

Interactive comment on “Variational regional inverse modeling of reactive species emissions with PYVAR-CHIMERE” by Audrey Fortems-Cheiney et al.

Anonymous Referee #1

Received and published: 9 September 2019

This paper describes a Bayesian inverse modeling system for reactive compounds, PYVAR-CHIMERE. It also provides an illustration of what this system can do with two one-day inversions of emissions over Europe. The paper is generally well written, and the topic of the paper is relevant for this journal. Although the results indicate that the system has potential, there are several major issues which should be dealt with before this work can be published in GMD.

1) poor quality of the figures and equations. On the screen, it is more or less acceptable (with the zoom in feature of my pdf viewer), but upon printing, many figures (especially Figs. 1-4 and Figs. 6a and 8a) and equations are impossible to read. Figure 3 is really

C1

a Table and should be inserted as such. Some mathematical symbols (e.g. gradient on line 158, multiplication on line 167) are inappropriate.

2) Section 3.2 is 'Development of the adjoint of CHIMERE'. But the adjoint of CHIMERE was developed a long time ago (publications by L. Menut, I. Pison). What are the specific developments realized for this study, besides the minor changes to CHIMERE mentioned in the text ?

3) the results of inverse modeling studies are very dependent on the inversion setup, in particular the definition of the control vector and the construction of the covariance matrices \mathbf{R} and especially \mathbf{B} . What is the strategy in this regard? I understand that the main purpose of the paper is less to present specific inversion studies than to describe the general modeling framework. But the very simplistic choices made for the two one-day inversions suggest an absence of any strategy. The chemical lifetimes of the target species, the duration of the experiments, the initial and boundary conditions, and the assumed a priori uncertainties should all play a role in the inverse setup definition.

4) the two illustrations of PYVAR-CHIMERE capabilities are unconvincing. Yes, the system finds a minimum to the cost function, and the a posteriori simulation matches the observations quite well; but no, the a posteriori emissions are not shown to be closer to reality. With its long lifetime, CO is largely determined by the initial and lateral boundary conditions, which are part of the control state vector being optimized. The paper does not provide information on the a priori uncertainties for these parameters. A discussion is needed, and possibly sensitivity simulations. Note that, although the a priori simulation overestimates CO over Central Europe (south of Poland), the inversion increases the emissions there by about a factor of 2! Over Germany, the emissions are almost doubled in the Southern part, but are unchanged elsewhere. How can this be justified? Even for NO_x, in spite of their shorter lifetime, the initial and boundary conditions play probably a very important role, as indicated by the strong reduction of the model biases over ocean (compare Fig. 7c and 7d). The discussion of the results for NO_x (lines 373-376) is impossible to understand. It says that the optimization of

C2

NOx fluxes has only a small impact on the model biases. This is not true. Comparison of Fig. 7c and 7d show that the optimization works very well ! The authors claim that PYVAR optimizes only the NO2 fluxes, not those of NO. I don't believe this, it doesn't make sense. Please check this. In any case, clarifications and possibly a sensitivity analysis are in order.

Other comments:

- throughout the text, replace "NO2 emissions" by "NOx emissions" (if indeed, as should be the case, NOx emissions are optimized, not just NO2)
- lines 107-110: please refer also to GEOS-Chem adjoint papers (Henze, Kopacz, Cao etc.)
- l. 149 what is meant by "the control of emissions"? Please rephrase
- l. 168-169 What are the thresholds for the ratio between final and initial gradient norm, and for the number of iterations?
- Figure 1 In the orange box on the right, the order of operators should go backwards, shouldn't they?
- l. 195 leads (instead of lead)
- Figure 3. Explain the meaning of the "correction type" and of the three numbers in column "B variance coefficients".
- Figure 4. The legends mentions text in blue and in grey. I don't see that on the figure. Is this figure useful?
- l. 298 What is the resolution of ECMWF data? Are those data interpolated to the model grid?
- l. 300 Derognat et al. 2003 does not present a chemical mechanism, but refers to earlier papers.

C3

- What are the a priori lateral boundary conditions?
- on l. 303 and legend of Fig. 8, the information required to run the inversion are said to be listed in Table 1. This is not correct.
- l. 304 Are only anthropogenic emissions optimized? Or the total of all emissions?
- l. 306-307 The 3D initial conditions at the model resolution are said to represent 8585 components of the control vector. But what about the vertical dependence?
- l. 311-314 What are the non-anthropogenic emissions used in the model? Please provide a webpage and reference for EMEP emissions. Aren't there any publication or webpage for the TNO emissions?
- l. 328 Why the median?
- l. 341 With errors of 100% on the emissions, how can negative a posteriori emissions be avoided? How is this dealt with?
- Section 4.2 The Figure 5 shows both underestimations and overestimations by the a priori simulation. This is not well reflected in the discussion.
- Figure 7c and 7d should show absolute differences. A better color scale should be possible for Fig. 7a and 7b.

Interactive comment on Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2019-186>, 2019.

C4