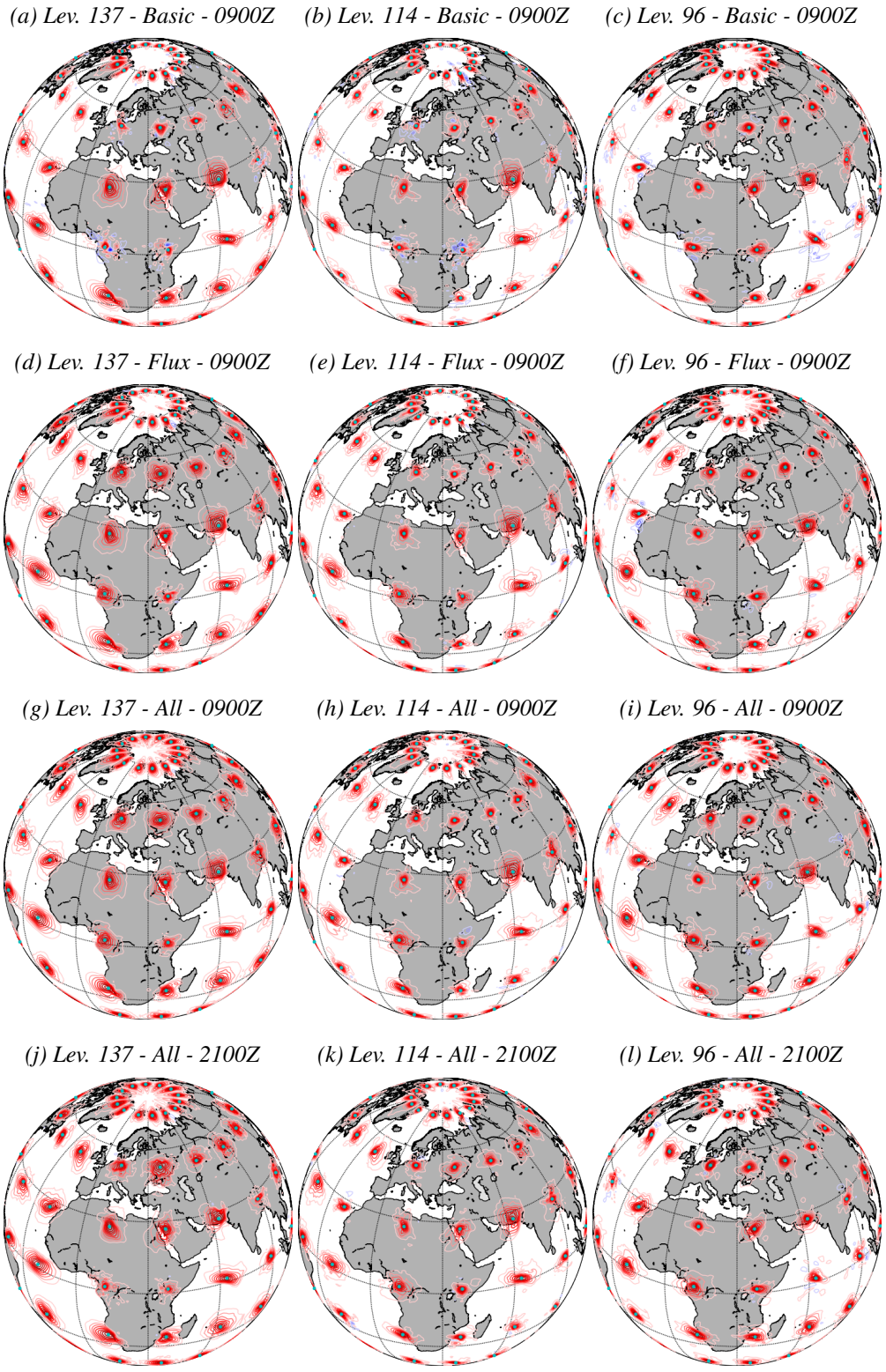


Figure 14: Background error spatial correlation of CO_2 *at various points of the Global separated by about 25° (cyan squares) and averaged over the period between 10 and 20 October for the 0900Z analysis. From left to right: model levels, from top to bottom:* BASIC , FLUX *and* ALL *experiment. The last row is for the 2100Z analysis of the* ALL *experiment. Dark red for correlation of 1 to dark blue for a correlation of -1, with a 0.1 contour interval.*

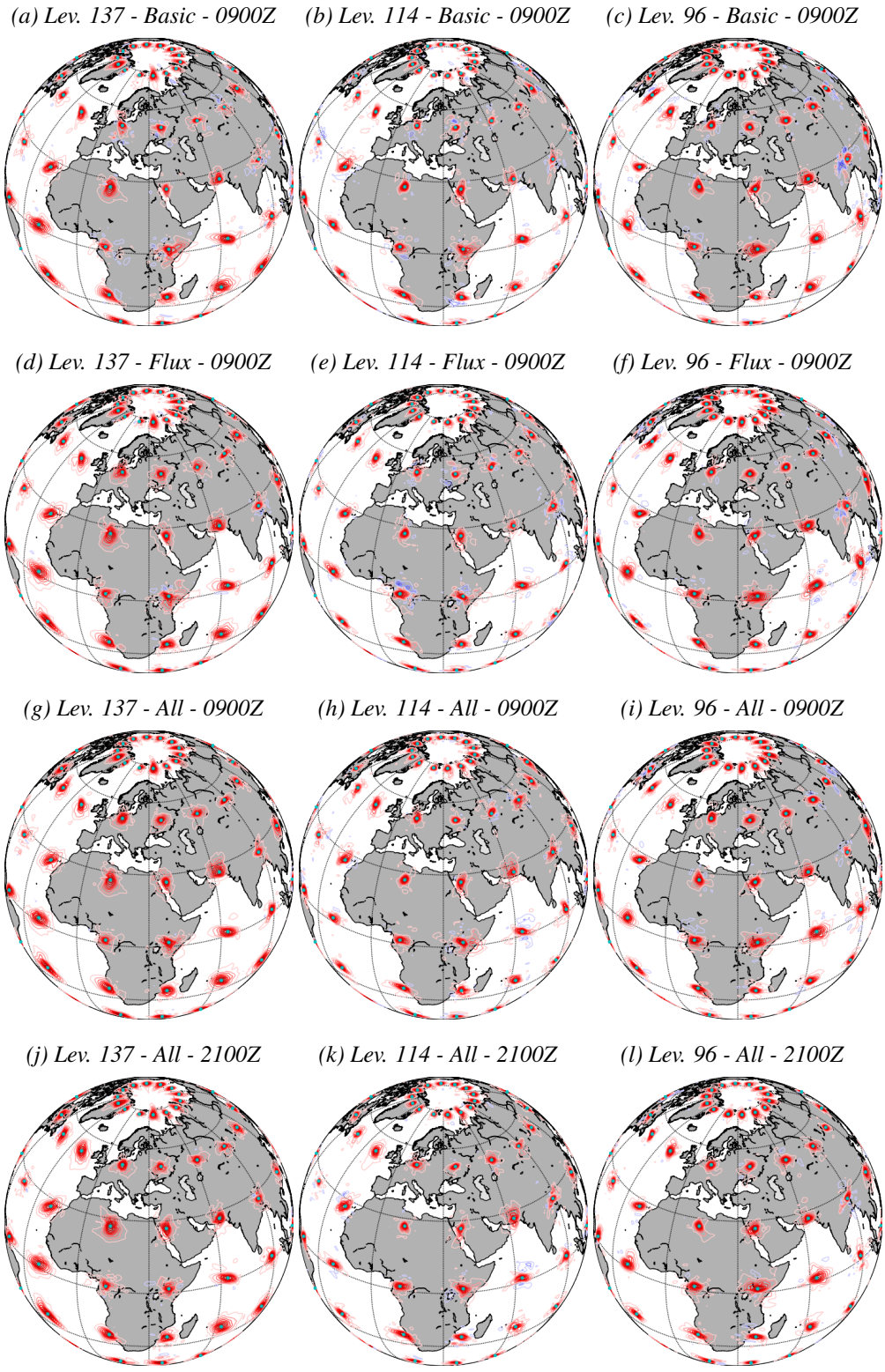


Figure 15: Same as Fig. 14 but for CH_4.

We repeated the above procedure for several points of the globe to have an overview of the spatial distribution of the spatial correlation of the background error. We also computed the spatial correlation for the same three model levels and for the three C-EDA experiments as before, and only for the 0900Z analysis.For the ALL experiment, we also computed the spatial correlation for the 2100Z analysis.

Figure 14 presents the results for CO_2. It shows that the shape of the spatial correlation varies a lot between the various selected locations at the surface for the BASIC experiment (Fig. 14a). For example, the spatial correlation of the CO_2 background error has a much larger spread for the point located in Northern Africa than for the points located in Central and West Africa. This reflects the difference between regions affected by local sources and other regions more representative of background CO_2. Another example is the different shapes of the spatial correlation over Europe from West to East, even if the correlation length is similar. A key feature is the weak correlation of the background error between land and ocean with for example the ocean points close to the Southern Africa coast or the land point close to Iran coast. Higher up in the atmosphere, this feature disappears and elsewhere the correlation structures are similar that the one at the surface (Figs. 14b and c).

The addition of more direct perturbation of the surface fluxes in the FLUX experiment increases the length scale of the spatial correlation where they were narrow in the BASIC experiment (Figs. 14d to 14f). This increase is particularly important at the surface for the two points of Central and West Africa, and for the points over Europe to a lesser extend. The CO_2 tendency perturbation in the ALL experiment hardly impacts the shape of the spatial correlation (Figs. 14g to i). There are nevertheless few differences like for the ocean point located near the west coast of Northern Africa for which the ALL experiments does not present negative correlations like the FLUX experiment at 500 hPa.

The difference in the ALL experiment between the 0900Z analysis and the 2100Z analysis is more the shape of the CO_2 spatial correlation of the background error than in the size of the characteristic length scale (Figs. 14j to l). For example, over Central Europe, the spatial correlation has a more noisy shape for the 2100Z analysis than for the 0900Z analysis at the surface. Another example is the presence of negative correlation for the point located in the middle of the Indian Ocean in the 2100Z analysis, where the correlation are smoother in the 0900Z analysis

Figure 15 presents the results for CH_4. The conclusions one can draw from it are similar to the conclusions for CO_2. The shape of the CH_4 spatial correlation varies across the globe. The characteristic length scale of the spatial correlation is nevertheless smaller for CH_4 than for CO_2. The increase of the characteristic length scale due to the addition of perturbation on the surface fluxes is also smaller than for CO_2. Lastly, there are some differences between the 0900Z analysis and the 2100Z analysis too.

5 Conclusions

The Ensemble of Data Assimilations (EDA) is currently the method chosen at ECMWF to estimate the background error statistics for the meteorological analysis. We are investigating the use of the EDA to also estimate the background error statistics for the atmospheric tracers analysis in Composition-IFS (C-IFS). This report presents the first step by describing specific modifications of the current EDA to account for particularities of atmospheric tracers. The resulting configuration is referred to as Composition-EDA (C-EDA).

We focused the document on two of the main greenhouse gases: carbon dioxide (CO_2) and methane (CH_4). As for other atmospheric tracers, CO_2 and CH_4 are highly sensitive to the surface fluxes. One modification brought to C-EDA was to perturb the surface fluxes in order to represent the uncertainty

associated with them. Two types of surface fluxes perturbations were implemented. For CO_2 we chose to perturb some parameters of the surface fluxes model. For CH_4 we chose a more familiar approach using a two-dimensional perturbation field. Another perturbation was added on the CO_2 and CH_4 tendencies as for the physical tendencies in the default EDA. This adds indirect perturbation of the CO_2 and CH_4 surface fluxes as the surface fluxes are transformed as tendencies in the boundary layer through the vertical mixing in C-IFS. For CH_4, this adds perturbation of the chemical tendency.

A default C-EDA experiment was run with only perturbation of the meteorological parameters. This experiment showed that the spread of the ensemble is larger at the surface than in the troposphere for both CO_2 and CH_4. Moreover the ensemble standard deviation has a strong diurnal cycle at the surface and has strong spatial variations at the surface and higher up in the troposphere. This is an important result as the current standard deviation of the background error is assumed to be constant in time and space for the assimilation. The first main conclusion of this work is that a time and space dependent standard deviation of the background error should be implemented in the assimilation of atmospheric tracers. We showed that the spatial correlation of the background error has a time dependence too comparing the 0900Z analysis with the 2100Z analysis. Having a fully time-dependent background error (standard deviation and correlation) should also benefit the analysis of atmospheric tracers.

Even if the spread of the ensemble is larger at the surface, it is estimated to be too low when comparing to the ensemble mean background error. Adding directly a perturbation on the surface fluxes helps increasing the ensemble spread, but it is found that the current amplitude of the perturbation (25%) is too large. On the other hand, posterior diagnostics suggest that the ensemble standard deviation is still too low with these perturbations on the surface fluxes. The computation of the calibration of the ensemble standard deviation may have to be revisited. The calibration of the ensemble standard deviation of the meteorological parameters is carried out using the high resolution deterministic forecast as the truth to estimate the ensemble mean background error. Here, we used an experiment with a configuration very similar to the ensemble control. Therefore the ensemble mean background error may be too low.

In this report, we chose two types of surface flux perturbations with a given amplitude, a given decorrelation time scale and a given spatial decorrelation scale for CH_4 surface fluxes. We proved that the perturbation of the surface fluxes have an influence on the ensemble standard deviation even after the calibration, and an influence on the spatial correlation of the background error.

The method implemented in this document to perturb the surface fluxes is derived from the methodology used for the EDA. One difference is that the surface fluxes are directly perturbed and not only the tendencies. This could be further improved by having different surface flux perturbations for different emission categories (anthropogenic and biogenic for example) or depending on the surface (sea or land). For instance, the localisation of anthropogenic surface fluxes is well known, with usually a low length scale for the error correlation and potentially high uncertainty in the amplitude. The localisation of CH_4 fluxes could vary significantly depending on atmospheric conditions and the length scale for the error correlation could be large, as the error is likely to be correlated over the surface covered by the wetland. Having a better representation of the surface fluxes uncertainty in the C-EDA would allow to have a better estimate of the CO_2 and CH_4 uncertainty and therefore should improve the quality of the analysis.

Acknowledgements

The authors would like to thank Anna Agustí-Panareda, Antje Inness and Richard Engelen for providing comments and suggestions on preliminary versions of this document. The authors are also grateful to Elías Hólm for his help while learning to run an EDA experiment at an early stage.

References

Agustí-Panareda, A., Massart, S., Chevallier, F., Balsamo, G., Boussetta, S., Dutra, E., and Beljaars, A. (2016). A biogenic CO_2 flux adjustment scheme for the mitigation of large-scale biases in global atmospheric CO_2 analyses and forecasts. *Atmos. Chem. Phys. Discuss.*, 2016:1–45.

Agustí-Panareda, A., Massart, S., Chevallier, F., Boussetta, S., Balsamo, G., Beljaars, A., Ciais, P., Deutscher, N. M., Engelen, R., Jones, L., Kivi, R., Paris, J.-D., Peuch, V.-H., Sherlock, V., Vermeulen, A. T., Wennberg, P. O., and Wunch, D. (2014). Forecasting global atmospheric CO_2. *Atmos. Chem. Phys.*, 14(21):11959–11983.

Bonavita, M., Hólm, E., Isaksen, L., and Fisher, M. (2015). The evolution of the ECMWF hybrid data assimilation system. *Q.J.R. Meteorol. Soc.*, 142(694):287–303.

Bonavita, M., Isaksen, L., and Hólm, E. (2012). On the use of EDA background error variances in the ECMWF 4D-Var. *Q.J.R. Meteorol. Soc.*, 138(667):1540–1559.

Boussetta, S., Balsamo, G., Beljaars, A., Agusti-Panareda, A., Calvet, J.-C., Jacobs, C., van den Hurk, B., Viterbo, P., Lafont, S., Dutra, E., Jarlan, L., Balzarolo, M., Papale, D., and van der Werf, G. (2013). Natural carbon dioxide exchanges in the ecmwf integrated forecasting system: Implementation and offline validation. *J. Geophys. Res. Atmos.*, 118:1–24.

Massart, S., Agusti-Panareda, A., Aben, I., Butz, A., Chevallier, F., Crevosier, 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. *Atmos. Chem. Phys.*, 14(12):6139–6158.

Massart, S., Agustí-Panareda, A., Heymann, J., Buchwitz, M., Chevallier, F., Reuter, M., Hilker, M., Burrows, J. P., Hase, F., Desmet, F., Feist, D. G., and Kivi, R. (2015). Ability of the 4D-Var analysis of the GOSAT BESD XCO_2 retrievals to characterize atmospheric CO_2 at large and synoptic scales. Submitted to Atmos. Chem. Phys.

Palmer, T., Buizza, R., Doblas-Reyes, F., Jung, T., Leutbecher, M., Shutts, G., Steinheimer, M., and Weisheimer, A. (2009). Stochastic parametrization and model uncertainty. ECMWF Technical Memorandum 598, ECMWF.

Peuch, V.-H., Engelen, R., Flemming, J., Huijnen, V., and Inness, A. (2015). Atmospheric composition in ECMWF's Integrated Forecasting System. *ECMWF Newsletter*, 143:20–25.