

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

### **Anonymous Referee #2**

Received and published: 5 December 2019

The paper describes the variational data assimilation version of the CHIMERE, PYVAR-CHIMERE, which is capable of inversions of reactive gases. As a demonstration to the newly developed code, the inversion of CO and NO<sub>2</sub> is shown for two different days in late winter/early spring 2015. The paper's topic is of good relevance for GMD and contributes to a documented open source regional data assimilation system for reactive chemistry. Although the paper is generally well written, major changes are requested before publishing the manuscript in GMD. These major changes are:

- The quality of figures and formulas is unacceptable. Arrows should be larger/thicker (Fig. 1 and 2), annotations in Figs. 5, 6, 7, and 8 are too small, separation of subplots

C1

in Figs. 5 and 7 should be clearer. - The description of the inversion is unsatisfactory. The cost function and its gradient should explicitly show the model operator  $M$ , which is currently included in the state vector  $x$ . Further, it is unclear how the emissions are corrected. How can negative emissions be avoided? Are the emissions optimized for each time step or for the whole assimilation window? Are the emissions constant for the simulation time or does the inversion result in correction factors for the emissions. Then, the special treatment of 4D-var for emission factor optimization should be shown, e. g. how the positive definiteness of the correction factors is ensured. The manuscript must be more precise in this context.

- The calculation of the size of the control vector is erroneous. The vertical dependence of the initial conditions is missing in the calculation.

- In the experiment section (section 4) no information on the initial and boundary conditions is given. It should be illustrated to what degree both are changed during the inversion. Further, a comparison or sensitivity test should be shown on what the impact of emission optimization is compared to a joint optimization with initial and boundary conditions.

- Although the two test cases show a reduction of the difference between the assimilated observation and the analysis, this is not a proof of the successful operation of the data assimilation algorithm. A comparison with independent observations and a table with quality measures (e. g. bias, root mean square error, cost reduction) is necessary. It is advised to perform the analysis on a few consecutive days to assess the stability and quality of the inversion on different days.

- in the description of the test cases both, initial values and boundary conditions are included in the control vector, thus, the analysis is not complete without showing these two variables. A discussion is needed about the correction for all three variables, i. e. emissions, initial values, and boundary conditions, their relative influence on the analysis and about potential limitations of the inversion.

C2

- If the full adjoint of the chemical processes is used there should be an adjoint signal for other species than CO and NO<sub>2</sub> as well. This must be clarified. Are these signals simply not considered or not discussed? What is the reason for not optimizing NO emissions then?

- the model resolution of 0.5 x 0.5 square degrees seems to be a bit coarse for anthropogenic emission assessments. Is nesting available? A discussion on this point is needed.

- a better description of the B-matrix is needed in section 4.1.3. What about correlations for initial and boundary conditions?

Further minor comments:

- line 30: (VOCs) instead of "(VOCs))"

- line 39: reference for (LRTAP) would be appreciated

- line 43: no commas

- line 85: CO and NO<sub>x</sub> (instead of "CO, NO<sub>x</sub>")

- line 88: citation van der A. [2008] is not appropriate (van der A et al. [2008]), also in the reminder of the manuscript

- line 93/94: "... for which variational methods are more suitable than KFs by design": a reference would be appreciated for this statement.

- line 122/123: of the current inversion (instead "of the inversion")

- line 163: quasi-Newton (instead "quasi-Newtonian")

- line 165: Reference for incremental 4D-var approach is appreciated

- line 203: It would be appreciated if the manuscript contains a table with the available (and adjoint) processes of CHIMERE

### C3

- line 227/228: better: "PVAR, CHIMERE, and text sources are displayed in blue, orange, and grey boxes, respectively."

- caption of figure 4: better: "Simplified scheme of how PYVAR scripts prepare the observations  $y$  using satellite data. PYVAR and text sources are displayed in blue and grey boxes, respectively."

- line 264: Equation " $C_m =$ " is not a correct mathematical formulation,  $C_m(o)$  is a column,  $x_a$  is the state vector (a profile in this context).

- line 284/285 (general remark): The models speedup increases enormously if the master process is additionally used for calculation and parallel IO is included for suitable IO operations.

- line 290: ... days for CO and NO<sub>2</sub>, respectively (instead of "... days, respectively for CO and NO<sub>2</sub>")

- line 302: Table 1 is not control vector specific. This sentence can be removed

- line 304: for one day (instead "at a 1-day"); resolution (instead "resolutions")

- line 313/314: a spinup for the initial values is needed for an appropriate analysis, otherwise the model may be too far off the observations for a suitable correction

- line 317: Reference for MOPITT is missing

- line 328: MOPITT instead of "OMI"

- page 12, line 4: flown instead of "flying"

- line 368: parts (instead of "part"); present (instead of "presents")

- page 14, last line: particularly over the Po Valley (instead ", and particularly over Po Valley")

- caption Fig. 7 d: is it really the difference between prior and posterior? Inconsistency with text (see next point)

### C4

- line 374: Fig. 5c seems to be wrong here. Is it Fig. 7d?
- line 380/381: Using the full adjoint of CHIMERE, this must already be available. Please check for adjoint NO signals
- line 399: remove "for example"

---

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