

## General comments

This paper describes a newly developed toy model GQ-Chem for testing different data assimilation algorithms in the context of air quality forecasting. In addition, the authors present a series of cases testing existing widely used 3D-var methods against a relatively new technique, untested in an air quality framework, called 4DEnvr. The authors demonstrate the utility of GQ-Chem and the potential of the new data assimilation algorithm to address various key problems associated with forecasting air quality with chemistry transport models.

The paper clearly fits within the scope of GMD as the authors present a new model and test a new technique in the context of air quality and chemistry transport models. The paper is overall very well written and clear, but with numerous minor grammatical errors, and a couple of sections of unclear text. The methods and results are well presented. Scientifically, this paper is very interesting in the context of operational air quality forecasting, as the authors appear to demonstrate a method that addresses a set of key problems associated with forecasting and data assimilation for air quality without sacrificing, too much, model complexity or computational cost.

I therefore recommend this paper for publication following minor corrections outlined in the comments below.

## Specific comments

I found the last paragraph in section 2.1 to be unclear and somewhat confusing. For instance, it is not clear what was meant by “The sources of the QG model provided by ECMWF...”. Do the authors mean that the code for the dynamical model was given to them by ECMWF? There are some further examples listed below. I therefore recommend that the authors check this text carefully for its clarity and re-write it. Examples:

“The boundary conditions are taken cyclic...”, but it is not clear what “cyclic” means in this context.

“For all the experiments presented in this study a coarse resolution of 16x8 grid points has been used.” I have assumed that this is the horizontal resolution, but it is not clear whether 16 refers to the North-South dimension or the East-West. The authors should also write the horizontal resolution in terms of spatial resolution in km.

Figure 1 would probably be improved with an explanation of what the grid lines represent. One presumes these are the outlines of the grid boxes used in the model, but the authors should state this clearly.

Page 24, lines 1-2. The authors should probably remind readers that only the initial conditions are perturbed during these tests, and therefore NO<sub>2</sub> converges to the truth due to the shorter lifetime of NO<sub>2</sub> combined with the fact that there are no errors in the emissions.

Page 25, line 15. Can the authors think of any specific reasons why ozone and NO<sub>2</sub> show similar behaviour to CO and CO<sub>2</sub>?

The explanations for the emission perturbations at the beginning of Section 4.2 could benefit from slightly more explanation. It isn't stated directly, but I assume from the existing text that the perturbations result in reductions in NO emissions. This should be stated directly. Also, does the size of the perturbations vary with time, and if so at what frequency?

Please can the authors add some further text to the Figure 12 caption to explain what the red and blue colours represent.

## Technical comments

Page 1, line 5. I would recommend not using “reckon” in this context. I suggest changing from “Among the assets of 4DEnVar, we reckon the possibility to deal with multivariate aspects of atmospheric chemistry and to account for model errors of generic type.” to “The assets of 4DEnVar include the possibility to deal with multivariate aspects of atmospheric chemistry and to account for model errors of generic type.”

Page 1, line 10. Recommend removing comma “...of surface chemical emissions, for two meteorological and chemical...”.

Page 1, line 12. Misspelling of analysis “...for some species, analysys and next day forecast errors...”.

Page 1, line 22. Change to “ash” singular. “...modeling of volcanic ashes...”.

Page 2, line 19. Change “...involved into rapid chemical reactions...” to “...involved in rapid chemical reactions...”.

Page 2, line 31. Change “However, the model dynamics is neglected...” to “However, the model dynamics are neglected...”.

Page 3, line 4. Recommend changing “More important,...” to “More importantly,...”.

Page 3, line 6. Error in the text due to repetition and a fullstop that should not be there: “next day model forecast (Wu et al., 2008). next day forecasts of reactive gases such as O3 or NO2,”.

Page 3, line 7. Suggest changing “...,depend weakly on the initial condition...” to “... that depend weakly on the initial conditions...”.

Page 3, line 19. Recommend “EnKF naturally includes model uncertainties...” in place of existing text.

Page 3, line 25. Recommend using “determine” in place of “reckon” as determine sounds less informal.

Page 3, line 28. Life-times should not be hyphenated and should all be one word.

Page 4, line 12. Recommend changing “...which still lack for most of the operational CTMs...” to “...which are still lacking for most of the operational CTMs...”.

Page 4, line 19. “...supports naturally multivariate...” to “...naturally supports multivariate...”.

Page 5, line 11. “We remind that the objective...” to “We remind readers that the objective...”.

Page 5, line 14. Remove comma “...assimilation, to ease the exchange...”.

Page 5, line 28. Add “the” such that it now reads “...DA in the presence of complex...”.

Page 5, line 29. Again missing “the” prior to “...presence of...”.

Page 6, line 2. It should really be “2-Layer QG model”. The plural is not needed.

Page 6, line 19. “We remind that...” add “readers” before “remind”.

Page 7, line 17. "A summary of the chemical configuration is given in Table 1." Do the authors mean Table 2?

Page 7, line 20. Recommend changing to "...and a relatively clean atmosphere".

Page 7, line 21. Recommend changing to "...and a polluted atmosphere".

Page 7, line 4. Recommend changing to "...this gas behave like...".

Page 10, line 2. Consider changing to "It also participates in O<sub>3</sub> chemistry...".

Page 10, line 13. Spelling error of 'constant' "...due to costant...". I also recommend adding 'a' prior to 'constant'.

Page 11, line 2. Remove comma "...episodes, to demonstrate...".

Page 15, line 19. Recommend changing to "and a function of the chemical species."

Page 17, line 23. Do the authors mean to say "forecast on an hourly basis."

Page 18, line 5. Please change to "The meteorology is never observed nor...".

Page 18, line 21. There is no need to hyphenate "reanalysis".

Page 18, line 23. Please change to "different orders of magnitudes".

Page 19, line 2. Please change to "a horizontal length". It should be 'a' instead of 'an' in this case.

Page 19, line 2. I think there is a typo here when you say "Multimultivariatevariate". Please check.

Page 19, line 11. Figure 6 seems to be introduced before Figure 5 is mentioned.

Page 19, line 22. Please change to "...which lasts a few hours in a summer...".

Page 20, Figure 4. The top left panel seems to have been cut off by accident.

Page 20, line 3. Please change to "...can reach the same values...".

Page 20, line 4. Please change to "...emissions even in the presence...".

Page 20, line 5. Please change to "...4D<sub>EnVar</sub> provides...".

Page 20, line 7. Remove comma "...from 3D-Var, because daily trajectories...".

Page 22, line 3. Add comma before 'respectively' like so: "analysis (or forecast) at the grid point (i, j), respectively."

Page 24, line 2. Please change to "...the truth after a few hours...".

Page 24, line 18. Please change to "In the case of 3D-Var,...".

Page 24, line 19. Please change to "...species are kept at zero as well as for cross-variable correlations."

Page 25, line 8. Recommend changing "...when the sole initial condition is taken as source of uncertainty,..." to "...when the initial condition is solely taken as a source of uncertainty..."

Page 25, line 12. Please change to "...for the species that are perturbed initially."

Page 25, line 13. Please change to "This also justifies neglecting..."

Page 25, line 16. Please change to "...was able to provide similarly good results as in Sec. 4.1.1..."

Page 26, line 16. Please modify text "In a few cases,..."

Page 26, line 21. Recommend modifying text to "We readers remind that the..."

Page 26, line 24. Please change to "...found it more valuable..."

Page 28, line 4. Please change to "...experiments are conducted in the presence..."

Page 28, line 6. Please change to "In the case of reactive species..."

Page 29, line 1. Recommend changing to "The true NO emissions..."

Page 30, line 13. I recommend changing the text to "We remind readers that..." I recommend the same change for page 30, line 22.

Page 30, line 22. Recommend changing to "...more difficult for the third case."

Page 30, line 31. Recommend changing to "...matter what the spatial patterns..."

Page 31, line 1. Please modify text to "...observation number and observation errors."

The labelling of the colour bar for Figure 12 has been cut off. Please correct this.

Page 32, line 7. Please modify text to "...nor is it strongly coupled to..."

Page 33, line 2. Please modify text to "When the NO/O<sub>3</sub> relationship was strongly non-linear,..."

Page 34, line 1. Please modify text to "This represents the objective of..."

Page 34, line 5. I think the reference to Eq. 16 needs to be modified to include Eq. Check for other instances.

Page 36, line 5. Please change text to "The same values as before"

Page 36, line 17. Please correct text to "We confirm preceding findings"

Page 37, line 5. Please correct text to "We remind readers that the bias correction..."

Page 38, lines 18-19. Recommend changing text from “Among the main benefits we reckon:...” to “We determine the main benefits to be:...”. Reckon sounds informal in this instance.

Page 38, line 27. Change to “...for these types of studies.”

All instances of “i.e.” and “e.g.” should be written as: “, i.e.,” and “, e.g.,”. Such that both are bracketed at the front and back by commas.

When using ‘respectively’, the authors should pay attention that it is always placed in parenthetical commas if in the middle of the sentence or just before it if used at the end of a sentence, i.e., “..., respectively,...” or “..., respectively.”.