Variational regional inverse modeling of reactive species emissions with PYVAR-CHIMERE-v2019

Audrey Fortems-Cheiney[1], Isabelle Pison[1], Grégoire Broquet[1], Gaëlle Dufour[2], Antoine Berchet[1], Elise Potier[1], Adriana Coman[2], Guillaume Siour[2], and Lorenzo Costantino[2]

[1]Laboratoire des Sciences du Climat et de l'Environnement, LSCE-IPSL (CEA-CNRS-UVSQ), Université Paris-Saclay, 91191 Gif-sur-Yvette, France.

[2]Laboratoire Interuniversitaire des Systèmes Atmosphériques, UMR CNRS 7583, Université Paris Est Créteil et Université Paris Diderot, Institut Pierre Simon Laplace, Créteil, France.

Abstract

Up-to-date and accurate emission inventories for air pollutants are essential for understanding their role in the formation of tropospheric ozone and particulate matter at various temporal scales, for anticipating pollution peaks and for identifying the key drivers that could help mitigate their **concentrations**. This paper describes the Bayesian variational inverse system PYVAR-CHIMERE, which is now adapted to the inversion of reactive species. Complementarily with bottom-up inventories, this system aims at updating and improving the knowledge on the high spatio-temporal variability of emissions of air pollutants and their precursors. The system is designed to use any type of observations, such as satellite observations or surface station measurements. The potential of PYVAR-CHIMERE is illustrated with inversions of both CO and NO_x emissions in Europe, using the MOPITT and OMI satellite observations, respectively. **In these cases, local increments on CO emissions can reach more than +50%, with increases located mainly over Central and Eastern Europe, except in the south of Poland, and decreases located over Spain and Portugal. The illustrative cases for NO_x emissions also lead to large local increments (> 50%), for example over industrial areas (e.g., over the Po Valley) and over the Netherlands. The good behavior of the inversion is shown through statistics on the concentrations: the mean bias, RMSE, standard deviation and correlation between the simulated and observed concentrations. For CO, the mean bias is reduced by about 27% when using the posterior emissions, the RMSE and the standard deviation are reduced by about 50% and the correlation is strongly improved (0.74 when using the posterior emissions against 0.02); for NO_x, the mean bias is reduced by about 24%, the RMSE and the standard deviation are reduced by about 7% but the correlation is not improved.**

1. Introduction

The degradation of air quality is a worldwide environmental problem: 91% of the world's population have breathed polluted air in 2016 according to the World Health Organization (WHO), resulting in 4.2 millions of premature deaths every year [WHO, 2016]. The recent study of Lelieveld et al. [2019] even suggests that the health impacts attributable to outdoor air pollution are substantially higher than previously assumed (with 790,000 premature deaths in the 28 countries of the European Union against the previously estimated 500,000 [EEA, 2018]). The main regulated primary (i.e. directly emitted in the atmosphere) anthropogenic air pollutants are carbon monoxide (CO), nitrogen oxides (NO_x =NO+NO_2), sulfur dioxide (SO_2), ammonia (NH_3), volatile organic compounds (VOCs), and primary particles. These primary air pollutants are precursors of secondary (i.e. produced in the atmosphere through chemical reactions) pollutants such as ozone (O_3) and Particulate Matter (PM), which are also threatening to both human health and ecosystems. Monitoring concentrations and quantifying emissions are still challenging and limit our capability to forecast air quality to warn population and to assess i) the exposure of population to air pollution and ii) the efficiency of mitigation policies.

Bottom-up (BU) inventories are built in the framework of air quality policies such as The Convention on Long-Range Transboundary Air Pollution (LRTAP**,** http://www.unece.org) for air pollutants. Based on national annual inventories, research institutes compile gridded global or regional, monthly inventories (mainly for the US, Europe and China) with a high spatial resolution (currently regional or city scale inventories are typically finer than 0.1°x0.1°). These inventories are constructed by combining available (economic) statistics data from different detailed activity sectors with the most appropriate emission factors (defined as the average emission rate of a given species for a given source or process, relative to the unit of activity **in a given administrative area**). It is important to note that the activity data (often statistical data) has an inherent uncertainty and that its reliability may vary between countries or regions. In addition, the emission factors bear large uncertainties in their quantification [Kuenen et al., 2014; EMEP/EEA, 2016; Kurokawa et al., 2013]. Moreover, these inventories are often provided at the annual or monthly scale with typical temporal profiles to build the weekly, daily and hourly variability of the emissions. The combination of uncertain activity data, emission factors and emission timing can be a large source of uncertainties, if not errors, for forecasting or analyzing air quality [Menut et al., 2012]. Finally, since updating the inventories and gathering the required data for a given year is costly in time, manpower and money, only a few institutes have offered estimates of the gaseous pollutants for each year since 2011 (i.e, **European Monitoring and Evaluation Programme** EMEP updated until the year 2017, MEIC updated until the year 2017 to our knowledge). Nevertheless, using

knowledge from inventories and air quality modeling, emissions have been mitigated. For example, from 2010 to nowadays, emissions in various countries have been modified and/or regional trends have been reversed **downwards** (e.g., the decrease of NO_x emissions over China since 2011 [de Foy et al., 2016]), leading to significant changes in the atmospheric composition. Consequently, the knowledge of precise and updated budgets, together with seasonal, monthly, weekly and daily variations of gaseous pollutants driven, amongst other processes, by the emissions are essential for understanding their role in the formation of tropospheric ozone and PMs at various temporal scales, for anticipating pollution peaks and for identifying the key drivers that could help mitigate these **concentrations**.

In this context, complementary methods have been developed for estimating emissions using atmospheric observations. They operate in synergy between a chemistry-transport model (CTM) which links the emissions to the atmospheric concentrations, atmospheric observations of the species of interest, and statistical inversion techniques. A number of studies using inverse modeling were first carried out for long-lived species such as greenhouses gases (GHGs) (e.g., carbon dioxide CO_2 or methane CH_4) at the global or continental scales [Hein et al., 1997; Bousquet et al. 1999], using surface measurements. Later, following the development of monitoring station networks, the progress of computing power, and the use of inversion techniques more appropriate to non-linear problems, these methods were applied to shorter-lived molecules such as CO. For these various applications (e.g., for CO_2, CH_4, CO), the quantification of sources was solved at the resolution of large regions [Pétron et al., 2002]. Finally, the growing availability and reliability of observations since the early 2000s (in-situ surface data, remote sensing data such as satellite data), the improvement of the global CTMs, of the computational capacities and of the inversion techniques have increased the achievable resolution of global inversions, up to the global transport model grid cells, i.e. typically with a spatial resolution of several hundreds of square kilometers [Stavrakou and Muller, 2006; Pison et al., 2009; Fortems-Cheiney et al., 2011; Hooghiemstra et al., 2012; Yin et al., 2015; Miyazaki et al., 2017, Zheng et al., 2019].

Today, the scientific and societal issues require an up-to-date quantification of pollutant emissions at a higher spatial resolution than the global one and imply to widely use regional inverse systems. However, although they are suited to reactive species such as CO and NO_x, and their very large spatial and temporal variability, they have hardly been used to quantify pollutant emissions. Some studies inferred NO_x [Pison et al., 2007; Tang et al., 2013] and VOC emissions [Koohkan et al., 2013] from surface measurements. Konovalov et al. [2006, 2008, 2010], Mijling et al. [2012, 2013], van der A et al. [2008], Lin et al. [2012] and Ding et al. [2017] have also shown that satellite

observations are a suitable source of information to constrain NO_x emissions. These regional inversions using satellite observations were often based on Kalman Filter (KF) schemes [Mijling et al., 2012, 2013; van der A et al., 2008; Lin et al., 2012; Ding et al., 2017].

Variational inversion systems allow solving for high dimensional problems, typically solving for the fluxes at high spatial and temporal resolution, which can be critical to fully exploit satellite images. Here, we present the Bayesian variational atmospheric inversion system PYVAR-CHIMERE for the monitoring of anthropogenic emissions **of reactive species** at **the regional scale**. It is based on the Bayesian variational assimilation code PYVAR [Chevallier et al. 2005] and on the regional state-of-the-art CTM CHIMERE **[Menut et al., 2013; Mailler et al., 2017]**. **CHIMERE is dedicated to the study of regional atmospheric pollution events [e.g., Ciarelli et al., 2019; Menut et al., 2020], included in the operational ensemble of the Copernicus Atmosphere Monitoring Service (CAMS) regional services. The main strengths of PYVAR-CHIMERE come from the strengths of CHIMERE and from its high modularity for the definition of the control vector. CHIMERE is indeed an extremely flexible code, in particular for the definition of the chemical scheme.**
The PYVAR-CHIMERE system takes advantage of the previous developments for the quantification of fluxes of long-lived GHG species such as CO_2 [Broquet et al., 2011] and CH_4 [Pison et al., 2018] at the regional to the local scales, but now solves for reactive species such as CO and NO_x. It has also a better level of robustness, clarity, portability, and modularity than these previous systems. Variational techniques require the adjoint of the model to compute the sensitivity of simulated atmospheric concentrations to corrections of the fluxes. CHIMERE is one of the **few** CTMs **for which the adjoint has been coded**. For global models**, they include**: GEOS-CHEM [Henze et al., 2007], IMAGES [Stavrakou and Muller, 2006], TM5 [Krol et al., 2008], GELKA [Belikov et al., 2016] and LMDz [Chevallier et al., 2005; Pison et al., 2009] ; for limited-area models **they include**: CMAQ [Hakami et al., 2007], EURAD-IM [Elbern et al., 2007], RAMS/CTM-4DVAR [Yumimoto et Uno, 2006], WRF-CO2 4D-Var [Zheng et al., 2018]).

The principle of variational atmospheric inversion and the configuration of PYVAR-CHIMERE are described in Section 2 and in Section 3, respectively. Details about the forward**,** tangent-linear and adjoint codes of CHIMERE are also given. Then, the potential of PYVAR-CHIMERE is illustrated in Section 4 with the optimization of European CO and NO_x emissions, constrained by observations from the Measurement of Pollution in the Troposphere (MOPITT) and from the Ozone Monitoring Instrument (OMI) satellite instruments, respectively.

2. Principle of Bayesian variational atmospheric inversion

In what follows, we use the notations and equations used in the inverse modeling community [Rayner et al., 2019]. The Bayesian variational atmospheric inversion method adjusts a set of control parameters, including parameters related to the emissions whose estimate is the primary target of the inversion.

The prior information about the parameters x to be optimized during the inversion process is given by the vector $\mathbf{x^b}$. The parameters to be optimized can be surface fluxes but may also include initial or boundary conditions for example, as explained in Section 3.4. The adjustments are applied to prior values, usually taken, for the emissions, from pre-existing BU inventories. The principle is to minimize, on the one hand, the departures from the prior estimates of the control parameters, which are weighted by the uncertainties in these estimates (called hereafter "prior uncertainties"), and, on the other hand, the differences between simulated and observed concentrations, which are weighted by all other sources of uncertainties explaining these differences (called hereafter all together "observation errors"). In statistical terms, the inversion searches for the most probable estimate of the control parameters given their prior estimates, observations, CTM and their associated uncertainties. The solution, which will be called posterior estimate, is found by the iterative minimization of a cost function J [Talagrand et al., 1997], defined as:

$$J(\boldsymbol{x}) = \left(\boldsymbol{x} - \boldsymbol{x^b}\right)^T \boldsymbol{B}^{-1}\left(\boldsymbol{x} - \boldsymbol{x^b}\right) + (H(\boldsymbol{x}) - \boldsymbol{y})^T \boldsymbol{R}^{-1}(H(\boldsymbol{x}) - \boldsymbol{y}) \quad \text{(Eq. 1)}$$

H is the non-linear observation operator that projects the **control** vector **x** onto the observation space. In most of the variational atmospheric inversion cases (such as those described in Section 4), **the observation operator includes the operations performed by the CTM in linking the emissions to the concentrations and any other transformation to compute the simulated equivalent of the observations such as an interpolation or an extraction and averaging of the simulated concentration** fields (see Section 3.5). The observations in **y** could be surface measurements and/or remote sensing data such as satellite data. The prior uncertainties and the observation errors are assumed to be **unbiased** and to have a Gaussian distribution. Consequently, the prior uncertainties are characterized by their covariance matrix **B** and the observation errors are characterized by their covariance matrix **R.** By definition, the observation errors combine errors in both the data and the observation operator, in particular measurement errors and errors in the conversion of satellite measurement into concentration data, errors from the CTM, representativity errors due to the comparison between point measurements and gridded models or due to the representation of the fluxes as gridded maps at a given spatial resolution, and aggregation errors associated with the optimization of emissions at a given spatial and/or temporal resolution (as

specified in the control vector) that is different from (usually coarser than) that of the CTM [Wang et al., 2017].

For inversions with observation and control vectors having a high dimension, the minimum of J cannot be found analytically due to computational limitations. It can be reached iteratively with a descent algorithm. In this case, the iterative minimization of J is based on a gradient method. J is calculated with the forward observation operator (including the CTM) and its gradient relative to the control parameters **x is provided by the adjoint of the observation operator (including the adjoint of the CTM). The gradient is defined as:**

$$\nabla J(\boldsymbol{x}) = \boldsymbol{B}^{-1}\left(\boldsymbol{x} - \boldsymbol{x^b}\right) + H^{*}\boldsymbol{R}^{-1}(H(\boldsymbol{x}) - \boldsymbol{y}) \textbf{ (Eq. 2)}$$

where H^* is the adjoint of the observation operator.

The high non-linearity of the chemistry for reactive species makes it difficult to use its tangent-linear to approximate the actual observation operator, and, more generally, it makes the inversion problem highly **non-linear**. Therefore, in PYVAR-CHIMERE, we use the M1QN3 limited memory quasi-Newton minimization algorithm [Gilbert and Lemaréchal, 1989], which relies on the actual CHIMERE non-linear model to compute J at each iteration of the minimization. As most quasi-Newton methods, it requires an initial regularization of **x**, the vector to be optimized, for better efficiency. We adopt the most generally used regularization, made by minimizing in the space defined by:

$$\boldsymbol{\chi} = \boldsymbol{B}^{\frac{1}{2}}\left(\boldsymbol{x} - \boldsymbol{x^b}\right) \textbf{ (Eq. 3)}$$

instead of the control space defined by **x**. Although more advanced regularizations can be chosen, the minimization with χ is preferred for its simplifying the equation to solve. In the χ-space, Equation 2 can be re-written as follows:

$$\nabla J\boldsymbol{\chi} = \boldsymbol{\chi} + \boldsymbol{B}^{\frac{1}{2}}H^{*}\left(\boldsymbol{R}^{-1}(H(\boldsymbol{x}) - \boldsymbol{y})\right) \textbf{ (Eq. 4)}$$

The criterion for stopping the algorithm is based on a threshold set on the ratio between the final and initial gradient norms or on the maximum number of iterations to perform. **As shown in Figure 1, the minimization algorithm repeats the forward-adjoint cycle to get an estimate close to the optimal solution of the inversion problem for the control parameters. This approximation of the optimal estimate is found by satisfying the convergence criteria of the minimizer with a given reduction of the norm of the gradient of *J*. Nevertheless, due to the non-linearity of the problem, the minimization may reach a local minimum only, instead of the global minimum.**

Finally, the calculation of the uncertainty in the estimate of emissions from the inversion, known as "posterior uncertainty", is challenging in a variational inverse system **[Rayner et al., 2019]**. Even though the posterior uncertainty can be explicitly written in various analytical forms, it requires the inversion of matrices that are too large to invert given the current computational resources in our variational approach. As a trade-off between computing resources and comprehensiveness, the analysis error may be evaluated by an approach based on a propagation of errors through sensitivity tests (e.g., as in Fortems-Cheiney et al., [2012]). It can also be estimated through a Monte Carlo Ensemble [Chevallier et al., 2007], implemented in PYVAR. **Nevertheless, it should be noted that the cost of the Monte Carlo experiments used to derive these posterior uncertainties is huge.**

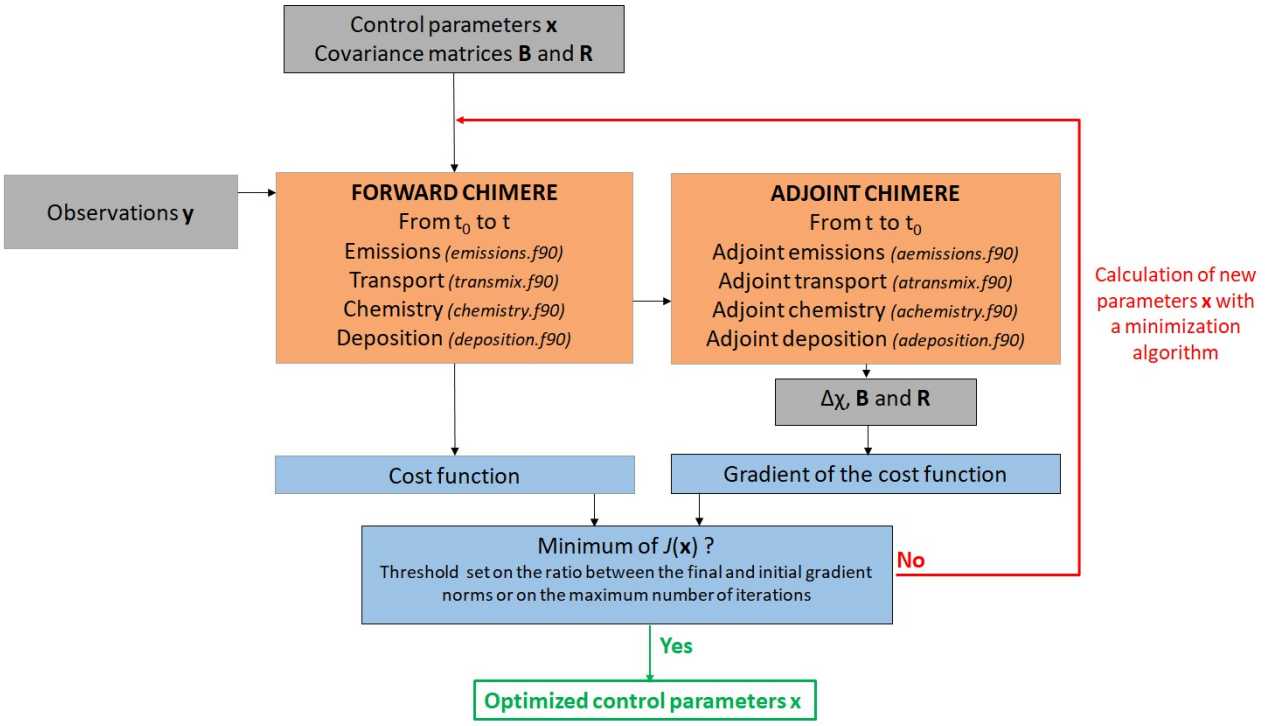


Figure 1. *Simplified scheme of the iterative minimization in PYVAR-CHIMERE. PYVAR, CHIMERE and text sources are displayed in blue, in orange and in grey, respectively.*

3. The PYVAR-CHIMERE configuration

3.1. PYVAR adapted to CHIMERE

The PYVAR-CHIMERE inverse modeling system is based on the Bayesian variational assimilation code PYVAR [Chevallier et al. 2005] and on a previous inversion system coupled to CHIMERE [Pison et al., 2007]. PYVAR is an ensemble of Python scripts, which deals with preparing the vectors and the matrices for the inversion, drives the required Fortran codes of the transport model and computes the minimization of the cost function to solve the inversion. Previously used for global inversions with the LMDz model [e.g., Pison et al., 2009; Chevallier et al., 2010; Fortems-Cheiney et al., 2011; Yin et al., 2015; Locatelli et al., 2015; Zheng et al., 2019], PYVAR has been

adapted to CHIMERE with an adjoint code without chemistry by Broquet et al. [2011]. In order to couple PYVAR to the new state-of-the-art version of CHIMERE (see Section 3.2), to include chemistry, and to increase its modularity, flexibility and clarity, the new system described here has been developed. It includes elements of the inversion system (coded in Fortran90) of Pison et al. [2007].

3.2. Development and parallelization of the adjoint and tangent-linear codes of CHIMERE

To compute the sensitivity of simulated atmospheric concentrations to corrections to the fluxes, the adjoint of CHIMERE has been developed. Originally, the sequential adjoint was coded [Menut et al., 2000; Menut et al., 2003; Pison et al., 2007]. The adjoint has been coded by hand line by line, following the principles formulated by Talagrand [1997]. It contains exactly the same processes as the CHIMERE forward model. **The code has been parallelized, which** required a redesigning of the **entire** code, associated with a full testing scheme **(see Section 3.3)**. Furthermore, the tangent-linear (TL) code has been developed and validated **(see Section 3.3)**. Changes have been implemented in the forward CHIMERE code embedded in PYVAR-CHIMERE to match requirements of the studies **conducted with this system**. These changes have been implemented in both the adjoint and the TL codes. **Compared to the CHIMERE 2013 version [Menut et al., 2013], the most important of these changes are, regarding geometry, the possibility of polar domains and the use of the coordinates of the corners of the cells instead of only the centers, allowing the use of irregular grids. Regarding transport, the non-uniform Van Leer transport scheme on the horizontal has been implemented, which is consistent with the use of irregular grids. Finally, various switches have been added to keep the system consistent for GHG studies. For example, we can avoid going into the chemistry, deposition or wet deposition routines when the focused species do not require them (e.g., no chemistry for methane or carbon dioxide at a regional scale).**

PYVAR-CHIMERE is currently **implemented with** a full module of gaseous chemistry. As a compromise between the robustness of the method for reactive species, the time required coding the adjoint and the computational cost with a full chemical scheme, the aerosols modules of CHIMERE have not been included in the adjoint of CHIMERE yet and are therefore not available in PYVAR-CHIMERE. The development and maintenance of the adjoint means that the version used is necessarily one or two versions behind the distributed CHIMERE version (http://www.lmd.polytechnique.fr/chimere/). It should also be noted that PYVAR-CHIMERE only

infers anthropogenic emissions at this stage. The optimization of biogenic emissions, which are linearly interpolated at the sub-hourly scale in CHIMERE, is currently under development.

As an example, Figure 2 presents a simplified scheme of how PYVAR scripts are used to drive this version of CHIMERE for forward simulations and inversions using satellite observations.

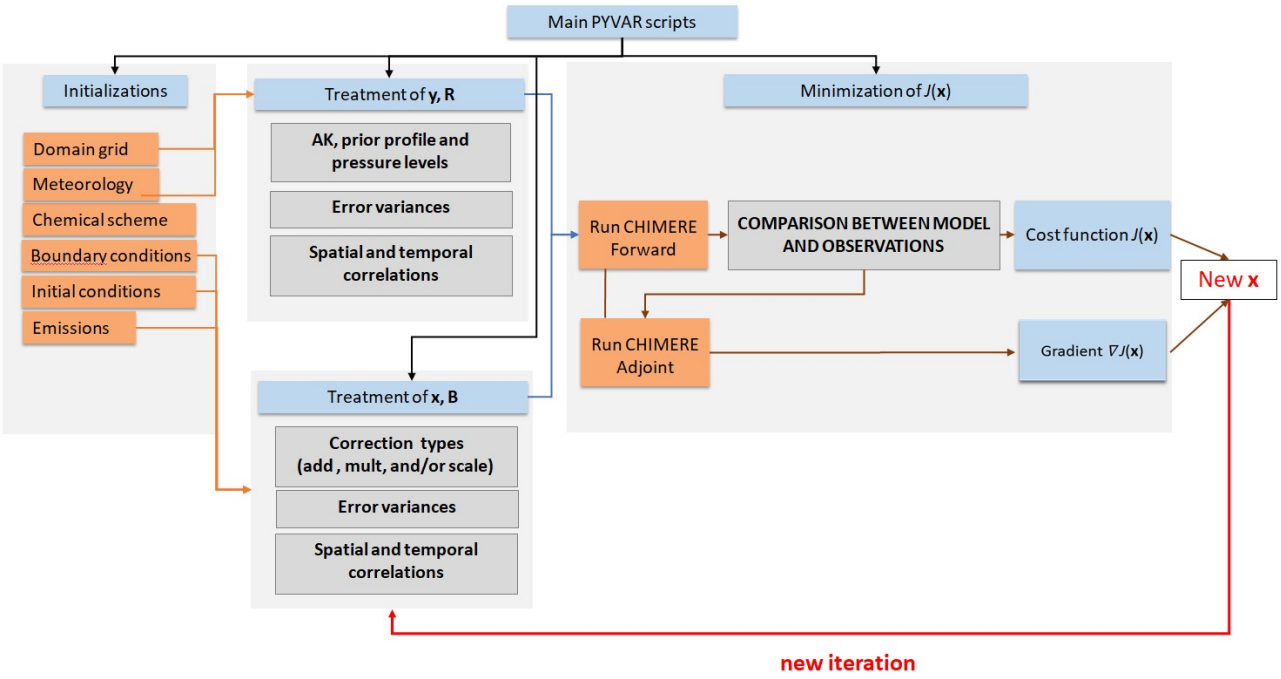


Figure 2. *Simplified scheme of how PYVAR scripts are used to drive CHIMERE for an inversion using satellite observations. PYVAR, CHIMERE and text sources are displayed in blue, in orange and in grey, respectively. "AK" refers to Averaging Kernels as detailed in Section 3.5.*

3.3. Accuracy of tangent-linear and adjoint codes

Different procedures have been implemented to test the accuracy of the TL and adjoint codes. To test the linearity of the TL, we compute a Taylor diagnostic. It consists in computing the TL at x_0 for given increments Δx, $dHx_0(\Delta x)$, then the TL at x_0 for $\lambda \times \Delta x$ with λ an arbitrary small number, $dHx_0(\lambda \Delta x)$. Theoretically, if the TL is well coded, $\lambda dHx_0(\Delta x) = dHx_0(\lambda \Delta x)$ by definition. In practice, the difference must be lower than 10 times the precision of the machine on which it is run.

The adjoint code is also tested, by verifying that $\langle H.\Delta x, H.\Delta x \rangle = \langle \Delta x, H^T H.\Delta x \rangle$ where H^T stands for the adjoint at x. What is actually computed is the ratio of the difference between the two scalar products to the second one and the accuracy of the computation. The difference should be a few times the precision of the machine on which it is run.

3.4. Definition of the control vector

The control vector is specified by the user in a text file. This file is formatted following Table 1. The parameters to **be inverted may** be fluxes and/or initial conditions and/or boundary concentration conditions, at the grid-cell resolution or for one region encompassing up to the whole domain. Several types of corrections can be applied, they are defined in the code as "add", "mult" or "scale". Both the corrections "add" and "mult" are applied to gridded control variables. For correction type “add”, the control variables are increments added to the corresponding components of the model inputs. For correction type “mult", the control variables are scaling factors multiplying the corresponding components of the model inputs. The difference between the two options "add" and "mult" plays a role when inverting fluxes which can switch from positive to negative values (like CO_2 natural fluxes). **For type “scale”, the control variables are scaling factors applied to maps different from the maps of emissions used as prior input of the forward model: for example, activity maps can be used and scaled to get emissions; the obtained values are then added to the corresponding components of the model inputs. With these various types, it is possible to define the control variables as the budgets of emissions for different regions, types of activities, and/or processes, which can thus be directly rescaled by the inversions, similarly to what is done in systems where the control vector is not gridded [Wang et al., 2018]).**

Different simple but efficient ways of building the error covariance matrix **B** are implemented in PYVAR-CHIMERE. The variances and correlations are defined independently. The variances are specified by the user through standard deviation coefficient (Table 1), which can be a fixed value ("fx") or a percentage ("pc") to define the diagonal standard deviation matrix $\sum$. For correction types "mult" and "scale", as well as for correction type "add" with a fixed value, the value is directly used as the standard deviation of the uncertainty in the corresponding components of the control vector. For correction type "add" with a percentage provided, maps of standard deviation of uncertainty are built by applying this percentage to the matching input fields (fluxes, initial conditions, boundary conditions). The user may also provide a script to build personalized maps of variances.

Potential correlations between uncertainties in different types of control variables, e.g. between fluxes and boundary conditions, and correlations between uncertainties in different species, e.g. between fluxes of CO and NO_x, are not coded yet. Only correlations for a given type of control variable and a given species are so far taken into account so that the **B** matrix is block diagonal. For a given type of control variable and a given species (in the illustration in section 4.2.2: CO, NO or NO_2 fluxes), spatial and temporal correlations can be defined using correlation lengths through time Lt and space Ls. Those lengths are used to model temporal and/or spatial auto-correlations using an

exponentially decaying function: the correlation r between parameters and at a given location but separated by duration $d(x_i, x_j)$, or at a given time but distant by $d(x_i, x_j)$ is given by $r(x_i, x_j) = exp\left(\frac{-d(x_i,x_j)}{L}\right)$ (***Eq.* 5**) where $L = L_T \vee L_S$ is the corresponding correlation length. There is no correlation between uncertainties in land and ocean flux. Note that the spatial correlations are computed for each vertical level independently when dealing with control variables with vertical resolution (3D fields of fluxes when accounting for emission injection heights, or boundary/initial conditions).Vertical correlations in the uncertainties in such variables have not been coded yet. Apart from this, the system assumes that temporal correlations and spatial correlations depend on the time lag and distance but not on the specific time and location of the corresponding parameters. It also assumes that the correlation between uncertainties at different locations and different time can be derived from the product of the corresponding autocorrelation in time and space.

Each block of **B** can thus be decomposed based on Kronecker products: $\mathbf{B}=\sum C_t \otimes C_s \sum$ (**Eq. 6**) where $\otimes$ is the Kronecker product, C_t and C_s are the temporal and spatial correlations, respectively. The calculations involving $\mathbf{B}^{1/2}$ **(in Eq. 3, Eq. 4)** are simplified in PYVAR-CHIMERE using the Eigen-decomposition of C_t and C_s. Its square root can be calculated according to: $C_t^{1/2} = V_{Ct} D_{Ct}^{1/2} V_{Ct}^T$ (**Eq. 7**) (and similarly for C_s), where V_{Ct} is the matrix with the Eigenvectors as columns, and D_{Ct} is the diagonal matrix of Eigenvalues of C_t. It is possible to chose a threshold under which the eigenvalues are truncated when computing the spatial correlations in order to save computation **time** and memory, but not when computing the temporal correlations.

Constrained species	Correction type : - Add - Mult - Scale	Spatial resolution - at the grid-cell resolution - for one region	Temporal resolution (in hours)	Input to constrain: -Fluxes -Initial conditions -Lateral Boundary conditions -Top Boundary conditions	B variance coefficient: -fx -pc	Decorrelation time (in hours)	Decorrelation length on land (in km)	Decorrelation length on sea (in km)
CO	add	0.5°x0.5°	168	Fluxes	100 %	-	-	-
CO	add	0.5°x0.5°	1	Initial conditions	15%	-	-	-
CO	add	0.5°x0.5°	168	Lateral Boundary conditions	15%	-	-	-
CO	add	0.5°x0.5°	168	Top Boundary conditions	15%	-	-	-

NO	add	0.5°x0.5°	24	Fluxes	50 %	-	50	50
NO	add	0.5°x0.5°	1	Initial conditions	15%	-	-	-
NO_2	add	0.5°x0.5°	24	Fluxes	50 %	-	50	50
NO_2	add	0.5°x0.5°	24	Initial conditions	15%	-	-	-

Table 1. ***Examples for the definition of the control vector and for the construction of the B matrix, as illustrated in Section* 4**.

3.5. Equivalents of the observations

During forward simulations, the equivalents of the components of **y** (i.e, the equivalents of the individual data) are calculated by PYVAR-CHIMERE. It includes the CTM and an interpolation (see below the vertical interpolation from the model's grid to the satellite levels) or an extraction and averaging (e.g. extracting the grid cell matching the geographical coordinates of a surface station and averaging over one hour). As a compromise between technical issues such as the time required for reading/writing files, the observation operator *H* that generates the equivalent of the observations by the model (i.e. $H(\mathbf{x})$) has been so far partly embedded in the code of CHIMERE. It makes it easier to use finer time intervals than available in the usual hourly outputs of CHIMERE to compute the required information (e.g., within the finer CTM physical time steps).

To make comparisons between simulations and satellite observations, the simulated vertical profiles are first interpolated on the satellite's levels (with a vertical interpolation on pressure levels) **in CHIMERE**. Then, the averaging kernels (AKs), when available, are applied to represent the vertical sensitivity of the satellite retrieval. Two types of formula, depending on the satellite observations used, have been detailed in PYVAR-CHIMERE for the use of AKs: $C_m = AK.C_{m(o)}$ (**Eq. 8**) or $C_m = x_a + AK\left(c_{m(o)} - x_a\right)$ (**Eq. 9**) where C_m is the modeled column, AK contains the averaging kernels, x_a is the prior profile (provided together with the AKs when relevant) and $C_{m(o)}$ is the vertical distribution of the original model partial columns interpolated to the pressure grid of the AKs.

3.6. Numerical language

The PYVAR code is in Python 2.7, the CHIMERE CTM is coded in Fortran90. The CTM requires several numerical tools, compilers and libraries. The PYVAR-CHIMERE system was developed and tested using the software versions as described in Table 2.

		URL	Version
Software	Python	https://www.python.org/downloads/	2.7
	Fortran compiler ifort	https://software.intel.com/en-us/fortran-compilers	Composer-xe-2013.2.146
Libraries or packages	UnidataNetCDF	https://www.unidata.ucar.edu/	3
	Open MPI	https://www.open-mpi.org/	1.10.5
	GRIB_API	https://confluence.ecmwf.int/display/GRIB/Releases	1.14
	nco	http://nco.sourceforge.net/#Source	4.6.3

Table 2. *URL addresses for the development and the use of the PYVAR-CHIMERE system and its modules.*

PYVAR-CHIMERE's computation time for one node of 10 CPUs is about 4h for 1 day of inversion (with ~10 iterations) for the European domain size of 101 (longitude) x 85 (latitude) x 17 (vertical levels) used in Section 4. The model **parallelization** results from a Cartesian division of the main geographical domain into several sub-domains, each one being processed by a worker process. To configure the parallel sub-domains, the user has to specify two parameters in the model parameter file: the number of sub-domains for the zonal and meridian directions. The total number of CPUs used is therefore the product of these two numbers plus one for the master process.

4. Potential of PYVAR-CHIMERE for the inversion of CO and NO_x emissions

The potential of the PYVAR-CHIMERE system to invert emissions of reactive species is illustrated with the inversion of CO and NO_x anthropogenic emissions in Europe respectively based on MOPITT CO data and OMI NO_2 data. We have chosen to present an illustration of CO inversion over a 7-day window, the first week of March 2015. Considering the short lifetime of NO_x of a few hours [Valin et al., 2013; Liu et al., 2016], we have chosen to present illustration of NO_x inversion over a 1-day window, 19 February 2015. These particular periods have been chosen as they present **a representative** number of super-observations during winter, and as the emissions are high during that period. All the information required by the system to invert CO and NO_x emissions is listed in Table 1.

4.1. Data and model description

4.1.1. Observations y

We use CO data from the MOPITT instrument [Deeter et al., 2019]. MOPITT has been flown onboard the NASA EOS-Terra satellite, on a low sun-synchronous orbit that crosses the equator at 10:30 and 22:30 LST. The spatial resolution of its observations is about 22x22 km^2 at nadir. It has been operated nearly continuously since March 2000. MOPITT CO products are available in three

variants: thermal-infrared TIR only, near-infrared NIR only and the multispectral TIR-NIR product, all containing total columns and retrieved profiles (expressed on a ten-level grid from the surface to 100 hPa). We choose to constrain CO emissions with the MOPITT surface product for our illustration. Among the different MOPITTv8 products, we choose to work with the multispectral MOPITTv8-NIR-TIR one, as it provides the highest number of observations, with a good evaluation against in situ data from NOAA stations [Deeter et al., 2019]. The MOPITTv8-NIR-TIR surface concentrations are sub-sampled into "super-observations" in order to reduce the effect of errors that are correlated between neighboring observations*:* we selected the median of each subset of MOPITT data within each 0.5°×0.5° grid-cell and each physical time step (about 5-10 minutes). After this screening, 8437 "super-observations" remain in the 7-day inversion (from 10667 raw observations).It is important to note that the potential of MOPITT to provide information at a high temporal resolution, up to the daily scale, is hampered by the cloud coverage (see the blanks in Figure 5b).

The observational constraint on NO_2 emissions comes from the OMI QA4ECV tropospheric columns [Muller et al., 2016; Boersma et al., 2016, Boersma et al., 2017]. The Ozone Monitoring Instrument (OMI), a near-UV/Visible nadir solar backscatter spectrometer, was launched onboard EOS Aura in July 2004. It has been flown on a 705 km sun-synchronous orbit that crosses the Equator at 13:30 LT. Our data selection follows the criteria of the OMI QA4ECV data quality statement. As the spatial resolution of the OMI data is finer than that of the chosen CHIMERE model grid (13x24 km^2 against 0.5°×0.5°, respectively), the OMI tropospheric columns are sub-sampled into "super-observations" (median of the OMI data within the 0.5°×0.5° grid-cell and each physical time step and its corresponding AK).

4.1.2 CHIMERE set-up

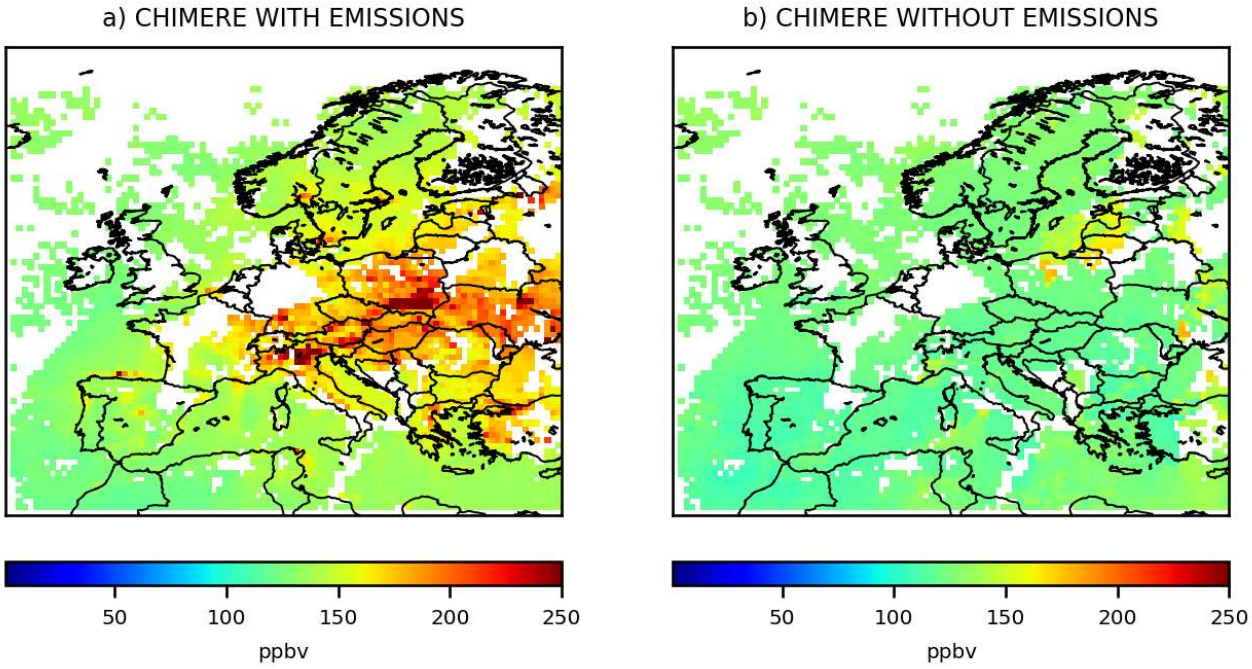
CHIMERE is run over a 0.5°×0.5° regular grid (about 50x50km^2) and 17 vertical layers, from the surface to 200hPa (about 12km), with 8 layers within the first two kilometers. The domain includes 101 (longitude) x 85 (latitude) grid-cells (15.5°W-35°E; 31.5°N-74°N, see Figure 3). CHIMERE is driven by the European Centre for Medium-Range Weather Forecasts (ECMWF) meteorological forecast [Owens and Hewson, 2018]. The chemical scheme used in PYVAR-CHIMERE is MELCHIOR-2, with more than 100 reactions [Lattuati, 1997; CHIMERE 2017], including 24 for inorganic chemistry. The prior anthropogenic emissions for CO and NO_x emissions **are obtained** from the TNO-GCHco-v1inventory [Super et al., 2020], the last update of the TNO-MACCII inventory [Kuenen et al., 2014]. **This inventory is based on the EMEP/Centre on Emission Inventories and Projections (CEIP) official country reporting for air pollutants done in 2017.**

It is an inventory at 6kmx6km horizontal resolution. From the annual and national budgets, each sector is assigned to a specific proxy to quantify the spatial variability of the emissions within each country. Temporal profiles are also provided per Gridded Nomenclature For Reporting (GNFR) sector code (variations due to the month, weekday and hour). Following the Generation of European Emission Data for Episodes (GENEMIS) recommendations [Kurtenbach et al., 2001; Aumont et al., 2003], NO_x emissions are speciated as 90% of NO, 9.2% of NO_2, and 0.8% of nitrous acid (HONO). The TNO-GHGco-v1 inventory has been aggregated to the CHIMERE grid.

The prior anthropogenic emissions for VOCs **are obtained** from the EMEP inventory [Vestreng et al., 2005; EMEP/CEIP website]. Biogenic emissions come from the **Model of Emissions of Gases and Aerosols from nature (MEGAN)** [Guenther et al., 2006]. Different climatological values from the LMDZ-INCA global model [Szopa et al., 2008] or from a Monitoring Atmospheric Composition and Climate (MACC) reanalysis are used to prescribe concentrations at the lateral and top boundaries and the initial atmospheric composition in the domain. Full access to and more information about the MACC reanalysis data can be obtained through the MACC-II web site (http://www.copernicus-atmosphere.eu). In order to ensure realistic fields of simulated CO and NO_2 concentrations from the beginning of the inversion period, runs have been preceded with a 10-day spin-up.

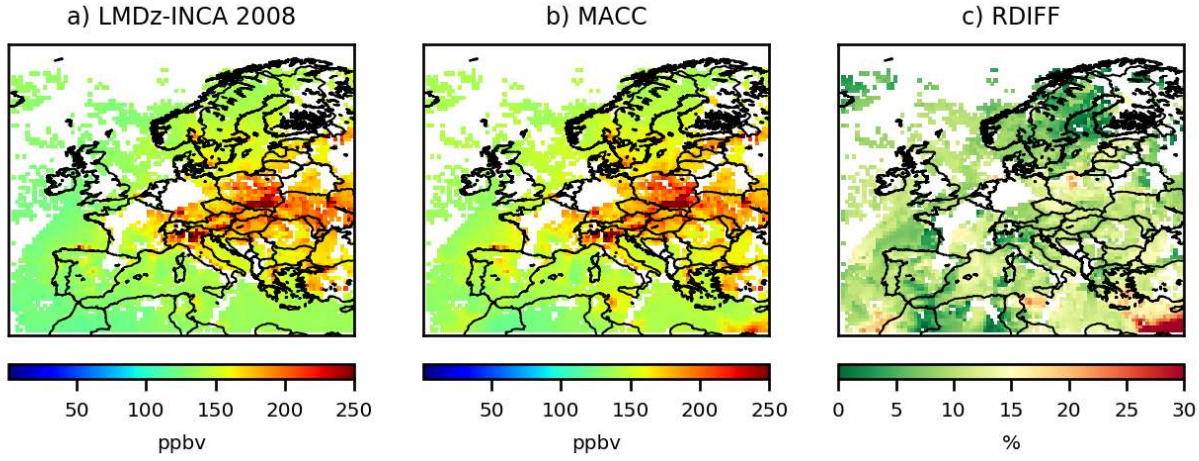
4.1.3. CO Sensitivity to emissions and to initial and boundary conditions

With its lifetime of about two months, CO could be strongly **influenced** by the initial and lateral boundary conditions prescribed in the CTM. In fact, as seen in Figure 4b, initial and boundary conditions provide a relatively flat background and the patterns which appear clearly over the background are linked to surface emissions (Figure 4a). To characterize the uncertainties in the concentration fields due to the initial and lateral boundary conditions, we performed a sensitivity test by using either climatological values from LMDZ-INCA or a MACC reanalysis: **maximum relative differences in concentrations of about 15% over continental land are estimated (Figure 4c)**. **The errors assigned to initial and boundary conditions in Section 4.2.2 are based on this sensitivity test.**



465

Figure 3. ***Mean CO surface concentrations from the 1st to the 7th, March 2015*** *simulated by CHIMERE a) with anthropogenic and biogenic emissions, and b) without emissions, in ppbv, at the 0.5°x0.5° grid-cell resolution.*



466

Figure 4. ***Mean CO surface concentrations from the 1st to the 7th, March 2015*** *simulated by CHIMERE using for initial and boundary conditions, a) the climatological values from the LMDZ-INCA global model b) the climatological values from a MACC reanalysis, in ppbv, and c) the relative differences between these two simulations , in %, at the 0.5°x0.5° grid-cell resolution.*

4.1.4. Comparison between CHIMERE and the observations

467 Large discrepancies are found between the MOPITT CO observations (Figure 5b) and the prior
468 simulation by CHIMERE over Europe (Figure 5a). For the first week of March 2015, CO

concentrations are generally under-estimated by CHIMERE, particularly over Central and Eastern Europe (excepted in the south of Poland). On the contrary, CO concentrations seem to be over-estimated over Spain and Portugal. Large discrepancies are also found between the OMI NO_2 super-observations and the prior simulation by PYVAR-CHIMERE (Figure 6**), as already noticed by Huijnen et al. [2010], with an inter-comparison of NO_2 OMI-DOMINO tropospheric columns with an ensemble of European regional air quality models including CHIMERE**. Over Europe, the prior simulation strongly underestimates the tropospheric columns over industrial areas (e.g., over the Netherlands and over Po Valley). **These discrepancies might be due to different causes, which can all interact. A source of uncertainties is related to the observations. For example, satellite data inter-comparison studies reveal large differences between different retrievals of the same compound [Qu et al., 2020]. It can be explained by uncertainties from the CTM (e.g., through the underestimation of the atmospheric production or the underestimation of the species lifetime). It could also be explained by an underestimation of the anthropogenic emissions in the BU inventory.**

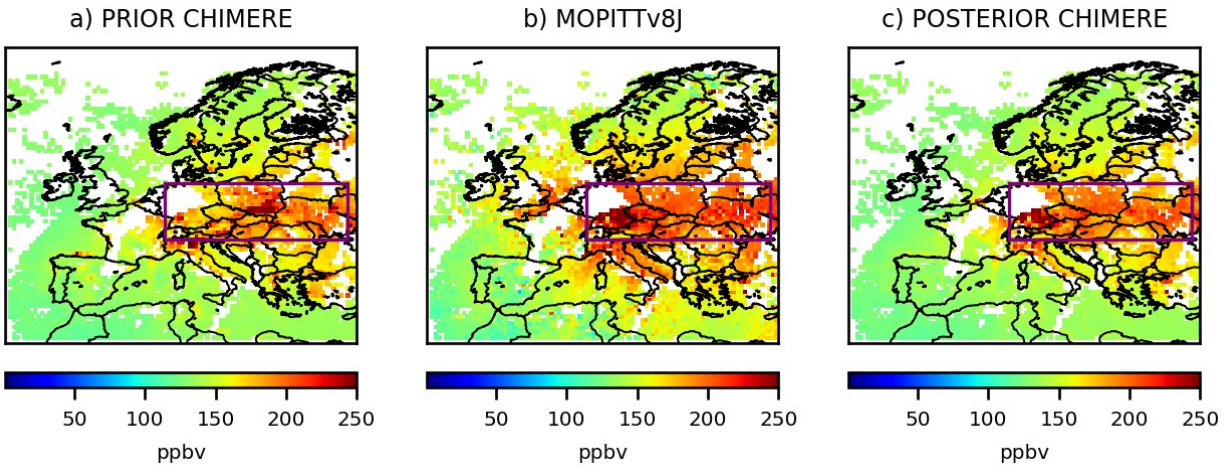


Figure 5. ***Mean CO collocated surface concentrations from the 1st to the 7th, March 2015*** *a) simulated by CHIMERE using the prior TNO-GHGco-v1 emissions and the climatological values from the LMDZ-INCA global model for initial and boundary conditions, b) observed by MOPITTv8-NIR-TIR and c) simulated by CHIMERE using the posterior emissions, in ppbv, at the 0.5°x0.5° grid-cell resolution.* ***Statistics for the comparison between simulations and observations are given in Table 4 for the area in the purple box****.*

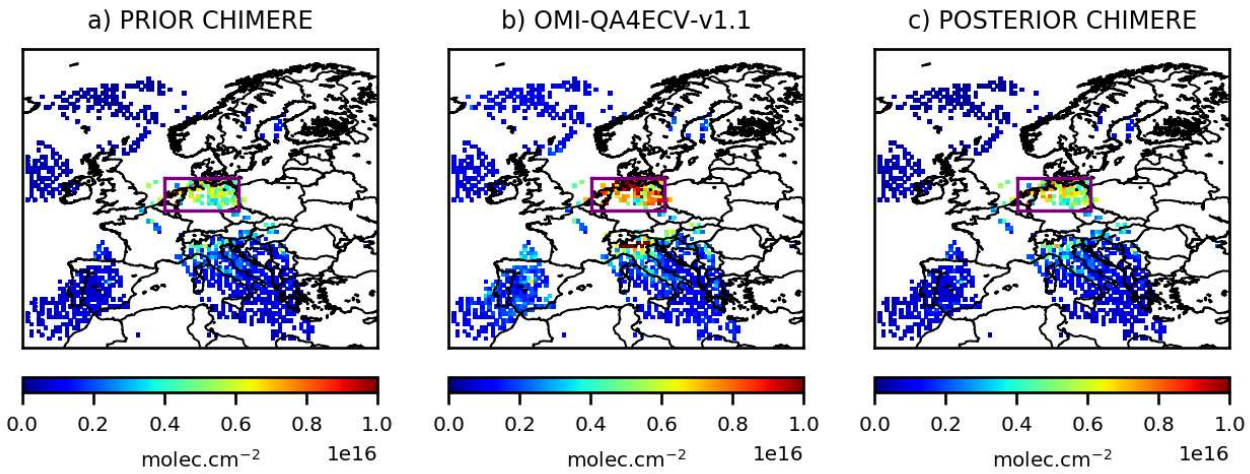


Figure 6. *NO_2 collocated **tropospheric columns** a) simulated by CHIMERE using the prior TNO-GHGco-v1 emissions and the climatological values from the LMDZ-INCA global model for initial and boundary conditions, b) observed by OMI and c) simulated by CHIMERE using the posterior emissions, in 10^{16}molec.cm^{-2}, at the 0.5°x0.5° grid-cell resolution, the 19th, February 2015.* ***Statistics for the comparison between simulations and observations are given in Table 5 for the area in the purple box.***

4.2. Inversions

4.2.1. Control vector x

For the CO inversion, the control vector **x** contains:

- the CO anthropogenic emissions **at a 7-day temporal resolution, a 0.5° ×0.5° (longitude, latitude) horizontal resolution, and 8 vertical levels, i.e. 101×85×8 components in x**,
- the CO 3D initial conditions at a 0.5° ×0.5° (longitude, latitude) resolution × 17 vertical levels,
- the CO lateral and top boundary conditions **at a** 7-day **temporal resolution**, at a 0.5° ×0.5° (longitude, latitude) resolution, i.e. (2x101 + 2x85) and 17 vertical levels.

Considering its short lifetime, there is no boundary conditions for NO_2. For the NO_x inversion, the control vector **x** contains:

- the NO and NO_2 anthropogenic emissions **at a** 1-day **temporal resolution**, at a 0.5° ×0.5° (longitude, latitude) resolution and 8 vertical levels, i.e. 101×85×8 grid cells,
- the NO and NO_2 3D initial conditions at a 0.5° ×0.5° (longitude, latitude) resolution and 17 vertical levels.

4.2.2. Covariance matrices B and R

To our knowledge, there are few available studies dealing with the estimates of the uncertainties in gridded bottom-up emission inventories at the 0.5°x0.5° resolution or higher.

The characterization of their statistics in the inversion configuration is consequently often based on crude assumptions from the inverse modelers. **Consequently, as an example for the NO_x inversion, different sensitivity tests described in Table 3 have been performed for the construction of the B matrix.** For both the prior NO and NO_2 emissions at 1-day and 0.5° resolution, the prior error standard deviations **are first assigned to 50% of the prior estimate of the emissions (test A), as in Souri et al. [2020]. Sensitivity tests have also been performed with prior error standard deviations assigned to 80 and 100% of the prior estimate of the emissions (test C and test D, respectively, Figure 8).**

With prior error standard deviations set at 15% of the initial conditions, the changes in initial conditions are very small (not shown) and do not affect the posterior emissions (test B, Figure 8). As indicated in Section 3.4 and in Table 1, it is possible to use correlations in **B**, as in Broquet et al. [2011], in Broquet et al. [2013] and in Kadygrov et al. [2015]. **We demonstrate the strong impact of spatial correlations, defined by an e-folding length of 50km over land and over the sea, on our inversions results (test E, Figure 8).**

Name of the sensitivity tests	Prior error standard deviations in B		Spatial correlation in B	Number of iterations	Reduction of the norm of the gradient of *J*
	On prior emissions	On prior initial conditions			
A	50%	-	-	4	99%
B	50%	15%	-	6	98%
C	80%	15%	-	7	97%
D	100%	15%	-	6	95%
E	50%	15%	50km	5	92%

Table 3. ***Description of the different sensitivity tests performed for the construction of the B matrix for the NO_x inversion.***

Even though annual CO emissions in Western Europe may be well known, with uncertainties of 6% according to Super et al., [2020], larger uncertainties could affect Eastern Europe. Moreover, large uncertainties still affect bottom-up emission inventories at the 0.5° resolution: spatial disaggregation of the national scale estimates to provide gridded estimates causes a significant increase in the uncertainty for CO [Super et al., 2020]. For the inversion of CO emissions, the error

standard deviations assigned to the prior CO emissions at 7-day and 0.5° resolution are 100%. **This value of 100% has already been chosen in Fortems-Cheiney et al. [2011] and in Fortems-Cheiney et al. [2012].** For this CO illustration, the covariance matrix **B** of the prior errors is defined as diagonal (i.e. only variances in the individual control variables listed in 4.2.1 are taken into account).With such a set-up, in theory, we could **obtain** negative posterior emissions since the inversion system does not impose a constraint of positivity in the results. Nevertheless, **even an uncertainty of** 100% lead**s** to a prior distribution mostly (>80%) on the positive side. The assimilation of data showing an increase above the background (at the edges of the domain; not shown) further drive the inversion towards positive emissions for both CO and NO_x inversions. In practice, our inversion does not lead to negative posterior emissions (Figure 7b). Spatial and temporal correlations in **B** would further limit the probability to get negative emissions locally by smoothing the posterior emissions at a spatial scale at which the "aggregated" prior uncertainty is smaller than 100%. However, a positivity constraint should be implemented in future versions of the system.

Based on the sensitivity test in Figure 4, the errors assigned to the CO lateral boundary conditions and to their initial conditions are set at 15%. As these relative errors are significantly lower than those for the emissions and as variations in the CO surface concentrations are mainly driven by emissions (Figure 3), we assume a small relative influence of the correction of initial and boundary conditions on our results. The variance of the individual observation errors in **R** is defined as the quadratic sum of the measurement error reported in the MOPITT and the OMI data sets, and of the CTM errors (including chemistry and transport errors and representativity errors) set at 20% of the retrieval values. The representativity errors could have been reduced with the choice of a finer CTM resolution (e.g., with a resolution closer to the size of the satellite pixel). Error correlations between the super-observations are neglected, so that the covariance matrix **R** of the observation errors is diagonal.

4.2.3. Inversion of CO emissions

Ten iterations are needed to reduce the norm of the gradient of *J* by 90% with the minimization algorithm M1QN3 and obtain the increments, i.e. the corrections provided by the inversion. The prior CO emissions over Europe for the first week of March 2015 and their increments are shown in Figure 7. As expected from the large differences between the prior surface concentrations (Figure 5a) and the MOPITT observations (Figure 5b), local increments can reach more than +50% (Figure 7b). CO emissions are increased over Central and Eastern Europe, except in the south of Poland. On

the contrary, CO emissions are decreased over Spain and Portugal. The analyzed concentrations are the concentrations simulated by CHIMERE with the posterior fluxes: as expected, the optimization of the fluxes improves the fit of the simulated concentrations to the observations (Figure 5c), particularly over Central and Eastern Europe. Over this area (see the **purple** box in Figure 5), the mean bias between the simulation and the observations has been reduced by about 27% when using the posterior emissions **(mean bias of 11.6 ppbv against 15.9 ppbv with the prior emissions, Table 4). The RMSE and the standard deviation have been reduced by about 50% and the correlation has been strongly improved (0.74 when using the posterior emissions against 0.02).**

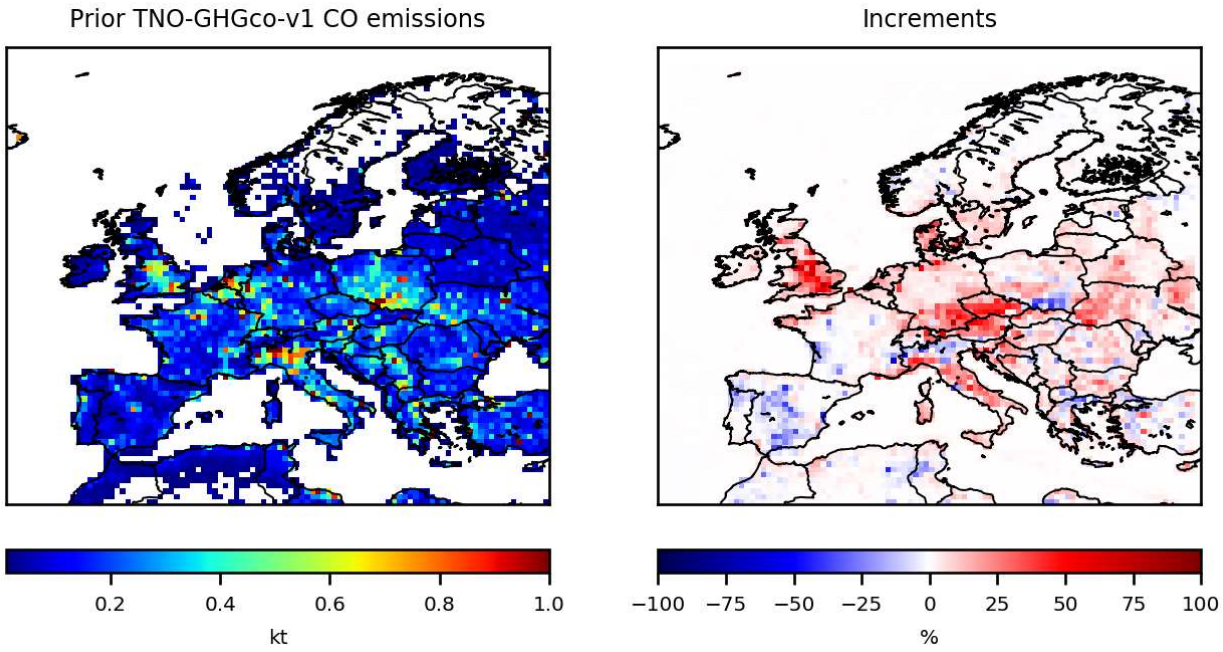


Figure 7. *a) TNO-GHGco-v1 CO anthropogenic prior emissions, in ktCO/grid-cell and b) increments provided by the inversion with constraints from MOPITTv8-NIR-TIR from the 1st to the 7th, March 2015, in %.*

prior				posterior			
MB	RMSE	STD	r	MB	RMSE	STD	r
15.88	41.95	38.82	0.02 (p value = 0.99)	11.58	21.14	17.69	0.74 (p value = 2.08×10^{-11}

Table 4. ***Statistics for the comparison between simulated and observed CO surface concentrations over Central and Eastern Europe (see the area in purple in Figure 5). MB= Mean Bias, RMSE= Root Mean Square Error, STD= Standard Deviation are in ppbv. The spatial correlations r are presented with their p value.***

4.2.4. Inversion of NO_x emissions

The prior NO_x emissions and the corrections provided by the **different sensitivity tests of Table 3 are shown in Figure 8. Here, we analyzed the results from inversion E.** As expected from the underestimation of the prior tropospheric columns in Figure 6, local increments may be large, for example over industrial areas (e.g., over the Po Valley) and over the Netherlands, with increments of more than +50% (Figure 8b). The analyzed NO_2 tropospheric columns in Figure 6c are the columns simulated by CHIMERE with the NO_2 posterior fluxes: as expected, the optimization of the fluxes improves the fit of the simulated concentrations to the observations over the Netherlands. Over this area (see the **purple** box in Figure 6), the mean bias between the simulation and the observations has been reduced by about 24% when using the posterior emissions (mean bias of 1.9**x10^{15}** molec.cm^{-2} **against 2.6x10^{15} molec.cm^{-2} with the prior emissions**, **Table 5**). **The RMSE and the standard deviation have been reduced by about 7%. The correlation has not been improved**. **Even with high emission increments, the impact on the tropospheric columns is rather small.** The posterior emissions and their uncertainties will have to be evaluated and may bring hints to the cause of the discrepancies **between simulated and observed NO_2 tropospheric columns. The biases between OMI and simulated NO_2 tropospheric columns is a complex topic that is not related to our CHIMERE simulations only [Huijnen et al., 2010; Souri et al., 2020; Elguindi et al., 2020]. It requires a fully comprehensive scientific study which is out of the scope of this paper.**

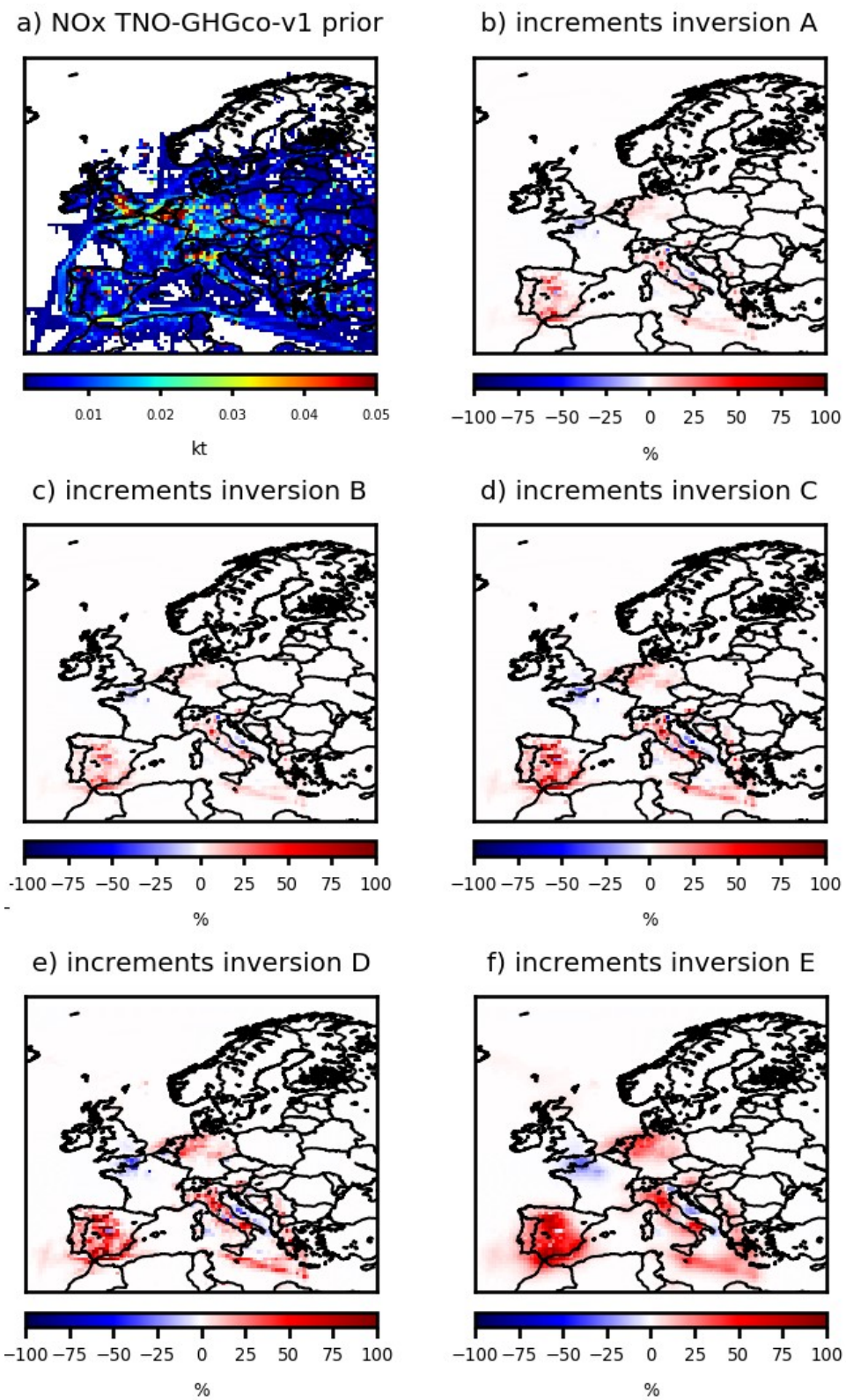


Figure 8. ***a) TNO-GHGco-v1 NO_x anthropogenic prior emissions, in $ktNO_2$/grid-cell and increments provided by the inversion b) A, c) B, d) C, e) D and f) E with constraints from OMI the 19th, February 2015, in %. The description of the different inversions is given in Table 3.***

	prior				posterior			
	MB	RMSE	STD	r	MB	RMSE	STD	r
NO_2	$2.6x10^{15}$	$4.0x10^{15}$	$3.0x10^{15}$	0.008 (p=0.96)	$1.9x10^{15}$	$3.74x10^{15}$	$2.9x10^{15}$	0.01 (p=0.91)

Table 5. ***Statistics for the comparison between simulated and observed NO_2 tropospheric columns for the inversion E, mainly over the Netherlands (see the area in purple in Figure 6). MB= Mean Bias, RMSE= Root Mean Square Error, STD= Standard Deviation are in molecm.cm^{-2}. The spatial correlations r are presented with their p value.***

5. Conclusion/Discussion

This paper presents the Bayesian variational inverse system PYVAR-CHIMERE, which has been adapted to the inversion of reactive species such as CO and NO_x, taking advantage of the previous developments for long-lived species such as CO_2 [Broquet et al., 2011] and CH_4 [Pison et al., 2018]. We show the potential of PYVAR-CHIMERE, with inversions for CO and NO_x illustrated over Europe. PYVAR-CHIMERE will now be used to infer CO and NO_x emissions over long periods, e.g. first for a whole season or year and then for the recent decade 2005-2015 in the framework of the H2020 VERIFY project over Europe, and in the framework of the ANR PolEASIA **project** over China, to quantify their trend and their spatio-temporal variability.

The PYVAR-CHIMERE system can handle any large number of both control parameters and observations. It will be able to cope with the dramatic increase in the number of data in the near future with, for example, the high-resolution imaging (pixel of 7x3.5 km^2) of the new Sentinel-5P/TROPOMI program, launched in October 2017. These new space missions with high-resolution imaging have indeed the ambition to monitor atmospheric chemical composition for the quantification of anthropogenic emissions. Moreover, a step forward in the joint assimilation of co-emitted pollutants will be possible with the PYVAR-CHIMERE system and the availability of TROPOMI co-localized images of CO and NO_2. This should improve the consistency of the inversion results and can be used to inform inventory compilers, and subsequently improve emission inventories. Moreover, this development will help in further understanding air quality problems and addressing air quality related emissions at the national to subnational scales.

Author Contribution

All authors have contributed to the manuscript writing (main authors: AFC, GB, IP and GD) and to the development of the present version of the PYVAR-CHIMERE system (main developer: IP). IP and GD have parallelized the adjoint version from Menut et al., [2000], Menut et al., [2003] and

Pison et al., [2007]. IP has complemented the adjoint of new parameterizations since the CHIMERE release in 2011 and the tangent-linear model.

Code and Data Availability

OMI QA4ECV NO_2 product can be found here: http://temis.nl/qa4ecv/no2.html.
MOPITTv8-NIR-TIR CO product can be found here: ftp://l5ftl01.larc.nasa.gov/MOPITT/
The CHIMERE code is available here: www.lmd.polytechnique.fr/chimere/.

The associated documentation of PYVAR-CHIMERE is available on the website https://pyvar.lsce.ipsl.fr/doku.php/3chimere:headpage. The documentation includes a whole description of PYVAR-CHIMERE and several tutorials on how to run a first PYVAR-CHIMERE simulation or how to run an inversion.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

We acknowledge L. Menut and C. Schmechtig for their contributions to the development work on the adjoint code of CHIMERE and its parallelization. We acknowledge the TNO team (H.A. Denier van der Gon, J. Kuenen, S. Dellaert, S.Jonkers, A. Visschedijk, et al.) for providing NO_x and CO emissions over Europe.We also acknowledge the free use of tropospheric NO_2 column data from the OMI sensor from http://temis.nl/qa4ecv/no2.htmland the free use of CO surface concentrations from the MOPITT sensor from ftp://l5ftl01.larc.nasa.gov/MOPITT/. For this study, A. Fortems-Cheiney was funded by the French Space Agency-Centre National d'Etudes Spatiales CNES and by the H2020 VERIFY project, funded by the European Commission Horizon 2020 research and innovation programme, under agreement number 776810. L. Costantino was funded by the PolEASIA ANR project under the allocation ANR-15-CE04-0005. This work was granted access to the HPC resources of TGCC under the allocations A0050107232 and A0070102201 made by GENCI. Finally, we wish to thank F. Marabelle (LSCE) and his team for computer support.

References

Aumont, B., Chervier, F., and Laval, S.: Contribution of HONO sources to the $NO_x/HO_x/O_3$ chemistry in the polluted boundary layer.Atmospheric Environment, 37(4):487 – 498, 2003.

Belikov, D. A., Maksyutov, S., Yaremchuk, A., Ganshin, A., Kaminski, T., Blessing, S., Sasakawa, M., Gomez-Pelaez, A. J., and Starchenko, A.: Adjoint of the global Eulerian–Lagrangian coupled atmospheric transport model (A-GELCA v1.0): development and validation, Geosci. Model Dev., 9, 749-764, https://doi.org/10.5194/gmd-9-749-2016, 2016.

Boersma, K. F., Vinken, G. C. M., and Eskes, H. J.: Representativeness errors in comparing chemistry transport and chemistry climate models with satellite UV–Vis tropospheric column retrievals, Geosci. Model Dev., 9, 875-898, https://doi.org/10.5194/gmd-9-875-2016, 2016.

Boersma, K. F., Eskes, H., Richter, A., De Smedt, I., Lorente, A., Beirle, S., Van Geffen, J., Peters, E., Van Roozendael, M. and Wagner, T.: QA4ECV NO_2 tropospheric and stratospheric vertical column data from OMI (Version 1.1) [Data set], Royal Netherlands Meteorological Institute (KNMI), http://doi.org/10.21944/qa4ecv-no2-omi-v1.1, 2017.

Bousquet, P., P. Ciais, P. Peylin, M. Ramonet, and P. Monfray: Inverse modeling of annual atmospheric CO_2 sources and sinks: 1. Method and control inversion, *J. Geophys. Res*., 104(D21), 26,161 – 26,178, doi:10.1029/1999JD900342, 1999.

Broquet, G., Chevallier, F., Rayner, P., Aulagnier, C., Pison, I., Ramonet, M., Schmidt, M., Vermeulen, A. T., and Ciais, P.: A European summertime CO_2 biogenic flux inversion at mesoscale from continuous in situ mixing ratio measurements, *J. Geophys. Res*., 116, D23303, doi: 10.1029/2011JD016202, 2011.

Broquet, G., Chevallier, F., Bréon, F.-M., Kadygrov, N., Alemanno, M., Apadula, F., Hammer, S., Haszpra, L., Meinhardt, F., Morguí, J. A., Necki, J., Piacentino, S., Ramonet, M., Schmidt, M., Thompson, R. L., Vermeulen, A. T., Yver, C., and Ciais, P.: Regional inversion of CO_2 ecosystem fluxes from atmospheric measurements: reliability of the uncertainty estimates, Atmos. Chem. Phys., 13, 9039–9056, https://doi.org/10.5194/acp-13-9039-2013, 2013.

Chevallier, F., M. Fisher, P. Peylin, S. Serrar, P. Bousquet, F.-M. Bréon, A. Chédin, and P. Ciais: Inferring CO2 sources and sinks from satellite observations: method and application to TOVS data, *J. Geophys. Res*., 110, D24309, doi:10.1029/2005JD006390, 2005.

Chevallier, F., F.-M. Bréon, and P. J. Rayner: The contribution of the Orbiting Carbon Observatory to the estimation of CO_2 sources and sinks: Theoretical study in a variational data assimilation framework. J. Geophys. Res., 112, D09307, doi:10.1029/2006JD007375, 2007.

Chevallier, F., Ciais, P., Conway, T. J., Aalto, T., Anderson, B. E., Bousquet, P., Brunke, E. G., Ciattaglia, L., Esaki, Y., Fröhlich, M., Gomez, A., Gomez-Pelaez, A. J., Haszpra, L., Krummel, P. B., Langenfelds, R. L., Leuenberger, M., Machida, T., Maignan, F., Matsueda, H., Morguí, J. A., Mukai, H., Nakazawa, T., Peylin, P., Ramonet, M., Rivier, L., Sawa, Y., Schmidt, M., Steele, L. P., Vay, S. A., Vermeulen, A. T., Wofsy, S., and Worthy, D.: CO_2 surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements, J. Geophys. Res., 115, 1–17, https://doi.org/10.1029/2010jd013887, 2010.

Ciarelli, G., Theobald, M. R., Vivanco, M. G., Beekmann, M., Aas, W., Andersson, C., Bergstrom, R., Manders-Groot, A., Couvidat, F., Mircea, M., Tsyro, S., Fagerli, H., Mar, K., Raffort, V., Roustan, Y., Pay, M.-T., Schaap, M., Kranenburg, R., Adani, M., Briganti, G., Cappelletti, A., D'Isidoro, M., Cuvelier, C., Cholakian, A., Bessagnet, B., Wind, P., and Colette, A.: Trends of inorganic and organic aerosols and precursor gases in Europe: insights from the EURODELTA multi-model experiment over the 1990-2010 period, Geosci. Model Dev., 12, 4923-4954, https://doi.org/10.5194/gmd-12-4923-2019, 2019.

CHIMERE documentation, https://www.lmd.polytechnique.fr/chimere/docs/CHIMEREdoc2017.pdf, Last update of this documentation: June 8, 2017,2017.

Deeter, M. N., Edwards, D. P., Francis, G. L., Gille, J. C., Mao, D., Martínez-Alonso, S., Worden, H. M., Ziskin, D., and Andreae, M. O.: Radiance-based retrieval bias mitigation for the MOPITT instrument: the version 8 product, Atmos. Meas. Tech., 12, 4561–4580, https://doi.org/10.5194/amt-12-4561-2019, 2019.

Ding, J., Miyazaki, K., van der A, R. J., Mijling, B., Kurokawa, J.-I., Cho, S., Janssens-Maenhout, G., Zhang, Q., Liu, F., and Levelt, P. F.: Intercomparison of NO_x emission inventories over East Asia, *Atmos. Chem. Phys*., 17, 10125-10141, https://doi.org/10.5194/acp-17-10125-2017, 2017.

EEA, Air quality in Europe - 2018 report, 12/2018, https://www.eea.europa.eu/publications/air-quality-in-europe-2018.

Elbern, H., Strunk, A., Schmidt, H., and Talagrand, O.: Emission rate and chemical state estimation by 4-dimensional variational inversion, Atmos. Chem. Phys., 7, 3749-3769, https://doi.org/10.5194/acp-7-3749-2007, 2007.

Elguindi, N., Granier, C., Stavrakou, T., Darras, S., Bauwens, M., Cao, H., et al.: Intercomparison of magnitudes and trends in anthropogenic surface emissions from bottom-up inventories, top-down estimates, and emission scenarios. Earth's Future, 8, e2020EF001520. https://doi.org/10.1029/2020EF001520, 2020.

EMEP/EEA air pollutant emission inventory guidebook, 2016.

EMEP/CEIP,

https://ceip.at/ms/ceip_home1/ceip_home/webdab_emepdatabase/emissions_emepmodels/

de Foy, B., Lu. Z. and Streets, D.G.: Satellite NO_2 retrievals suggest China has exceeded its NOx reduction goals from the twelfth Five-Year Plan, *Nature Scientific Reports*, 6:35912, 2016.

Fortems-Cheiney, A., et al: Ten years of CO emissions as seen from MOPITT, Journal of Geophysical Research, 116, D5, https://doi.org/10.1029/2010JD014416, 2011.

Fortems-Cheiney, A., Chevallier, F., Pison, I., Bousquet, P., Saunois, M., Szopa, S., Cressot, C., Kurosu, T. P., Chance, K., and Fried, A.: The formaldehyde budget as seen by a global-scale multi-constraint and multi-species inversion system, Atmos. Chem. Phys., 12, 6699–6721, https://doi.org/10.5194/acp-12-6699-2012, 2012.

Gilbert, J., and C. Lemaréchal (1989), Some numerical experiments with variable storage quasi Newton algorithms,Math. Program., 45, 407–435.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmos. Chem. Phys., 6, 3181–3210, https://doi.org/10.5194/acp-6-3181-2006, 2006.

Hakami, A., Henze, D. K., Seinfeld, J. H., Singh, K., Sandu, A., Kim, S., Byun, D., and Li, Q.: The adjoint of CMAQ, Environ. Sci. Technol., 41, 7807–7817, https://doi.org/10.1021/es070944p, 2007.

Hein, R., et coll.: An inverse modeling approach to investigate the global atmospheric methane cycle, *Global. Biogeochem.Cycles*, 11, 43-76, 1997.

Henze, D. K., Hakami, A., and Seinfeld, J. H.: Development of the adjoint of GEOS-Chem, Atmos. Chem. Phys., 7, 2413–2433, https://doi.org/10.5194/acp-7-2413-2007, 2007.

Huijnen, V., Eskes, H. J., Poupkou, A., Elbern, H., Boersma, K. F., Foret, G., Sofiev, M., Valdebenito, A., Flemming, J., Stein, O., Gross, A., Robertson, L., D'Isidoro, M.,

Kioutsioukis, I., Friese, E., Amstrup, B., Bergstrom, R., Strunk, A., Vira, J., Zyryanov, D., Maurizi, A., Melas, D., Peuch, V.-H., and Zerefos, C.: Comparison of OMI NO_2 tropospheric columns with an ensemble of global and European regional air quality models, Atmos. Chem. Phys., 10, 3273–3296, https://doi.org/10.5194/acp-10-3273-2010, 2010.

Hooghiemstra, P. B., Krol, M. C., Bergamaschi, P., de Laat, A. T. J., van der Werf, G. R., Novelli, P.C., Deeter, M. N., Aben, I., and Rockmann, T.: Comparing optimized CO emission estimates usingMOPITT or NOAA surface network observations, *J. Geophys. Res.*, 117, D06309,doi:10.1029/2011JD017043, 2012.

Kadygrov, N., Broquet, G., Chevallier, F., Rivier, L., Gerbig, C., and Ciais, P.: On the potential of the ICOS atmospheric CO_2 measurement network for estimating the biogenic CO_2budget of Europe, Atmos. Chem. Phys., 15, 12765–12787, https://doi.org/10.5194/acp-15-12765-2015, 2015.

Konovalov, I. B. et coll.: Inverse modelling of the spatial distribution of NO emissions on a continental scale using satellite data, *Atmos. Chem. Phys.*, 6, 1747-1770, doi:10.5194/acp-6-1747-2006, 2006.

Konovalov, I. B., Beekmann, M., Burrows, J. P., and Richter, A.: Satellite measurement based estimates of decadal changes in European nitrogen oxides emissions, *Atmos. Chem. Phys.*, 8, 2623-2641, doi:10.5194/acp-8-2623-2008, 2008.

Konovalov, I. B., Beekmann, M., Richter, A., Burrows, J. P., and Hilboll, A.: Multi-annual changes of NO_x emissions in megacity regions: nonlinear trend analysis of satellite measurement based estimates, *Atmos. Chem. Phys*., 10, 8481-8498, doi:10.5194/acp-10-8481-2010, 2010.

Koohkan, M. R., Bocquet, M., Roustan, Y., Kim, Y., and Seigneur, C.: Estimation of volatile organic compound emissions for Europe using data assimilation, *Atmos. Chem. Phys*., 13, 5887-5905, https://doi.org/10.5194/acp-13-5887-2013, 2013.

Krol, M. C., Meirink, J. F., Bergamaschi, P., Mak, J. E., Lowe, D., Jöckel, P., Houweling, S., and Röckmann, T.: What can 14CO measurements tell us about OH?, Atmospheric chemistry and physics, 8, 5033–5044, 2008.

Kuenen, J. J. P., Visschedijk, A. J. H., Jozwicka, M., and Denier van der Gon, H. A. C.: TNO-MACC_II emission inventory; a multi-year (2003–2009) consistent high-resolution European emission inventory for air quality modelling, Atmos. Chem. Phys., 14, 10963-10976, https://doi.org/10.5194/acp-14-10963-2014, 2014.

Kurokawa, J., Ohara, T., Morikawa, T., Hanayama, S., Janssens-Maenhout, G., Fukui, T., Kawashima, K., and Akimoto, H.: Emissions of air pollutants and greenhouse gases over Asian regions during 2000–2008: Regional Emission inventory in ASia (REAS) version 2, *Atmos. Chem. Phys*., 13, 11019-11058, doi:10.5194/acp-13-11019-2013, 2013.

Kurtenbach, R., Becker, K., Gomes, J., Kleffmann, J., LŽrzer, J., Spittler, M.,Wiesen, P., Ackermann, R., Geyer, A., and Platt, U.: Investigations of emissions and heterogeneousformation of HONO in a road traffic tunnel.Atmospheric Environment, 35(20):9506 – 9517. 3385Ð3394, 2001.

Lattuati, M., Impact des émissions européennes sur le bilan de l'ozone troposphérique a l'interface de l'europe et de l'atlantique nord : apport de la modélisation lagrangienne et des mesures en altitude, Ph.D. thesis, Université Paris VI, 1997.

Lelieveld, J., Klingmüller, K., Pozzer, A., Pöschl, U., Fnais, M., Daiber, A., Münzel, T.; Cardiovascular disease burden from ambient air pollution in Europe reassessed using novel hazard ratio functions, *European Heart Journal*, , ehz135, https://doi.org/10.1093/eurheartj/ehz135, 2019.

Lin, J.-T., McElroy, M. B., and Boersma, K. F.: Constraint of anthropogenic NO_x emissions in China from different sectors: a new methodology using multiple satellite retrievals, *Atmos. Chem. Phys.*, 10, 63-78, doi:10.5194/acp-10-63-2010, 2010.

Liu, F., Beirle, S., Zhang, Q., Dörner, S., He, K., and Wagner, T.: NO_x lifetimes and emissions of cities and power plants in polluted background estimated by satellite observations, Atmos. Chem. Phys., 16, 5283–5298, https://doi.org/10.5194/acp-16-5283-2016, 2016.

Locatelli, R., Bousquet, P., Saunois, M., Chevallier, F., and Cressot, C.: Sensitivity of the recent methane budget to LMDz sub-grid-scale physical parameterizations, Atmos. Chem. Phys., 15, 9765-9780, https://doi.org/10.5194/acp-15-9765-2015, 2015.

Mailler S., L. Menut, D. Khvorostyanov, M. Valari, F. Couvidat, G. Siour, S. Turquety, R. Briant, P. Tuccella, B. Bessagnet, A. Colette, L. Letinois, and F. Meleux, CHIMERE-2017: from urban to hemispheric chemistry-transport modeling ,*Geosci. Model Dev.*, 10, 2397-2423, https://doi.org/10.5194/gmd-10-2397-2017, 2017.

Menut, L., R. Vautard, M. Beekmann, and C. Honoré: Sensitivity of photochemical pollution using the adjoint of a simplified chemistry-transport model, *J. Geophys. Res.*, 105, 15,379–15,402, 2000.

Menut L.: Adjoint modelling for atmospheric pollution processes sensitivity at regional scale during the ESQUIF IOP2, Journal of Geophysical Research - Atmospheres, 108, D17, https://doi.org/10.1029/2002JD002549, 2003.

Menut, L., Goussebaile, A., Bessagnet, B., Khvorostiyanov, D., and Ung, A.: Impact of realistic hourly emissions profiles on air pollutants concentrations modelled with CHIMERE, Atmospheric Envionment, 49, 233–244, doi:10.1016/j.atmosenv.2011.11.057, 2012.

Menut, L., Bessagnet, B., Khvorostyanov, D., Beekmann, M., Blond, N., Colette, A., Coll, I., Curci, G., Foret, G., Hodzic, A., Mailler, S., Meleux, F., Monge, J.-L., Pison, I., Siour, G.,Turquety, S., Valari, M., Vautard, R., and Vivanco, M. G.: CHIMERE 2013: a model for regional atmospheric composition modelling, *Geosci. Model Dev.*, 6, 981–1028, doi:10.5194/gmd-6-981-2013, 2013.

Menut, L., Bessanet, B., Siour, G., Mailler, S., Pennel, R. and Cholakian, A. : Impactof lockdown measures to combat Covid-19 on air quality over western Europe, Science of The Total Environment, 741,https://doi.org/10.1016/j.scitotenv.2020.140426, 2020.

Mijling, B., and R. J. van der A: Using daily satellite observations to estimate emissions of short-lived air pollutants on a mesoscopic scale, *J. Geophys. Res.*, 117, D17302, doi:10.1029/2012JD017817, 2012.

Mijling, B., et al., Regional nitrogen oxides emission trends in East Asia observed from space, *Atmos. Chem. Phys.*, 3, 12003, 2013.

Miyazaki, K., Eskes, H., Sudo, K., Boersma, K. F., Bowman, K., and Kanaya, Y.: Decadal changes in global surface NO_x emissions from multi-constituent satellite data assimilation, Atmos. Chem. Phys., 17, 807–837, https://doi.org/10.5194/acp-17-807-2017, 2017.

Muller, J.-P., Kharbouche, S., Gobron, N., Scanlon, T., Govaerts, Y., Danne, O., Schultz, J., Lattanzio, A., Peters, E., De Smedt, I., Beirle, S., Lorente, A., Coheur, P. F., George, M., Wagner, T., Hilboll, A., Richter, A., Van Roozendael, M., and Boersma, K. F.: Recommendations (scientific) on best practices for retrievals for Land and Atmosphere ECVs (QA4ECV Deliverable 4.2 version 1.0), 186 pp., available at: http://www.qa4ecv.eu/sites/default/ files/D4.2.pdf (last access: 12 April 2018), 2016.

Owens, R. G. and Hewson, T.: ECMWF Forecast User Guide, Reading, https://doi.org/10.21957/m1cs7h,https://software.ecmwf.int/wiki/display/FUG/Forecast+User+Guide, 2018.

Pétron, G., Granier, C., Khattatov, B., Lamarque, J.F., Yudin, V., Muller, J.F. and Gille, J.: Inverse modeling of carbon monoxide surface emissions using CMDL networks observations, *J. Geophys. Res,* 107, D24, 2002.

Pison, I., Menut, L., and Bergametti, G.: Inverse modeling of surface NOx anthropogenic emission fluxes in the Paris area during the ESQUIF campaign, *J. Geophys. Res. Atmos.*, 112, D24302, doi:10,1029/2007JD008871, 2007.

Pison, I.,Bousquet, P., Chevallier, F., Szopa, S., and Hauglustaine, D.: Multi-species inversion of CH_4, CO and H_2 emissions from surface measurements, *Atmospheric Chemistry and Physics*, 9, 5281-5297, 2009.

Pison, I., Berchet, A., Saunois, M., Bousquet, P., Broquet, G., Conil, S., Delmotte, M., Ganesan, A., Laurent, O., Martin, D., O'Doherty, S., Ramonet, M., Spain, T. G., Vermeulen, A., and Yver Kwok, C.: How a European network may help with estimating methane emissions on the French national scale, Atmos. Chem. Phys., 18, 3779–3798, https://doi.org/10.5194/acp-18-3779-2018, 2018.

Qu, Z., Henze, D. K., Cooper, O. R., and Neu, J. L.: Improving NO_2 and ozone simulations through global NO_x emission inversions, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2020-307, in review, 2020.

Rayner, P. J., Michalak, A. M., and Chevallier, F.: Fundamentals of data assimilation applied to biogeochemistry, Atmos. Chem. Phys., 19, 13911–13932, https://doi.org/10.5194/acp-19-13911-2019, 2019.

Souri, A. H., Nowlan, C. R., González Abad, G., Zhu, L., Blake, D. R., Fried, A., Weinheimer, A. J., Wisthaler, A., Woo, J.-H., Zhang, Q., Chan Miller, C. E., Liu, X., and Chance, K.: An inversion of NO_x and non-methane volatile organic compound (NMVOC) emissions using satellite observations during the KORUS-AQ campaign and implications for surface ozone over East Asia, Atmos. Chem. Phys., 20, 9837–9854, https://doi.org/10.5194/acp-20-9837-2020, 2020.

Stavrakou, T. and J.-F. Müller: Grid-based versus big region approach for inverting CO emissions using Measurement of Pollution in the Troposphere (MOPITT) data, *Journal of Geophysical Research: Atmospheres*,111, D15, 2006.

Stavrakou, T., Muller, J.-F., Boersma, K. F., De Smedt, I., and van der A, R. J.: Assessing the distribution and growthrates of NOx emission sources by inverting a 10-year record of NO_2 satellite columns, *Geophys. Res. Lett.*, 35, 1–5, doi:10.1029/2008GL033521, 2008.

Super, I., Dellaert, S. N. C., Visschedijk, A. J. H., and Denier van der Gon, H. A. C.: Uncertainty analysis of a European high-resolution emission inventory of CO_2 and CO to support inverse modelling and network design, Atmos. Chem. Phys., 20, 1795–1816, https://doi.org/10.5194/acp-20-1795-2020, 2020.

Szopa, S., Foret, G., Menut, L., and Cozic, A.: Impact of large scale circulation on European summer surface ozone: consequences for modeling, *Atmospheric Environment*, 43, 1189–1195,doi:10.1016/j.atmosenv.2008.10.039, 2008.

Talagrand, O. : Assimilation of observations : an introduction, J. Met. Soc., Japan, 75, 191-209, 1997.

Tang, X., ZhuJ.,WangZ.F., WangM., GbaguidiA., LiJ., ShaoM., TangG. Q., and JiD.S.: Inversion of CO emissions over Beijing and its surrounding areas with ensemble Kalman filter, Atmospheric Environment,81, 676-686, 2013.

Valin, L. C., Russell, A. R., and Cohen, R. C.: Variations of OH rad-ical in an urban plume inferred from NO2 column measurements, Geophys. Res. Lett., 40, 1856–1860, doi:10.1002/grl.50267, 2013.

van der A, R. J., Eskes, H. J., Boersma, K. F., van Noije, T. P. C., van Roozendael, M., De Smedt, I., Peters, D. H. M. U., and Meijer, E. W.: Trends, seasonal variability and dominant NOx source derived from a ten year record of NO_2 measured from space, *J. Geophys. Res*., 113, 1–12, doi:10.1029/2007JD009021, 2008.

Vestreng, V., Breivik, K., Adams, M., Wagner, A., Goodwin, J., Rozovskaya, O., andOacyna, J.: Inventory Review 2005 - Emission Data reported to CLRTAP and under the NEC Direc-tive - Initial review for HMs and POPs .EMEP Status report, Norwegian Meteorological Institute, Oslo, 2005.

Yin, Y., Chevallier, F., Ciais, P., Broquet, G., A. Fortems-Cheiney, Pison, I. and Saunois, M: Decadal trends in global CO emissions as seen by MOPITT, *Atmos. Chem. Phys*., 15, 13433-13451, 2015.

Yumimoto, K. and Uno, I.: Adjoint inverse modeling of CO emissions over Eastern Asia using four-dimensional variational data assimilation, Atmospheric Environment, 40, 35, 6836-6845, DOI: 10.1016/j.atmosenv.2006.05.042, 2006.

Wang, Y., G. Broquet, P. Ciais, F. Chevallier, F. Vogel, N. Kadygrov, L. Wu, Y. Yin, R. Wang and S. Tao: Estimation of observation errors for large-scale atmospheric inversion of CO_2emissions from fossil fuel combustion, Tellus B: Chemical and Physical Meteorology, 69:1, DOI: 10.1080/16000889.2017.1325723, 2017.

Wang, Y., Broquet, G., Ciais, P., Chevallier, F., Vogel, F., Wu, L., Yin, Y., Wang, R., and Tao, S.: Potential of European $^{14}CO_2$ observation network to estimate the fossil fuel CO_2 emissions

via atmospheric inversions, Atmos. Chem. Phys., 18, 4229–4250, https://doi.org/10.5194/acp-18-4229-2018, 2018.

WHO World Health Organization: Ambient Air Pollution: a global assessment of exposure and burden of disease, 2016.

Zheng, T., French, N. H. F., and Baxter, M.: Development of the WRF-CO2 4D-Var assimilation system v1.0, Geosci. Model Dev., 11, 1725–1752, https://doi.org/10.5194/gmd-11-1725-2018, 2018.

Zheng, B., Chevallier, F., Yin, Y., Ciais, P., Fortems-Cheiney, A., Deeter, M. N., Parker, R. J., Wang, Y., Worden, H. M., and Zhao, Y.: Global atmospheric carbon monoxide budget 2000–2017 inferred from multi-species atmospheric inversions, Earth Syst. Sci. Data, 11, 1411–1436, https://doi.org/10.5194/essd-11-1411-2019, 2019.

Report #1

I thank the authors for the numerous additions and improvements to their paper. The description of the model and inversion system is now much more complete. Nevertheless, I still see some important issues (see below) that should be addressed before I can recommend the paper for publication in GMD.

We wish to thank the referee for his/her helpful comments and for this positive acknowledgment of the improvements of our manuscript. The full review is copied hereafter and our responses are inserted in bold.

Main comments

1) There is an unusually large number of typos, grammatical and orthographical errors. This is quite distracting while reviewing a paper. I encourage the authors to put more efforts into this aspect in the future.

We apologize for the inconvenience. We have conducted a much more cautious and proofreading of the new version of the manuscript to avoid it.

2) The performance of the inversions is relatively poor, especially in the NOx case. Two possible explanations should be explored: 1) the number of iterations -- I doubt that the criterion of 90% reduction of the gradient of J is sufficient. Is computing time really such a hard constraint that more iterations cannot be tested? I would be curious to see the results for a 99% reduction.

We now present results with reduction of the norm of the gradient of *J* for the NO_x case higher than 90%, going up to 99% (see the new Table 3 describing sensitivity tests for the NO_x inversion).

Nevertheless, it should be noted that in some cases, the minimizer finds a local minimum. Once in such a local minimum, the minimizer runs many simulations but cannot get out into a new direction and perform more iterations. This would of course not be the case for a linear problem. We have added details in Section 2: "As shown in Figure 1, the minimization algorithm repeats the forward-adjoint cycle to get an estimate close to the optimal solution of the inversion problem for the control parameters. This approximation of the optimal estimate is found by satisfying the convergence criteria of the minimizer with a given reduction of the norm of the gradient of *J*. Nevertheless, due to the non-linearity of the problem, the minimization may reach a local minimum only, instead of the global minimum."

2) the choice of a priori errors is perplexing: 30% in the case of NOx, 100% for CO. There is no possible justification for this. Furthermore, correlations are used for one compound, not for the other

Our examples had been chosen to illustrate the parameters of the system configuration and the way the code works. The aim was not to provide a fully comprehensive scientific study on NO_x and CO inversions. Still, our choice of prior error parameters now follows insights from previous studies. In particular, Souri et al. [2020] performed NO_x inversions with prior error standard deviations set at 50% of the prior estimate of the anthropogenic emissions. For the inversion of CO emissions, the error standard deviations assigned to the prior CO emissions are set at 100%. This value of 100% has already been chosen in Fortems-Cheiney et al. [2011] and in Fortems-Cheiney et al. [2012].

We have added such justifications for our choices of prior errors in Section 4.2.2. Furthermore, we now present some tests of sensitivity of the NO_x inversions to the prior error assigned to prior emissions and initial conditions.

Souri, A. H., Nowlan, C. R., González Abad, G., Zhu, L., Blake, D. R., Fried, A., Weinheimer, A. J., Wisthaler, A., Woo, J.-H., Zhang, Q., Chan Miller, C. E., Liu, X., and Chance, K.: An inversion of NO_x and non-methane volatile organic compound (NMVOC) emissions using satellite observations during the KORUS-AQ campaign and implications for surface ozone over East Asia, Atmos. Chem. Phys., 20, 9837–9854, https://doi.org/10.5194/acp-20-9837-2020, 2020.

Fortems-Cheiney, A., Chevallier, F., Pison, I., Bousquet, P., Saunois, M., Szopa, S., Cressot, C., Kurosu, T. P., Chance, K., and Fried, A.: The formaldehyde budget as seen by a global-scale multi-constraint and multi-species inversion system, Atmos. Chem. Phys., 12, 6699–6721, https://doi.org/10.5194/acp-12-6699-2012, 2012.

The small NOx emission errors likely explain the poor match of a posteriori NO2 with OMI. I think we can all agree that the choice of a priori errors and correlations is arbitrary to some (large) degree. This is why it is important to assess whether the inversion results are dependent on such choices. Clearly this exploration is not considered important here, which I think is a mistake. Especially in view of the weak bias reduction achieved with the current setup.

We now present tests of sensitivity to the amplitude of prior errors assigned to prior initial conditions and emissions to provide insights the impact of these choices on the

inversions results. We also performed a test of sensitivity to spatial correlations in these prior errors.

However, the biases between OMI and simulated NO_2 tropospheric columns is a complex topic that is not related to our CHIMERE simulations only [Huijnen et al., 2010; Souri et al., 2020; Elguindi et al., 2020] and that would require very careful analysis. Addressing it properly is thus clearly out of the scope of this paper.

We have added text: "The biases between OMI and simulated NO_2 tropospheric columns is a complex topic that is not related to our CHIMERE simulations only [Huijnen et al., 2010; Souri et al., 2020; Elguindi et al., 2020]. It requires a fully comprehensive scientific study which is out of the scope of this paper. »

Huijnen, V., Eskes, H. J., Poupkou, A., Elbern, H., Boersma, K. F., Foret, G., Sofiev, M., Valdebenito, A., Flemming, J., Stein, O., Gross, A., Robertson, L., D'Isidoro, M., Kioutsioukis, I., Friese, E., Amstrup, B., Bergstrom, R., Strunk, A., Vira, J., Zyryanov, D., Maurizi, A., Melas, D., Peuch, V.-H., and Zerefos, C.: Comparison of OMI NO_2 tropospheric columns with an ensemble of global and European regional air quality models, Atmos. Chem. Phys., 10, 3273–3296, https://doi.org/10.5194/acp-10-3273-2010, 2010.

Elguindi, N., Granier, C., Stavrakou, T., Darras, S., Bauwens, M., Cao, H., et al.: Intercomparison of magnitudes and trends in anthropogenic surfaceemissions from bottom-up inventories, top-down estimates, and emissionscenarios. Earth's Future, 8,e2020EF001520. https://doi.org/10.1029/2020EF001520, 2020.

Minor comments

Figures 4-5 Is it a coincidence that the right panel of Fig. 4 presents high values off the coast of Egypt, which is where the MACC-based run is most different from the standard one using LMDZ-INCA? It looks like the run without emissions used boundary and initial conditions from MACC.

Simulations with null emissions have been conducted with two different sets of boundary conditions LMDz-INCA and MACC. In Figure 4, indeed, the one that is shown used boundary and initial conditions from MACC. It has been corrected.

l 365 Why is the median appropriate to take proper account of the AK ?

For each model grid cell, we have several observations and corresponding averaging kernels. Taking the average of both would not make sense, as the averaged averaging kernel would not be consistent with the average observation. Conversely, taking the median of the observations allows selecting a specific observation with its corresponding averaging kernel which is consistent.

Technical/language corrections

The appearance of mathematical symbols appears variable throughout the manuscript. For example, the B1/2 notation of l 170 (also 173) and 293. Please use consistent fonts throughout the text and equations.

We apologize for the inconvenience. Efforts have been put into this aspect.

l 95 "Van" -->"van"

It has been corrected.

l 162 "high-non linearity" ->"high non-linearity"

It has been corrected.

l 162-165 : Difficult to read due to sentence within parentheses; please rephrase.

We have removed the sentence within parentheses.

l 200: delete "a first time"

It has been done.

l 204 Problem with reference (parentheses)

It has been corrected.

l 210 A space is missing before "The adjoint"

It has been corrected.

l 212 "Then, it has been parallelized. This work" -->"The code has been parallelized. This task"

It has been corrected.

l 213 "whole code, associatedwith" -->"entire code, associated with"

It has been corrected.

l 213 missing space between sentences

It has been corrected.

l 214 missing space between sentence

It has been corrected.

l 216 "lead" -->"conducted"

It has been corrected.

l 219, 221, 222 Insert space after the bullet

It has been done.

l 219 "For the geometry" -->"Regarding geometry". Id. for next bullets

It has been corrected.

l 233 "infer" ->"infers"

It has been corrected.

Figure 2 The blue could be a bit lighter for lisibility (also in Figure 3)

It has been done, for Figure 1 and Figure 2. Figure 3 has been removed.

l 248 "to constrain could be" -->"to be constrained might be"

It has been corrected.

l 252 Insert a comma after "add"

It has been done.

l 256-260 I cannot understand what is explained here. What is meant by "activity maps and/or masks for regions"? What is meant by "control of budgets"? Etc. Please rephrase.

We have rephrased the sentences: “For type “scale”, the control variables are scaling factors applied to maps different from the maps of emissions used as prior input of the forward model: for example, activity maps can be used and scaled to get emissions; the obtained values are then added to the corresponding components of the model inputs. With these various types, it is possible to define the control variables as the budgets of

emissions for different regions, types of activities, and/or processes, which can thus be directly rescaled by the inversions, similarly to what is done in systems where the control vector is not gridded [Wang et al., 2018]).”

l 264-265 "to define the diagonal standard deviation matrix SIGMA" : do you mean that the matrix is diagonal, or are your referring to the diagonal of the matrix? SIGMA has not been defined previously.

We do not agree, SIGMA was defined l256-258 with the sentence “The variances are specified by the user through standard deviation coefficient (Table 1), which can be a fixed value ("fx") or a percentage ("pc") to define the diagonal standard deviation matrix ∑.”

l 281 Insert space before "where"

It has been done.

l 291 Subscript "t" in "Ct"

It has been corrected.

l 293 "The calculations involving B1/2" : it would be helpful to use equation numbers, to help the reader figure out where such calculations were mentioned.

It has been done.

l 294 Subscripts for Ct and Cs

It has been corrected.

l 298 "computation time"

It has been corrected.

l 313 "Simplified scheme describing how...". "prepares" -->"prepare".

It has been corrected.

l 313 In the colored box, what does mean "PYVAR y building"?

Figure 3 has been removed.

l 344 Insert a space before "OMI"

It has been corrected.

l 344 "to present an illustration"

It has been corrected.

l 381 and 383 "come from" --> e.g. "are obtained from"

It has been corrected.

l 388 Insert a space after "site"

It has been corrected.

l 391 "strongly driven" -->"strongly influenced"

It has been corrected.

Figure 4, the right panel is denoted using "a)" instead of "b)", please correct.

It has been corrected.

Fig 4 & 5 The legend could be more clear, e.g. "CO surface concentrations between 1-7 March 2015, simulated by CHIMERE...". You might drop "over Europe"

It has been corrected. The new legend for Figure 4 is: "Mean CO surface concentrations from the 1st to the 7th, March 2015 simulated by CHIMERE a) with anthropogenic and biogenic emissions, and b) without emissions, in ppbv, at the 0.5°x0.5° grid-cell resolution." The new legend for Figure 5 is: "Mean CO surface concentrations averaged from the 1st to the 7th, March 2015 simulated by CHIMERE using for initial and boundary conditions, a) the climatological values from the LMDZ-INCA global model b) the climatological values from a MACC reanalysis, in ppbv, and c) the relative differences between these two simulations , in %, at the 0.5°x0.5° grid-cell resolution ».

Figure 5 The legend says "relative difference" but the figure shows "ppbv"... Which one is it? In any case, the color scale is poorly chosen since almost all values are positive. Relative differences of 15% are still significant.

The legend has been corrected. The color scale has also been changed.

The sentence describing Figure 5 has been changed: "the results were similar, with maximum relative differences in concentrations of about 15% over continental land (Figure 5c)."

l 406 Drop "of their equivalents"

It has been done.

l 408-411 What is meant by "general trend in emissions"? Furthermore, if the underestimation persists throughout the year, it might still be due to specific activity sectors, isn't it? Please rephrase in a more logical manner.

The paragraph has been changed: "These discrepancies might be due to different causes, which can all interact. A source of uncertainties is related to the observations. For example, satellite data inter-comparison studies reveal large differences between different retrievals of the same compound [Qu et al., 2020]. It can be explained by uncertainties from the CTM (e.g., through the underestimation of the atmospheric production or the underestimation of the species lifetime). It could also be explained by an underestimation of the anthropogenic emissions in the BU inventory."

Qu, Z., Henze, D. K., Cooper, O. R., and Neu, J. L.: Improving NO_2 and ozone simulations through global NO_x emission inversions, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2020-307, in review, 2020.

l. 412 "chemistry with OH" rephrase.

A shown just above, the paragraph has been changed.

Figures 6 & 7 Legend: "Mean bias" -->"Mean biases"

It has been corrected.

l 419 & 423 "for 7-day" is ambiguous. Only one value for the period, or several?

We agree. We have one corrective value per pixel for the 7-day period. We have changed the description of the control vector with for example: "the CO anthropogenic emissions at a 7-day temporal resolution, a 0.5° ×0.5° (longitude, latitude) horizontal resolution, and 8 vertical levels, i.e. 101×85×8 components in x».

l 427 "for 1-day" : as above.

It has been corrected.

l 433 "ofthe"

It has been corrected.

l 433 "We hardly have sources of estimates" : weird, re-phrase.

To our knowledge, there are only few available studies dealing with the estimates of the uncertainties in gridded bottom-up emission inventories at the 0.5°x0.5° resolution or higher. For NO_x, to our knowledge, there is no study dealing with such uncertainties. For CO, Super et al. [2020] estimated uncertainties for specific grid cells at a 1km^2 resolution over Western Europe. This study was already cited in the text.

We have rephrased: "To our knowledge, there are few available studies dealing with the estimates of the uncertainties in gridded bottom-up emission inventories at the 0.5°x0.5° resolution or higher".

l 437 "is" -->"are"

It has been corrected.

l 439 remove the commas before the year in the references.

It has been done.

l 449 "obtained" -->"obtain"

It has been corrected.

l 451 "even 100% of uncertainty lead to" -->"even an uncertainty of 100% leads to"

It has been corrected.

l 463 "deriven by" -->"driven by"

It has been corrected.

l 480 Space missing after Portugal

It has been corrected.

l 500 Please write 1.4Ee+15 as 1.4(times)10^{15}. Same elsewhere in the paper (e.g. legend of Figure 7)

It has been corrected.

Report #2

General Comments:

This manuscript presents an added inverse modeling system capability for CHIMERE based on PYVAR, which has been successfully used with LMDz for GHG flux estimation. The manuscript describes its model development and show some very promising results for CO and NOx. This version of the manuscript has improved relative to original submission, addressing several of previous reviewers comments. I commend the authors for this revision. However, while this is relevant to the scope of GMD, I have several concerns with regards to the clarity and presentation of its description and substantive presentation of results, especially on quantifying some uncertainties or describing its fidelity. I suggest minor revisions to address the following concerns before publication:

We wish to thank the referee for his/her helpful comments and for this positive acknowledgment of the improvements of our manuscript. The full review is copied hereafter and our responses are inserted in bold.

1) Description of the system could be improved by:

a) clear differentiation of new developments in PYVAR for reactive species. In its current form, it appears to be just an implementation of PYVAR to CHIMERE. If so, detailed testing strategies is warranted to show the fidelity of i) CHIMERE adjoint/TLM, ii) PYVAR global minimization especially for the current application to reactive species. How are these tested and quantified?

The "PYVAR" framework in PYVAR-CHIMERE is actually quite different from that of the PYVAR system that has been used along with LMDZ. The development of an inversion system based on the PYVAR code and on CHIMERE that was suitable for our regional applications required new options for the control and observation vectors, for the set-up of uncertainties and for the interfaces with the model (that are detailed in this paper). This explains that the development of PYVAR-CHIMERE did not just consist in plugging CHIMERE to the PYVAR code.

Detailed testing strategies for the TL and for the adjoint are now described in Section 3.3.

b) clear differentiation between PYVAR and 4D-Var. How is PYVAR different than other 4DVar approaches? What are its advantages and disadvantages? What are its strengths and limitations? Please add some context.

The mathematical scheme we use to control the concentration initial and boundary conditions along with the surface fluxes using the chemistry transport model and its tangent and adjoint codes is very similar to the 4D-Var schemes used in meteorological

data assimilation. However, for GHG/pollutant surface flux variational inversions, the main part of the control vector is fluxes over the whole data assimilation period and not the initial conditions of the state vector of the dynamical concentrations. The dynamical model connects these controlled fluxes to the observation variables (concentrations over the whole data assimilation period). Therefore, the term "4D-Var" does not really apply. For this reason, the atmospheric inverse modeling community rather uses the term "variational inversion". We thus assume that the question here is about the differences between PYVAR and other variational inversion systems.

Historically, inversion systems have been built around specific CTMs and to our knowledge, there is no study comparing different variational assimilation codes around the same CTM. In general, variational codes differ mostly in the options and features for the definition of the control and observation vectors, or of the prior and observation errors. The use of different CTMs can bring different constraints for such parameters. The underlying minimization algorithms are generally identical (CONGRAD or M1QN3 in most cases, with minor differences in implementation). With CMAQ [Hakami et al., 2007], PYVAR-CHIMERE is one of the few systems based on an Eulerian area-limited CTM. The main strengths of PYVAR-CHIMERE come from the strengths of CHIMERE and from its high modularity for the definition of the control vector (as detailed in this paper). CHIMERE is an extremely flexible code, in particular for the definition of the chemical scheme; chemical reactions are defined in a configuration file and can be switched on/off, added/removed smoothly, making it especially relevant for air quality studies.

We have added information about CHIMERE in the introduction: "CHIMERE is dedicated to the study of regional atmospheric pollution events [Ciarelli et al., 2019, Menut et al., 2020], included in the operational ensemble of the Copernicus Atmosphere Monitoring Service (CAMS) regional services. The main strengths of PYVAR-CHIMERE come from the strengths of CHIMERE and from its high modularity for the definition of the control vector. CHIMERE is indeed an extremely flexible code, in particular for the definition of the chemical scheme."

c) Clear description of Figure 3. In its current form, it shows a list or table of variables and parameters. It would be clearer if details on this Figure are discussed. What do these

variables represent? Are there utilities in PYVAR system that are used? If so, it may be better to name and describe those utilities.

Figure 3 has been removed, as the Section 3.4 already well describes the calculations of the equivalents of the observations.

d) clear mathematical expressions and their representation (including consistent use of nomenclature). In particular, how does the time component of the cost function (i.e., 4DVAr versus PYVAR) treated in PYVAR? Can the "add", "mult", and "scale" be incorporated in the cost function notation? Where are these correction estimated and applied in the algorithm?

See our point regarding the use of the term "variational inversion" rather than "4DVAr".

The control vector and observation vectors gather fluxes and concentrations (respectively) over the whole data assimilation temporal window. Decomposing the cost function into a sum of comparisons between the projection of the control variables in the observation space and the observations at different times, similarly to what is often done in meteorology, would actually make it quite confusing and uselessly complicated. The *H* operator and the H matrix bear the "time component". We think that the text makes it clear. We now refer to the paper of Rayner et al. [2019] lying the basis for variational atmospheric transport inversion to avoid having to redevelop such fundamental points.

We think that developing the details of *H* to show how different its columns can be in practice depending on the options for each type of control variable 'add', 'mult', and 'scale' (i.e. depending on the way the components of the control vector are defined) could be uselessly complicated (e.g., requiring a lot of indices, blocks of *H*, and different types of notations associated to the different types of components in x). In the case of 'add', the corresponding variables in the control vector and B matrix are in the physical space: in the case of surface emissions, molecules.cm^{-2}.s^{-1} are stored in the control vector. For 'mult', the corresponding variables in the control vector and B matrix are unitless numbers: these numbers are multiplied by reference emissions before running CHIMERE. The 'scale' option is similar to the 'mult' option, but applies to activity maps, instead of reference emissions.

Figure 2 has been updated and now better represents where these corrections are applied in the algorithm.

Rayner, P. J., Michalak, A. M., and Chevallier, F.: Fundamentals of data assimilation applied to biogeochemistry, Atmos. Chem. Phys., 19, 13911–13932, https://doi.org/10.5194/acp-19-13911-2019, 2019.

2) Presentation of results could be improved by:

a) Some diagnostics to check optimality of the inversion algorithm. For example,
i) posterior error covariances – if calculated especially error reduction estimates,
As already discussed in Section 2, posterior error covariances are not a straight forward product of variational inversion systems. We have added a reference to Rayner et al. [2019] in this paragraph. We have also added details: "Nevertheless, it should be noted that the cost of the Monte Carlo experiments used to derive these posterior uncertainties is huge."
We documented such computations and tests of consistency in Broquet et al. [2011] and Kadygrov et al. [2015]) for CO_2 and the assimilation of data from surface networks. However, such an exercise here for satellite data and highly reactive species is much more difficult.

ii) RMSEs.This reviewer understands that comparison with independent measurements is beyond the scope of this paper. But at least provide some indication of its optimality. Results on scaling factors and increments can only be interpreted if compared to independent datasets and other approaches. However, one can show for example that the minimization reached its minimum or show the breakdown of the cost function to show that the observations are able to constrain some elements of the control vector. In its current form, this is shown qualitatively in Figure 5 to 8. At least quantify these by statistics other than mean biases. These are very promising results and should be highlighted more.

We agree that the decrease of the misfit to the observation can be an indicator of the good behavior of the system. We have added diagnostics to check the efficiency of our system, such as RMSEs, standard deviations and correlations. We have added Table 3 and Table 4, respectively for CO and NO_2.
Sentences have been added in Section 4.2.3 for CO: "Over this area (see the grey box in Figure 6), the mean bias between the simulation and the observations has been reduced by about 27% when using the posterior emissions (mean bias of 11.6 ppbv, Table 3) instead of the prior emissions (mean bias of 15.9 ppbv, Table 3). The RMSE and the standard deviation have been reduced by about 50% and the spatial correlation has been strongly increased."

Sentences have also been added in Section 4.2.4:"Over this area (see the purple box in Figure 6), the mean bias between the simulation and the observations has been reduced by about 24% when using the posterior emissions (mean bias of $1.9x10^{15}$ molec.cm^{-2} against $2.6x10^{15}$molec.cm^{-2} with the prior emissions, Table 5). The RMSE and the standard deviation have been reduced by about 7%. The correlation has not been improved. Even with high emission increments, the impact on the tropospheric columns is consequently rather small. The posterior emissions and their uncertainties will have to be evaluated and may bring hints to the cause of the discrepancies between simulated and observed NO_2 tropospheric columns."

As an indication, in the illustration for NO_x inversion (inversion E), J decreases from 622 to 594. Jo, the part of the cost function corresponding to the misfits to observation, is decreased by the inversion from 622 to 567. Jb, the component corresponding to the misfits to the prior estimate of the fluxes, increased from 0 to 27.

b) Spell and grammar checks, as well as clearer formats of tables.
We apologize for the inconvenience. We have conducted a much more cautious and proofreading of the new version of the manuscript to avoid it.

Specific Comments:

1) Abstract: Please add some numbers for your results (especially error reduction)
Information about the biases has been added in the abstract: "In these cases, local increments on CO emissions can reach more than +50%, with increases located mainly over Central and Eastern Europe, except in the south of Poland, and decreases located over Spain and Portugal. The illustrative cases for NO_x emissions lead to large local increments (> 50%), for example over industrial areas (e.g., over the Po Valley) and over the Netherlands. The good behavior of the inversion is shown through statistics on the concentrations: the mean bias, RMSE, standard deviation and correlation between the simulated and observed concentrations. For CO, the mean bias is reduced by about 27% when using the posterior emissions, the RMSE and the standard deviation are reduced by about 50% and the correlation is strongly improved; for NO_x, the mean bias is reduced by about 24% , the RMSE and the standard deviation are reduced by about 7% but the correlation is not improved."

2) Line 16: “in addition to greenhouse gases”. I understand that PYVAR has been used for GHG before, but I don’t think it’s used here. Good for the intro but perhaps not in the abstract. Focus on what exactly is new here.
The reference to greenhouse gases has been removed.

3) Introduction: While discussion on emissions and inversions can be interesting, some of these are general statements which may not be exactly what this paper is specifically addressing and demonstrating. Focus on what issues/problems exactly the study addresses. For example, is this study addressing high resolution emission estimations (at 0.5 degrees?) or O3 or GHG? I understand that some of these statements motivate future work but perhaps make it more concise and focus more on what exactly does this paper at its current form address in terms of scientific problems?

Ozone (O_3) is mentioned in the first lines of the introduction to describe general air quality issues. We choose to keep it. GHGs are mentioned to describe the evolution of the PYVAR-CHIMERE system, initially developed for the quantification of fluxes of long-lived GHG species such as CO_2 and CH_4. We also choose to keep it.

We agree that a 0.5° resolution is not a high spatial resolution. We have changed the sentence: “Here, we present the Bayesian variational atmospheric inversion system PYVAR-CHIMERE for the monitoring of anthropogenic emissions of reactive species at the regional scale”.

Why are other approaches not sufficient? What are the limitations of these approaches?
Ensemble, reduced rank Kalman filter or analytical methods are popular in the inverse modeling community. Analytical methods and Kalman filters rely on strong assumptions of the linearity of the response of the observation operator to correction to the prior fluxes, which can be problematic for inverse modeling problems with highly reactive species. All these methods are low or reduced rank inversion techniques that rely on small control vectors or on a strong reduction of the space in which the inversion searches for its optimal solution. Variational inversion systems allow solving for high dimensional problems, typically solving for the fluxes at high spatial and temporal resolution, which can be critical to fully exploit satellite images. We have added this last sentence in the introduction.

4) Line 102. Check spacing
It has been done.

5) Line 121. “in input of the CTM”. Please clarify
We have removed this part of the sentence.

6) Line 124: “during the inversion process (surface fluxes …)” Perhaps make this two separate sentences?

This has been done: “The control vector x contains these variables to be optimized during the inversion process. It can be surface fluxes but it may also include initial or boundary conditions for example, as explained in Section 3.3.”

7) Line 121-134: parameters versus variables? Are they the same?

They are the same, but we have homogenized the sentences by keeping the same vocabulary “parameters”.

8) Line 130: “observation errors”. Why not call it model-data mismatch?

Observation errors is the appropriate vocabulary here.

9) Line 132: “given their prior estimates, the observations, the CTM and the associated uncertainties”. First, it may be better if “the” article is omitted on the succeeding segments of the sentence. Second, “given the CTM” may not be accurate way to describe this minimization.

The sentence has been changed: “In statistical terms, the inversion searches for the most probable estimate of the control parameters given their prior estimates, observations, CTM and their associated uncertainties”.

10) Line 133: “in the following”. Please clarify or omit.

It has been removed.

11) Line 135: What is xb? There is no mention of time in this section. Be consistent in notations (italic versus non-italic, bold vs not bold) for all expressions not just Eq. 1. What are the dimensions of these vectors and matrices?

$\mathbf{x_b}$ is now defined with the sentence: “The prior information about the parameters x to be optimized during the inversion process is given by the vector $\mathbf{x_b}$”.

The notations have been checked. The dimensions of these vectors and matrices depend on the inversion configuration chosen by the user. Illustrations are given in Section 4.

12) Line 137: “state vector x”. Is this the same as control vector? What do you mean by state? Is a parameter a state?

It was a mistake. We have changed the sentence: “*H* is the non-linear observation operator that projects the control vector x onto the observation space”.

13) Line 139: “includes the CTM”. Please clarify.

The sentence has been changed: “the observation operator includes the operations performed by the CTM in linking the emissions to the concentrations and any other

transformation to compute the simulated equivalent of the observations such as an interpolation or an extraction and averaging of the simulated concentrations fields".

14) Line 157: I suggest to may Eq 2 as a separate line. Should B be bold?

Equation 2 is now presented as a separate line. B indeed should be bold, it has been corrected.

15) Line 142: "errors are assumed to be centered and to have Gaussian distribution". What do you mean by centered? Unbiased? Is it necessary to assume Gaussianity? If so, then why not just solve the analytical solution? Because H is non-linear?

We indeed mean unbiased. The sentence has been corrected.

The fact that *H* is of high dimension mainly justifies the use of a variational approach. See our answer above.

16) Line 159: "optimal solution". What does optimality mean here?

The sentence has been detailed: "As shown in Figure 1, the minimization algorithm repeats the forward-adjoint cycle to get an estimate close to the optimal solution of the inversion problem for the control parameters. This approximation of the optimal estimate is found by satisfying the convergence criteria of the minimizer with a given reduction of the norm of the gradient of *J*. Nevertheless, due to the non-linearity of the problem, the minimization may reach a local minimum only, instead of the global minimum."

17) Line 163: Careful on spacing between words

It has been corrected.

18) Line 172: I suggest to make the modified equation 2 as a separate line. How is this linearization implemented? i.e., where is this linearization point relative to the iteration interval? Does making shorter intervals improve minimization and representation of non-linearity?

We do not understand this question and in particular what the reviewer calls the "iteration interval". CHIMERE uses inputs, such as emission fluxes, at an hourly resolution. Then, the computations of mixing and chemistry are done at a finer time resolution. The linearization of CHIMERE relative to anthropogenic emission fluxes for example is done for linearization points which are the emissions at the beginning of each hour and the computations for mixing and chemistry are run for the sub time steps inside each hour. The number of sub-hourly time steps is the same for the forward and

tangent-linear runs and is chosen to ensure a good representation of the physical and chemical processes in all cases without inducing too large a computing time.

The modified equation 2 is now presented as a separate line.

19) Line 173: what do you mean by "norms".

Here, "norm" is the mathematical term. It refers to Euclidean or *l*2 norms, as explained in Gilbert and Lemaréchal. [2009].

Gilbert, J.C.†and Lemaréchal, C.: The module M1QN3, Version 3.3, https://who.rocq.inria.fr/Jean-Charles.Gilbert/modulopt/optimization-routines/m1qn3/m1qn3.pdf, 2009.

20) Line 175: how do you address local minimum?

There is no way to know when you are in a local minimum rather than the global one. In practice, sensitivity studies can be performed.

21) Line 177-184: If there's an estimate of posterior uncertainty in this system, is this used in the study? Please state which approach is used.

Please refer to our response above.

22) Figure 1. Are B and R fixed? Caption has bold fonts.

B and R are matrices. They are represented with bold fonts.

23) Line 199: "without chemistry a first time". Please clarify.

We have removed "a first time" in this sentence.

24) Section 3.2 How do you diagnose if these adjoints are calculated accurately? Are there tests conducted for this purpose?

We have added a subsection describing the different tests of our system: "3.3. Accuracy of tangent-linear and adjoint codes

Different procedures have been implemented to test the accuracy of the tangent-linear (TL) and adjoint codes.

To test the linearity of the TL, we compute a Taylor diagnostic. It consists in computing the TL at x_0 for given increments Δx, $dHx_0(\Delta x)$, then the TL at x_0 for $\lambda \times \Delta x$ with λ an arbitrary small number, $dHx_0(\lambda \Delta x)$.

Theoretically, if the TL is well coded, $\lambda dHx_0(\Delta x) = dHx_0(\lambda \Delta x)$ by definition. In practice, the difference must be lower than 10 times the precision of the machine on which it is run.

The adjoint code is also tested, by verifying that $\langle H.\Delta x, H.\Delta x \rangle = \langle \Delta x, H^T H.\Delta x \rangle$ where H^T stands for the adjoint at x. What is actually computed is the ratio of the difference between the two scalar products to the second one and the accuracy of the computation. The difference should be a few times the precision of the machine on which it is run."

25) Line 207-216. Please check spacing between words.

It has been corrected.

26) Line 214: "lead with". Please clarify.

We have changed the sentence: "Changes have been implemented in the forward CHIMERE code embedded in PYVAR-CHIMERE to match the requirements of the studies conducted with this system".

27) Line 217-222. While important, I suggest to have them numbered but part of the paragraph rather than bulleted. And please elaborate each one.

We have changed the sentences: "Compared to the CHIMERE 2013 version [Menut et al., 2013], the most important of these changes are, regarding geometry, the possibility of polar domains and the use of the coordinates of the corners of the cells instead of only the centers, allowing the use of irregular grids. Regarding transport, the non-uniform Van Leer transport scheme on the horizontal has been implemented, which is consistent with the use of irregular grids. Finally, various switches have been added to keep the system consistent for GHG studies. For example, we can avoid going into the chemistry, deposition or wet deposition routines when the focused species do not require them (e.g. no chemistry for methane or carbon dioxide at a regional scale)."

28) Line 221: "when no species requires them". Please clarify. Do you mean for GHG for all – chem, dep? Or for a particular species that do not have either of these processes?

We mean for species that do not have either of these processes. As seen above, the sentences have been changed.

29) Line 224: "currently operational". Please clarify. Does this mean it is used in operational mode to forecast and predict? Also, is there a particular version of PYVAR and CHIMERE and PYVAR-CHIMERE used in this study?

PYVAR-CHIMERE is not used in operational mode to forecast and predict. We have changed the sentence: "PYVAR-CHIMERE is currently implemented with a full module of gaseous chemistry".

The CHIMERE version has been already described: "The development and maintenance of the adjoint means that the version used is necessarily one or two versions behind the distributed CHIMERE version". The inversion code in PYVAR-CHIMERE follows the general framework of the PYVAR codes documented in the series of paper using LMDZ from Chevallier et al. [2005] to Zheng et al., [2019] (and a significant amount of lines of codes). But in practice, this code differs in various ways, in particular due to its specific adaptations to CHIMERE and the types of problems it can deal with (e.g. non-linear), and regarding the definition of the control and observation vectors, or of the B and R matrices. In some sense, the name PYVAR-CHIMERE can be taken as a version of the PYVAR code. As given in the title, we have named the PYVAR-CHIMERE version "PYVAR-CHIMRE-v2019".

Chevallier, F., M. Fisher, P. Peylin, S. Serrar, P. Bousquet, F.-M. Bréon, A. Chédin, and P. Ciais: Inferring CO2 sources and sinks from satellite observations: method and application to TOVS data, *J. Geophys. Res*., 110, D24309, doi:10.1029/2005JD006390, 2005.

Zheng, B., Chevallier, F., Yin, Y., Ciais, P., Fortems-Cheiney, A., Deeter, M. N., Parker, R. J., Wang, Y., Worden, H. M., and Zhao, Y.: Global atmospheric carbon monoxide budget 2000–2017 inferred from multi-species atmospheric inversions, *Earth Syst. Sci. Data*, 11, 1411–1436, https://doi.org/10.5194/essd-11-1411-2019, 2019.

30) Table 1 is very informative. Please format accordingly, especially separating the header as it becomes confusing to read. Not sure if the "example of the definition.." row should be there. Can it be in the title?

We agree, it has been changed.

31) Section 3.3. Discussion of correction types is very informative as well. Is it possible to show how these are related to Eq. 1 to 3?

As explained above, the 'add', 'mult', and 'scale' correction types do not need to be explicitly included in the cost function.

Isnt it that the control vector consists of elements –corresponding to each grid point and species? If so, how is "scale" implemented to maps or masks for regions?

We have rephrased the sentences: "For type "scale", the control variables are scaling factors applied to maps different from the maps of emissions used as prior input of the forward model: for example, activity maps can be used and scaled to get emissions; the obtained values are then added to the corresponding components of the model inputs. With these various types, it is possible to define the control variables as the budgets of emissions for different regions, types of activities, and/or processes, which can thus be directly rescaled by the inversions, similarly to what is done in systems where the control vector is not gridded [Wang et al., 2018])."

32) Line 254: "which is similar to the control vector of budgets…" Please elaborate.

Please see the answer above.

33) Line 256: "adding the obtained values to the …" please rephrase.

Please see the answer above.

34) Line 259: "standard deviation coefficient". Please clarify. Is it really a coefficient? And since this is an error covariance matrix, should the diagonal elements be error variance not error standard deviation?

Indeed the diagonal elements are variances, as already explained in Section 3.4: "The variances are specified by the user through standard deviation coefficient (Table 1), which can be a fixed value ("fx") or a percentage ("pc") to define the diagonal standard deviation matrix ∑."

35) Line 260-262: Very important statement. But please elaborate or rephrase.

The description of the correction types has been detailed above.

36) Line 266: "variances". Are these error variances?

Yes, they are.

37) Line 270: " error correlation between fluxes of CO and NOx, are not coded yet". Please elaborate on its potential effect on your estimation?

We can not quantify this potential effect at this stage and to our knowledge, there is no study about this in the literature.

38) Line 296: How about calling this "Observation Operators"?

We do not agree. The section title remains "Equivalents of the observations".

39) Line 298: Please note spacing between words.

It has been corrected.

40) Section 3.4. I think this is very relevant. Please elaborate Figure 3. In its current form, it is not clear what this Figure represents and how we can use it to interpret results. I think coding of these operators is a vital step in the assimilation and should be given more emphasis. Are these utilities also available? How good are the adjoints of these operators? Are there tests to diagnose their accuracy?

As said above, Figure 3 has been removed. **We have added a subsection describing the different tests of our system: "3.3. Accuracy of tangent-linear and adjoint codes".**

41) Please check bold fonts in line 311 to 312

It has been done.

42) Line 314-318: Please highlight in your notations if these are scalars or vectors.

And please add corresponding dimensions. What is the difference between small (c_m(o)) and big C_m(o. What is x_a?

There is no difference between these two notations. Notations have been homogenized. As already noticed in Section 3.5, x_a is the prior profile provided with averaging kernels, when this is relevant.

43) Line 328-334: This is also informative. Is there a reference for parallelization approach in PYVAR and CHIMERE?

The parallelization approach for CHIMERE is described in the Section 2.2 of Menut et al. [2013].

How does it scale with more CPUs? 4 hours seem to be a long time, isn't it? Please elaborate and compare with other systems.

The optimal number of CPUs for the parallelization of the transport scheme depends on the size of the tiles (for the Van Leer scheme, they must be at least 6 grid cell large because of the upwind and downwind information required) and also of the technical characteristics of the machine, because of the time required to exchange halos. A set-up with many tiles on many CPUs requiring large amounts of exchanges for halos may be less efficient than a setup with less tiles (each being larger). The performances quantified on a given type of machine is not transposable to another since they are sensitive to CPU types etc.

44) Line 336-343. Check spacing between words.

It has been done.

45) Line 391-392. Why are they not different?

The sentence has been changed: "To characterize the uncertainties in the concentration fields due to the initial and lateral boundary conditions, we performed a sensitivity test by using either climatological values from LMDZ-INCA or a MACC reanalysis: maximum relative differences in concentrations of about 15% over continental land are estimated (Figure 5c)."

46) Figure 5 caption. "differences are in %" is in contrast to the units in the figure.

It has been corrected.

47) Figure 6 and 7. Is it possible to show difference plots? And more statistics (RMSEs, correlation, bias? Error reduction? Are these really surface concentrations? They are column measurements, right?

Figure 7 indeed represents tropospheric columns, the legend has been corrected.

What about initial conditions? Has this change as well since these are part of the control vector? Superscript on units?

The initial conditions are slightly changed. This is now described in Section 4.2.2: "With prior error standard deviations assigned to 15% of the initial conditions, the changes in initial conditions are very small (not shown) and do not affect the posterior emissions (test B, Figure 8)."

The superscripts on units have been corrected.

48) Section 4.2. Should this be presented prior to section 4.1.3 since some of the plots are for the posterior estimates?

We do not agree, we have kept the sections as initially presented.

49) Section 4.2.1. Can this be summarized in a table and discuss a little bit in the text as to the rational of the choice of these parameters? Am I to assume that NOx emissions are estimated only for 1 day, and all days are the same? For CO, what do you mean by 7-day? Average? How are emissions incorporated in CHIMERE in terms of time? Is there a distribution? i.e., diurnal and weekly cycle?

The choice of the parameters has already been summarized in Table 1. NO_x emissions are indeed estimated for one day in our illustration. As mentioned above, we have changed the description of the control vector in Section 4.2.1. Choices for prior error statistics are now better described in Section 4.2.2: "… different sensitivity tests described in Table 3 have been performed for the construction of the B matrix. For both

the prior NO and NO_2 emissions at 1-day and 0.5° resolution, the prior error standard deviations are first assigned to 50% of the prior estimate of the emissions (test A), as in Souri et al. [2020]. Sensitivity tests have also been performed with prior error standard deviations assigned to 80 and 100% of the prior estimate of the emissions (test C and test D, respectively, Figure 8). **With a prior error standard deviation assigned to 15% of the initial conditions, the changes in initial conditions are very small (not shown) and do not affect the posterior emissions (test B, Figure 8). As indicated in Section 3.4 and in Table 1, it is possible to use correlations in B, as in Broquet et al. [2011], in Broquet et al. [2013] and in Kadygrov et al. [2015].We have also demonstrated the strong impact of spatial correlations, defined by an e-folding length of 50km over land and over the sea, on our inversions results (test E, Figure 8).”**

Description about the TNO anthropogenic emissions have been added in Section 4.1.2:“The prior anthropogenic emissions for CO and NO_x emissions are obtained from the TNO-GCHco-v1 inventory [Super et al., 2020], the last update of the TNO-MACCII inventory [Kuenen et al., 2014]. This inventory is based on the EMEP/CEIP official country reporting for air pollutants done in 2017. It is an inventory at 6kmx6km horizontal resolution. From the annual and national budgets, each sector is assigned to a specific proxy to quantify the spatial variability of the emissions within each country. Temporal profiles are also provided per GNFR sector code (variations due to the month, weekday and hour). Following the Generation of European Emission Data for Episodes (GENEMIS) recommendations [Kurtenbach et al., 2001; Aumont et al., 2003], NO_x emissions are speciated as 90% of NO, 9.2% of NO_2, and 0.8% of HONO. The TNO-GHGco-v1 inventory has been aggregated to the CHIMERE grid.”

50) Section 4.2.2. Please check spacing of words and bold fonts.

It has been done.

51) Section 4.2.3. How about emission error reduction?

The emission error reduction is not a straightforward product of the method.

How do you ensure that these increments are “resolved by the observations”.

The method ensures it by design. Sensitivity tests will be done later over longer periods for example with constant emissions to quantify the impact of the observations in the system.

It would be great to see error reduction plots, if posterior error covariances are calculated.

They aren't, as explained before.

How about initial conditions? Did this change as well?

See the answer above.

52) Line 508-516. What is the implication of this to overall cost and computing and optimality of minimization including error correlation of CO and NOX (and spatial correlation against superobbing) as well as increase in dimension of control vector? This also entails using this system at higher spatiotemporal resolution, right? It would be great to have a section on limitations before future implication.

The high-resolution imaging of TROPOMI will indeed entail using PYVAR-CHIMERE at higher spatio-temporal resolutions, but for smaller domains (i.e., over countries rather than over Europe) as a compromise between resolution and the computational cost.