

Commentary on **Impact of a time-dependent background error covariance matrix on air quality analysis**

E. Jaumouillé¹, S. Massart¹, A. Piacentini¹, V.-H. Peuch² and D. Cariolle¹

¹ URA CNRS/CERFACS 1875, Toulouse, France.

² CNRM-Game/CNRS-URA 1357, F-31057 Toulouse, France.

The authors would like to thank the referees for their pertinent comments and suggestions on this paper.

Response to comments by referees

Reply to Referee #1 comments

- ▷ **Referee: Section 3. Line 16 : It is not sufficiently clear what authors mean under specific 3D-Var system? Is it the same 3D-Var assimilation procedure, which was applied for the Mocage CTM at the global scale? Please provide more details and description about specific implementation of this system and its difference to previous assimilation studies with Mocage at the global scale.**

Is the assimilation system the same 3D-Var FGAT Method as mentioned in previous Mocage studies? If FGAT, then what about mentioned issues in Massart et al. (2010), that 3D-Var FGAT does not localize well (in space) the increment when the dynamics are rapid. Rapid dynamics are typical cases at regional scale. Are FGAT issues explored for the current analysis?

The assimilation scheme used for this study derives from the 3D-Var FGAT method initially implemented for the Mocage CTM at the global scale. Nonetheless, this study is targeting air quality modelling in a regional domain and therefore we have adapted the Mocage-Valentina system : the algorithm has been coupled with the regional domain of Mocage and specific treatments for the regional boundary conditions of the model and of the approximation of the BECM have been introduced.

The 3D-Var FGAT weakness mentioned by the referee is particularly problematic at the global scale, where the dynamics (e.g. wind velocity, stream direction,...) have large geographical variation and it is therefore hard to choose an appropriate assimilation window length for the entire domain. At the regional scale, the dynamics are active, but their extremes are quite homogeneously distributed over the domain. In such a case we avoided any problem in the increment location by reducing the size of the assimilation window to one hour. As the frequency of the used synoptic observation is also one hour, this particular configuration of the 3D-Var FGAT makes it a simple 3D-Var.

To bring more details as suggested by the reviewer, we modified the text toward :

“To have a regional analysis at a better resolution than the one we could extract from an analysis performed at the global scale, a specific version of Valentina has been implemented. First Valentina was coupled with the regional domain of Mocage instead of its global domain. This allowed to perform the analysis only in the regional domain of Mocage. Used surface measurements are performed each

full hour. To avoid any problem we previously encounter with the 3D-Var FGAT method when the dynamics are rapid (Massart et al., 2010), we decided to use a simple 3D-Var method that produces an analysis each full hour when the data are available. Thus for each day, we have 24 successive analyses.”

- ▷ **Referee: Section 4. Line 18 : The note about values exclusion over ocean looks a little strange. Assuming that low emissions and small variability are the reasons of the inappropriate statistics over ocean, the authors should have similar problems over some land regions, like northern Europe, for example. So one should also exclude from current study these regions ? From another point of view statistics and BECM typically is not able to "see" land/ocean difference. So the motivation to exclude ocean regions is not clear. To improve understanding, may be more details about inappropriate statistics over ocean should be provided in this paper.**

We agree with the referee when he states that we should have similar behavior of our statistics over oceans and over some land regions. Due to this comment, we decided to remove the mask we had over the ocean and to show the statistics over the whole domain. This causes a few changes in the structure of the paper. The last paragraph of section 4.1 (page 887) has been suppressed. And more discussions have been added to the section where the statistics are presented.

Paragraph 4.2, page 888, from line 26, we modified the paragraph with : *“The length-scale values are relatively high over Ireland, Eastern-Europe, some oceanic regions and in the vicinity of domain boundaries (Fig. 3). This is related to the way the ensemble is built. First, in the vicinity of domain boundaries, when the wind is such that concentrations are influenced by boundary conditions, we lost some variability. This is for example the case for the region located West of the United Kingdom. In the upper left corner of the domain, the high values could be explained by a second phenomenon. In regions without observations (like oceans or Eastern-Europe) the ensemble has a lower variability because of the lack of constraint by the perturbed observations. The only source of variability comes there from the perturbed emissions. And in the region of low emissions like the North West part of the North Sea, due to the low amount of emissions, perturbed emissions do not bring variability. This is less the case in some other oceanic regions where the ships produce NO_x emissions that play a role in the ozone chemistry. In all the regions where the variability is low, the different simulations from all the members are thus very similar, which give a high correlation between them and results in large length-scales. To avoid too huge length-scale values, we have chosen to limite them to 200 km. Note that we also have high values for the length-scales in the Western part of the Strait of Gibraltar. There the ensemble of simulations are probably similar due to the dynamics, the strong wind advecting ozone from the Mediterranean Sea. ”*

Paragraph 4.3, page 889, from line 20, we modified the paragraph with : *“The ensemble-diagnosed standard deviations are inhomogeneous over Europe (Fig. 6). As already explained in the previous section, where there is no constraint by the perturbed observations and near the boundaries of the domain (like Eastern Europe) or over region with low emissions (like North Sea), we have low variability in the ensemble that results in low standard deviation. Note that over the Mediterranean Sea, in spite of the absence of observations, we produce some variability thanks to the perturbation of the emissions (from ships). Other geophysical features appear such (...)”*

▷ **Referee: Section 5.5** The Taylor diagrams in this study have not caught the difference between different versions for the BECM implementation. Presumably the results can be improved applying target-diagrams instead of the Taylor diagrams. In this scope, at conclusion section was also mentioned that the impact of the BECM formulation has been also difficult to evaluate because the Mocage model shows a systematic bias in situations with low ozone concentration. This systematic bias in the model is one more argument for the target-plots, which are able to differ between systematic and unsystematic errors.

As the reviewer suggested, target-diagrams have been plotted for the two studied periods as shown in Fig. 1 of this answer. As with the Taylor diagrams, the statistics of the five experiments issued from the assimilation process are very similar. But these diagrams give some information on bias that Taylor diagrams do not show : the bias (i.e. the difference between the mean of the analysis and the mean of the observations) is negative for the six experiments. The five experiments issued from assimilation have a greater bias than the Direct experiment during the summer period whereas these experiments reduce the bias during the winter period. We noted that during the summer period the standard deviation of the Percent experiment is larger than the standard deviation of the observations (that is not the case for the other experiments).

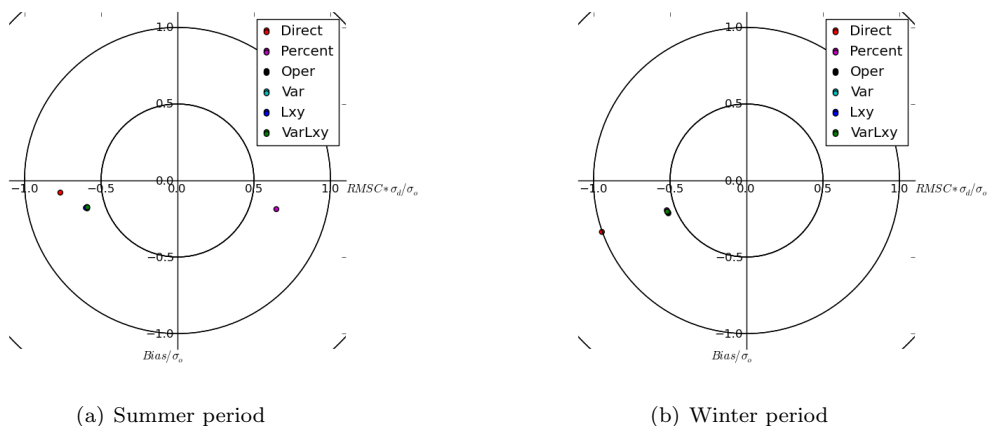


FIG. 1 – Target-diagram for several experiments studied, statistics averaged in the period between 1 and 10 July 2010 (left) and 1 and 10 December 2012 (right) for validation stations.

We will keep Taylor diagrams in the manuscript. But to bring more details on the statistics of the experiments, we have modified in section 5.5 the text toward : “(...) A systematic bias (i.e. the difference between the mean of the analysis and the mean of the observations) appears in the Mocage model, larger during the winter period than during the summer period (-3.37 ppb in winter against -1.62 in summer). The bias is reduced thanks to the five experiments issued from the assimilation process during the winter period but the bias is larger with these experiments during the summer period (Figures not shown). Those diagnostics (...)”

Reply to Referee #1 technical corrections

- ▷ **Referee:** The plots of the domains on Fig. 9 are inconsistent with previous plots (Fig. 1,2,3,6,8) and with domain definition at Section 3.1 : longitudes 16° East to 36° West and latitudes 32° South to 72° North . Please use the same consistent map with North and South-East Europe regions for all plots.

As the reviewer suggested, the same map domain for all plots appears in the new version of the manuscript.

Reply to Referee #2 comments

- ▷ **Referee:** p875 lines 17-26 : maybe some references are missing concerning air quality and especially the way that they are prescribing their BECM (Elbern et al, Blond et al, Wu et al . . .).

The paragraph has been rewritten to take more references into account :

“Recent studies of Massart et al. (2011) or Elbern et al. (2010) investigate the evaluation of the BECM but for global atmospheric chemistry. In air quality applications, several methodologies are currently used, but BECMs are mainly obtained with ensemble Kalman filter (EnKF) approaches. For example, Coman et al. (2011) used an EnKF in their analysis of partial lower tropospheric ozone columns to provide estimations of the background errors. Constantinescu et al. (2007) have also investigated the EnKF for the simulation of air-pollution in the Northeastern United States. Alternatively, in atmospheric chemistry as in other geophysical applications, one can use a statistical interpolation to specify an anisotropic and heterogeneous BECM (Blond et al., 2003) or can combine an ensemble approach with a variational data assimilation approach (Massart et al., 2011; Desroziers et al., 2008; Buehner, 2005). We have used this approach for global ozone analyses, and we examine it here to provide a time-dependent BECM for a regional chemical application.”

- ▷ **Referee:** p883 : concerning the hypothesis made in equation 13, is there a way (a priori or a posteriori) to verify that this hypothesis is sound ?

A simple one-dimensional case with three grid points can further explain the hypothesis. In that case, let's write $\mathbf{H} = (\alpha, \beta, 0)$ with $\beta = 1 - \alpha$, and $b_{i,j}$ the (i, j) component of \mathbf{B} . Then $\mathbf{HBH}^T = \alpha^2 b_{1,1} + 2\alpha\beta b_{1,2} + \beta^2 b_{2,2}$. If the observation is exactly in the middle of the first grid point $\alpha = 1$, $\beta = 0$ and $\mathbf{HBH}^T = b_{1,1}$ that is exactly the variance of \mathbf{B} in this grid point.

For different values of α and depending on the correlation between the two grid points, we make an error (see Fig. 2 where $b_{1,1} = b_{2,2}$ and the error is the normalized difference between \mathbf{HBH}^T and $b_{1,1}$). The correlations we diagnosed are mainly over 0.7. Even if an observation is near the boundary between two grid points, in this particular framework, the error is mainly under 10%. Based on this idealized study, we also tested the hypothesis in a 2 dimensional domain, using the observation locations, and our conclusions were similar. We chose not to detail the results of our

preliminary study as it is not a major results of the paper.

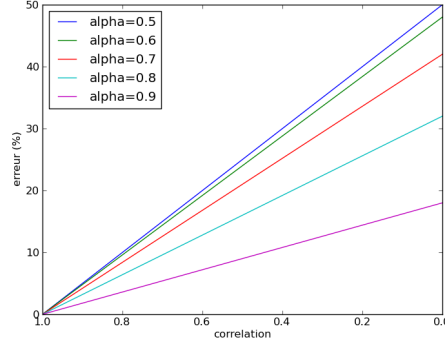


FIG. 2 – Example of error assuming that the standard deviation diagnosed in the observation space is equal to the grid point standard deviation, as a function of the correlation between the neighbor grid point and the location of the observation.

- ▷ **Referee: p883** : authors explain that the estimate of the background error correlation is out of the scope of the paper. It is a bit disappointing for the reader but can you at least discuss the way it could impact the results of the assimilation ? Have we some insight on the respective weight of variances and correlation formulation on the assimilation results ? Is it a way that could explain that results remains close what ever your BECM choice is ? If the answer is yes maybe it should be mention in the conclusion.

The estimate of the background error correlations is within the scope of the paper, for sure. But using the \mathbf{HBH}^T matrix to achieve this goal is out of the scope. As we estimated the background error correlations using an ensemble method and we assess their impact on the analysis (comparing for example Oper and Lxy experiment), we don't really understand the comment of the reviewer.

- ▷ **Referee: p 887 section 4.1** : Here authors are describing how their system (ensemble and choices for OECM and BECM) is built. Some choices should maybe be explained such like 50% for emissions perturbations, 45km for length-scales, why only emissions are perturbed ? is it enough ? Probably that reference to the existing literature would help (Hanea, Boynard, Garaud, Mallet). Moreover, the results of previous analysis tend to show that OECM > BECM, is it meaning that the system tends to have more confidence in the model than in the observations ? Could authors discuss this ? Lastly for this section, authors explain that ocean emissions are missing ? Considering the amount of NOx emissions by ships is it not a problem (especially along coastal areas) ?

Aside from the observations, we chose to perturb only the emissions and not other parameters because of computational cost considerations. The other things we could have perturbed are the meteorological forcing (winds, temperature, ...) or the chemical reaction rates. Concerning the me-

teology, it is not so simple to generate perturbations that are coherent with the flow properties (like the mass conservation). Concerning the chemical reaction rates, it's difficult to have an uncertainty value on them. So we chose the simplest way as we knew that emissions (in particular NO_x emissions) influence ozone concentrations. The aim of this study was also to know if it's worth spending computational time of an ensemble method to diagnose the BECM. And if the answer would be positive, we would spend time on creating a better ensemble by perturbing other parameters.

Moreover, it was difficult to select which emissions we can perturb or not in order to have an impact on ozone. NO_x emissions were selected as mandatory but we could not modify these emissions without modifying others because for example, why a vehicle would emit twice the amount of NO_x and the same amount of CO when the ratio NO_x/CO is well known for this vehicle. So we decided to modify the emissions all together and a standard deviation of 50% is not so far from what we can find for some emissions. And this value was validated a posteriori thence the standard deviation obtained with the ensemble approach presents similar values than the ones obtained with the a posteriori diagnostics.

Few studies have been investigated on the appropriate length-scales to be used in air-quality studies. Chai et al. (2007) have shown the correlation length-scale cannot be greater than four times the grid spacing in order to avoid the illconditioning problem (i.e. approximatively not greater than 100 km for Mocage). Considering the ozone dispersion in the troposphere, we have chosen to have some 45 km horizontal length-scales for the first analysis. This value corresponds approximatively to two contiguous grid cells of the Mocage model.

The estimate of length-scales obtained with the ensemble method shows that 45 km is an appropriate value for surface ozone assimilation and it is a little underestimating in the North-South direction during the summer period.

A posteriori diagnostics done on a previous analysis give $\text{OECM} > \text{BECM}$. These values of BECM and OECM are taken to initialize our data assimilation system. This means that it tends to have more confidence in the model than in the observations. The analyses obtained from the ensemble method show the same tendency and proved that the values used as initial conditions for the error covariance matrices are adapted to our study.

- ▷ **Referee: p887-888-889 section 4.2 : Authors consider a 10 days period for their study. Such a period do not sample a statistically representative number of synoptic situations. Now, we do suspect that synoptic situations (stream direction, cloud cover . . .) could modify the length-scales. I suggest that authors discuss this in this section. Maybe, it would be interesting to know the meteorological conditions of this period. It could help to interpret more deeply the results.**

As the reviewer suggested, we were interested in the meteorological conditions. The surface pressure and surface velocity are represented in Fig. 3 of this answer for the two periods under study. These meteorological fields confirm that high length-scales values are located in areas where a strong wind dominates (i.e. over Ireland, Eastern-Europe and in the vicinity of domain boundaries). The length-scales also depend on the wind direction : the highest North-South length-scales are located in areas where a North-South wind appears (i.e. in the Rhone valley or over Eastern-Europe).

We have concluded that the length-scale isotropy is lost during the summer period due to the photochemistry effect. This hypothesis is confirmed since an anticyclone situation appears during this period over a large part of Europe.

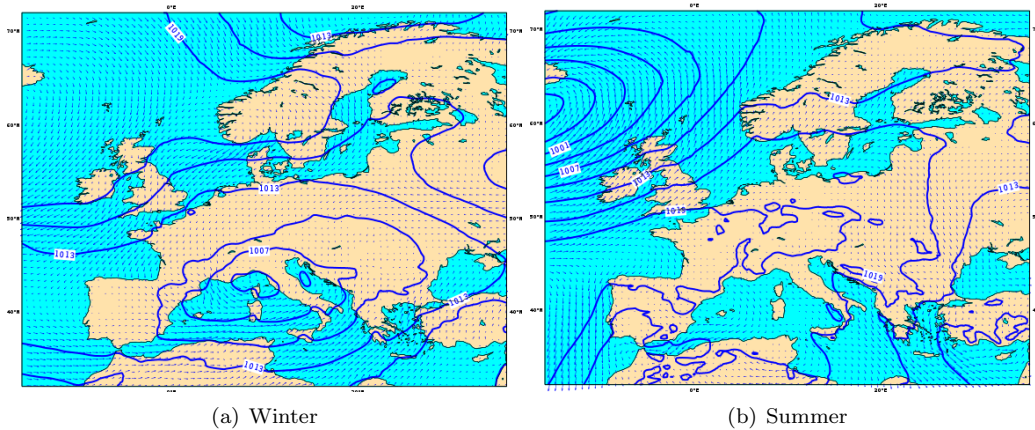


FIG. 3 – Time average of the surface pressure (contour lines) and surface velocity (vectors) for the two periods under study.

Figure 3 of this answer does not appear in the new version of the manuscript but we have completed some paragraphs of this section :

Paragraph 4.2, page 888, from line 9 to 25 : “(...) The values of the length-scales depend on many parameters such as wind, orography, type of grid cell (i.e. rural, urban, suburban). For example results show that the largest length-scales are located in areas where a strong wind appears (i.e. in the Rhone valley or over Eastern-Europe, Figure not shown). The isotropy is lost during summer. (...) ”

Paragraph 4.2, page 888, from line 26 (already modified in a previous response) : “(...) Note that we also have high values for the length-scale in the Western part of the Strait of Gibraltar. There the ensemble of simulations are probably similar due to the dynamics, the strong wind advecting ozone from the Mediterranean Sea (Figure not shown). ”

Paragraph 4.3, page 889, from line 24 : “ (...) Larger standard deviations during the summer season could be due to high ozone concentration and variability. The anticyclonic condition over Europe during this period (Figure not shown) leads to this chemical situation. A diurnal variation is observed during summer with an increase during the day. During winter, the variations of the standard deviations are smooth and the diurnal cycle is not visible (Fig. 7). ”

- ▷ Referee: p892 line1-11 : It will improve the section 5.2 if authors could recall (briefly) why they use this Joly and Peuch classification instead of the metadata classification proposed classically by airbase. I think it is obvious for the authors but it could be an

interesting point for less skilled readers.

We added a new explanation in the following paragraph :

“From the MACC ozone ground-based stations presented in section 3, a selection of station was done. A classification by type of stations (i.e. urban, suburban and rural) has been developed at Centre National de Recherches Météorologiques (CNRM) by Joly and Peuch (2012). This classification in ten classes is based on the measurement data itself, using all the data available in the European AirBase data set and the French data set named BDQA (Base De données de Qualité de l’Air, i.e. Air Quality Data Base). Each class allows to group past time series of measured pollutants that are homogeneous from the point of view of their statistical properties. Thanks to this classification we retain exclusively rural stations in the assimilation process. Rural stations give a relevant representation of large-scale conditions because they are less influenced by local phenomena (e.g. emissions)”.

- ▷ **Referee: p893 line 25 : Why authors do choose this particular date ? Is it representative of the whole period ? Is is a particular situation ?**

This date is representative of the whole summer period at this time. This time has been chosen because highest ozone concentration are observed at 14 :00 :00 UTC on polluted places. It is then simpler to compare the different analyses.

- ▷ **Referee: p895 line 8-10 : I think that it would be a real plus for the paper if authors could discuss the fact that it is difficult to distinguish the best configuration of the BECM. Do authors have any leads to explain that ? You do not consider different formulation of the correlation, could it be important ? Other things ? ? ? Whatever author can propose maybe they can emphasize more on the fact that there are few differences in their results as a function of the BECM choice. Indeed, it is valuable information for other group working on such topic.**

In order to estimate the correlation, we have studied two formulations : the currently used Gaussian length-scales formula and the non Gaussian formula from Belo Pereira and Berre (detailed in the section 2.3). As results from these two methods provided similar length-scales values, we have only presented the results from the Gaussian length-scales formula. In this study, the choice of the correlation formulation do not have an impact on analyses from the different configurations of the BECM.

The choice of the validation station locations can probably be improved. Few validation stations are located in the area where the ozone concentrations differ significantly between the analyses from the different configurations of the BECM (surface ozone concentrations show in Fig. 9 for the summer period). Choosing more validation stations in the polluted areas could provide more differences in the statistics of the studied configurations of the BECM and a configuration could be distinguished as the best configuration.

The authors could advise to other group working on such topic to use the Oper configuration during winter (i.e the configuration with a fixed horizontal length-scale and a standard deviation derived from a posteriori diagnostics). Indeed this configuration gives some good statistics and its computational cost is negligible. This configuration is daily used at Météo-France within the European

MACC project and provides some pertinent analyses. During summer, the authors could advise to use a fully time-dependent BECM (i.e. the StdLxy configuration). We have shown that the summer photochemical effects tends to produce anisotropic length-scales and a diurnal variation of standard deviations is observed.

- ▷ **Referee: Section 5.5 : I am surprised that authors choose so few validation stations. How the choice of the stations has been done? I can understand this choice for operational set-up but for this study it appears not so well appropriate. Authors should discuss this point. We can imagine that if the station network is dense enough the choice you are making to prescribe properly the BECM especially concerning length-scales is useless. This will be impossible to verify if validation stations are few and located just in the neighbourhood of assimilation stations. It would be interesting to consider aspect.**

The ozone ground-based stations used in this study are the ones used within the European MACC project. When this study began, a list of validation stations was not available within MACC. To select the validation stations, we have chosen to take the ones which are the most represented in a grid cell. As the rural station network is not dense everywhere in Europe (Fig. 8), we wanted to avoid removing isolated stations from the assimilated stations (as the ones in Spain, Sweden or Norway). It could be effectively better to use validation stations far from assimilated stations in order to have independent statistics. But the rural station network given do not let us having a better choice.

- ▷ **Referee: Section 5.6 : This section as well as the corresponding conclusions missed comparisons with previous studies of this type. Considering papers of Blond, Elbern will help authors to strengthen their analysis and conclusions on these forecasting aspects.**

We added new references as suggested by the reviewer in the last paragraph in section 5.6 : “ *Improvements in ozone forecasts subsequent to the assimilation procedure were found as in the studies of Elbern and Schmidt (2001) and Blond and Vautard (2004). In our study, results indicate that the impact of the assimilation process persists longer during the winter period than the summer period : one day in summer and about three days in winter (Fig. 11). After that, forecasts from the five experiments from the assimilations are very similar to the DIRECT simulation. These results are consistent with the study of Blond and Vautard (2004) which shows that ozone analysis initialization improves the simulation for 24-48 hours later and beyond that time the analysis becomes identical to the simulation without assimilation. During the first 12 h, the influence of our assimilation is rapidly reduced (...)*”.

Reply to Referee #2 technical corrections

- ▷ **Referee: p875 line 7 : ”... the best estimate of the physical state given the input ... ”**
Done.
- ▷ **Referee: p877 line 22 : observation instead of observations**
Done.

- ▷ **Referee: p881 line 22 : flow-dependent instead of flow-independent ?**
Yes, it is flow-dependent.
- ▷ **Referee: p884 line 5 : maybe a web site or a reference on the MACC project is needed.**
Done.
- ▷ **Referee: p884 line12 : I understand that authors do not spend too much time on the description of tropospheric ozone but maybe that an "English" reference could also be added there such like the books of Seinfeld & Pandis or Finlayson-Pitts or Warneck.**
Done.
- ▷ **Referee: p885 line 9-19 : Can you specified the reference there where the assimilation system your are using is described in details ?**
Done.
- ▷ **Referee: p885 line 16-17 : "... measurements are available at hourly frequency from the MACC project, ..."**
Done.
- ▷ **Referee: p885 line 21 : If you do specify the web site and more specifically the link toward the informations about ground stations, this sentence is almost useless.**
Done.
- ▷ **Referee: P885 line23 : I do not think that you need to show the station's location for both summer and winter. One of these two figure is enough. If you want to underline differences between winter and summer (is it significative, I don't think though ?) just tell it or show a map of differences.**
Only an example of summer stations location is shown.
- ▷ **Referee: p885 You should probably mention what kind of ground stations are used and from which database.**
Done, it's airbase stations.
- ▷ **Referee: p894 line 10 : a word is missing after quite ?**
The word was added and the new sentence is : "... *distribution of ozone over the Alps and urban cities is also quite similar when the ensemble-based ...*"

Références

Blond, N. and Vautard, R. : Three-dimensional ozone analyses and their use for short-term ozone forecasts, J. Geophys. Res, 109, D17 303, 2004.

Blond, N., Bel, L., and Vautard, R. : Three-dimensional ozone data analysis with an air quality model over the Paris area, J. Geophys. Res, 108, 1993–1996, 2003.

- Buehner, M. : Ensemble-derived stationary and flow-dependent background-error covariances : Evaluation in a quasi-operational NWP setting, *Quarterly Journal of the Royal Meteorological Society*, 131, 1013–1043, 2005.
- Chai, T., Carmichael, G., Tang, Y., Sandu, A., Hardesty, M., Pilewskie, P., Whitlow, S., Browell, E., Avery, M., Nédélec, P., et al. : Four-dimensional data assimilation experiments with International Consortium for Atmospheric Research on Transport and Transformation ozone measurements, *J. Geophys. Res.*, 112, D12S15, 2007.
- Coman, A., Foret, G., Beekmann, M., Eremenko, M., Dufour, G., Gaubert, B., Ung, A., Schmechtig, C., Flaud, J.-M., and Bergametti, G. : Assimilation of IASI partial tropospheric columns with an Ensemble Kalman Filter over Europe, *Atmospheric Chemistry and Physics Discussions*, 11, 26 943–26 997, 2011.
- Constantinescu, E., Sandu, A., Chai, T., and Carmichael, G. : Ensemble-based chemical data assimilation. I : General approach, *QJR Meteorol. Soc.*, 133, 1229–1243, 2007.
- Desroziers, G., Berre, L., Pannekoucke, O., Stefanescu, S., Brousseau, P., Auger, L., Chapnik, B., and Raynaud, L. : Flow-dependent error covariances from variational assimilation ensembles on global and regional domains, *Hirlam Technical Report Y*, 68, 2–22, 2008.
- Elbern, H. and Schmidt, H. : Ozone episode analysis by four-dimensional variational chemistry data assimilation, *Journal of Geophysical Research*, 106, 3569–3590, 2001.
- Elbern, H., Schwinger, J., and Botchorishvili, R. : Chemical state estimation for the middle atmosphere by four-dimensional variational data assimilation : System configuration, *J. Geophys. Res.*, 115, D06 302, 2010.
- Joly, M. and Peuch, V.-H. : Objective classification of air quality monitoring sites over Europe, *Atmospheric Environment*, 47, 111–123, 2012.
- Massart, S., Pajot, B., Piacentini, A., and Pannekoucke, O. : On the merits of using a 3D-FGAT assimilation scheme with an outer loop for atmospheric situations governed by transport, *Monthly Weather Review*, 2010.
- Massart, S., Piacentini, A., and Pannekoucke, O. : Importance of using ensemble estimated background error covariances for the quality of atmospheric ozone analyses, *Quart. J. Roy. Meteor. Soc.*, 2011.