Anonymous Referee #3

Received and published: 26 August 2013

This study by A. Robichaud and R. Ménard presents the results of a multiyear Objective Analysis of the warm season surface ozone and particulate matter of diameter less than 2.5 microns (PM2.5). The first part of the paper describes the background, motivation, data and the methodology in some detail. The final few sections focus on some results: a nice trend analysis, and correlations with meteorological and economical indices. I especially like the trend analysis that shows the decrease of ozone and PM2.5 at high percentiles, indicative of emission reductions and an increase of background (low percentile) values. These results are not entirely new or surprising but they do corroborate and extend earlier findings (e.g. Cooper et al 2010) using a different methodology. In addition, the paper makes a solid case for Optimal Interpolation. The mathematics appears sound but a few items could be made clearer and there are a few minor mistakes, which I list in Specific comments below. Because of the methodology and scientifically important results, the paper should be of interest to the ACP readers, however, I have found some areas that need some serious editing.

General comments

Reviewer no. 3

I would put more emphasis on the importance of the cross-validation results. Basically what they show is that this OA yields good results in areas where there are no observations(but not too far away from observations). This builds a case for using OA rather than calculating trends directly from measurements. This is mentioned in the text but it should be emphasized more.

Reply from authors

These are very good points and we agree that they deserve to be emphasized more. This will be done with pleasure in the revised version of the paper and your comments and suggestions will also be added. Thank you for pointing out these important aspects of our work.

Reviewer no. 3

I would move Section 4 up before current Section 3 and state clearly why this validation is important, in the first paragraph (instead of just saying that it is required).

Reply from authors

We accept this change and it will be done with pleasure in the revised version. Note that reviewer no. 1 also requested that change. We will also emphasize more the purpose of the cross-validation in the revised text.

Reviewer no. 3

Since this impacts the entire analysis, please discuss the validity of the Gaussian assumption for errors. For example, the mixing ratios for ozone at the surface can be expected to be comparable to the standard deviations of the errors, in which case the Gaussian assumption fails. Figure 3 shows that the random model error can be over 15 ppbv. Isn't it comparable with the mean background surface ozone in rural areas?

Reply from authors
First as a point of clarification, in the manuscript, we are not talking about the distribution of the observed values but rather about the distribution of the observation errors, and likewise about the model forecast errors. The most reliable source of information for these errors is given by analyzing residuals O-P (Observation minus forecast). However, the O-P is a combination of observation error and model forecast error taken together. It is true that by inspection of the spatial autocorrelation of the O-P we can distinguish which part of the variance arises from observation error and which part is attributed to model forecast error, but we have no direct evidence of the distribution of model forecast error or of observation error separately, only of these two errors combined (in the O-P). The distribution of O-Ps for ozone and PM2.5 for instance was found to have roughly a near Gaussian shape but with more heavy tails or more leptokurtic than real Gaussian distribution. However, and as mentioned in the section 2.1 of the original manuscript, the analysis can still be derived without assuming Gaussian distributed errors, but from using the BLUE (best linear unbiased estimation) principle and assuming an analysis linear in y_o and x_f (Eq. 1 in the text). The significant reduction of variance due to the analysis (as compared to the model random error) is also verified in practice (see Section 4). For example, for ozone, the O-A have much lower variance than the O-P of about a factor of 3-4 indicating the success of the methodology (see Fig 3A).

Reviewer no. 3
English. There are a lot of ungrammatical sentences and style issues, especially in the first few sections. The manuscript should be seriously edited for language. I don't feel particularly competent here but I point out some of the most glaring problems in Specific comments. There are many more.

Reply from authors
We agree to completely review the document with the help of a English native speaker and will pay attention to making better grammatical sentences and correct the style wherever it is needed.

Reviewer no. 3
Figures. The fonts are too small. Color bars are impossible to read without some serious magnification. Essentially, almost all the figures should be edited and made more readable

Reply from authors
The figures will be re-done paying very much attention to the above details and magnifying all color bars and fonts. In the revised version, we will make sure that even after transformation for publishing purposes, this will not happen. We apologize for that and this will be corrected.

Specific comments

Reviewer no. 3
P 13970 L26. 'to provide to the public and health specialists with'. Drop the second 'to'

Reply from authors
This will be done easily with pleasure.

Reviewer no. 3
P 13974. L2. Reference to Menard and Robichaud (2005). This is a seminar talk. It would be better to cite peer-reviewed material if available

Reply from authors
The above reference (Ménard and Robichaud, 2005) is part of a ECMWF seminar and this, in our mind, is equivalent to a peer-reviewed manuscript because of the general high reputation of the ECMWF seminar series. In any case, up to now, there has been no official peer-reviewed material to refer aside from this seminar talk. Note that the concepts discussed in the above reference concern general concept of objective analysis which are not new and have been developed and used routinely in meteorology for several decades.

Reviewer no. 3
P13974, L5: 'analysis matrix' is fine but "analysis vector" would be better with the terminology you use later on.

Reply from authors
This modification will be done with pleasure.

Reviewer no. 3
P13975 L20 "Finally,σ _f and Lc represent respectively the background error variance and the correlation length". Shouldn't it be "standard deviation"?

Reply from authors
Yes, reviewer no. 3 is right and this will be corrected with pleasure.

Reviewer no. 3
P13975. Equation (3) needs a reference and some justification, e.g. why this gives a positive semi-definite background error covariance. Maybe Gaspari and Cohn 1998.

Reply from authors
We are glad that the reviewer no. 3 brings this point. Equations 3 and 4 are actually the working assumption that is used in many optimum interpolation schemes. In this paper we use a first order autoregressive model as a correlation model but in principle it can be any correlation models. See for example Menard 2000, Daley 1991 (ref. Chap 4) or Kalnay 2003 (ref. section 5.4). Also since the exponential model (or first-order autoregressive model) is positive definite so is the r.h.s. of Eqn 2 through Eqn 4.

Reviewer no. 3
P 13978, Eq. (11). Why are you making this assumption? Please, explain, as I don't think this is discussed in Menard 2010. An exponential decay of the model bias like this could result in essentially zero bias in between regions if they are sufficiently far away. Why were these particular values chosen for a and b? If there are only four elliptical regions then a map with the region boundaries would help visualize what's going on.

Reply from authors

Reviewer no. 3 is correct. Equation (11) was not discussed in Menard 2010, or in anywhere else. This is a new aspect related to the submitted paper. Let us first discuss about Figure 5 of the original manuscript. It clearly shows that analysis increments (related to model systematic biases) have the tendency to form an organized structure with maximum (negative or positive) in eastern U.S. and southwestern U.S. for both ozone and PM2.5. Two criteria were used to define the ellipse position and parameters (e.g. values of a and b): 1) the ellipses were centered where the density of stations are higher, 2) highest model biases (e.g. time-averaged analysis increments) where located inside the ellipses so that modeling of bias correction outside the ellipses are not critical and has little impact. The 4 ellipses are centered respectively over the southeastern and southwestern Canada, eastern and southwest portions of U.S. in such a way that the bias correction be effective over these locations and have significant influence on the validation. This is an empirical bias correction which was tailored for the problem under study. We are not claiming that this is a general bias correction algorithm. Note, however, that the adaptive scheme (Eq. 14 and 16 of the manuscript) is applied before the bias correction scheme in our methodology so a significant part of the bias correction has also been achieved by the adaptive scheme. The bias correction scheme Eq. 11 was therefore only applied to correct the residual bias left over. The values chosen for a and b were simply obtained through trial and errors so that they reduced the bias as much as possible in the cross-validation mode. Note finally, that values of the seasonal average bias correction is constant through the region (within a given ellipse) but varies with time of the day and changes with season. In the revised version, we agree to provide more details since this is a new feature. In figure 2 of the revised paper, we will also superimpose the location of elliptical regions on top of the observation network to facilitate the understanding.

Reviewer no. 3

P 13979, L12 'calculate' ! calculating
P 13979, L23 'Goddard Space Flight' ! Goddard Space Flight Center

Reply from authors

Thank you for this; we will introduce all corrections accordingly.

Reviewer no. 3

P 13980, Eqs. (12) & (13). Introducing coefficients alpha and beta seems unnecessary. They just represent the fact that L and sigma can be tuned by scaling. Otherwise, they are the same as Eqs (3) and (4). You could simply state that you are tuning the two parameters in (3) and (4), thus avoiding the repetition and making the presentation more concise.

Reply from authors

We agree with this change which makes the paper more concise by eliminating Eqn 12 and 13. This will be done accordingly.

Reviewer no. 3

P 13980, If Eq. (13) ends p being used after all I think the l.h.s. should read (HB)ˆt, not H(HB)ˆt. Also, 'T' and 't' are used interchangeably to indicate transpose. I would just stick to 'T'.

Reply from authors
Reviewer no. 3 is correct, this was just a typo error. In any case, we agree to remove Eqns 12 and 13 so that this is no longer a problem.

Reviewer no. 3
P 13980, Eqs. (14) - (16):
1) It's not clear to me why applying (14) would produce a Kalman gain that will bring the chi-square statistic closer to 1.0. Please explain. It's intuitively clear why Eq. (16) would do that (if chi_sq > 1 then an increased error variance should bring it down and vice cersa). Equivalently, why does the sequence in (14) converge?
2) 'until there is convergence or until chi-sq _ 1'. I think convergence will be there if and only if chi-squared converges to 1, so it's not either or. I would like to see more detailed justification for this adaptive scheme

Reply from authors
We will add up ANNEX A to address this (included in this reply, see below). The new ANNEX will give mathematical details for the reason why the formulation (14) works out: e.g. because a) it allows the value of chi-squared equals 1 as a possible fixed-point solution and b) Eq. 14 is a contracting transformation for an initial chi-squared greater than one and for such transformation, Banach (1922) demonstrated that it has a fixed-point which is unique so that the transformation converges. If the convergent value is not 1, then Eq. 16 (inflating the background error variance) is applied which brings chi-squared to 1. For initial chi-squared less than one, the scaling Eq. 14 inflates the correlation length until chi-squared reach 1. So in this particular case which happens less often, this is not a converging but rather forcing correlation length so that chi-squared =1. The rationale for having chi-squared=1 is based on the findings of Menard (2000) where it is demonstrated that this gives optimal error statistics. The scheme was clearly demonstrated to work properly in cross validation mode for both ozone and PM2.5 for all the regions and all hours on figure 6 (year 2005 for ozone) and figure 7 (year 2007 for PM2.5). Future research needs to be done to obtain a general theoretical framework on how to tune correlation length. Note however that the contraction of the correlation length have been done by other authors in a similar context such as Frydendall et al (2009) (their Eq.4) who modified the correlation length obtained from the Hollingsworth and Lonnberg`s method (1986) to take into account the density of station of a surface pollutant measurement network. Our procedure, although very different from Frydendall et al (2009), goes along the same line in the case that the initial chi-squared is greater than 1 (e.g. results improve significantly when the correlation length from H-L is reduced by a factor from 1 to 5).

Reviewer no. 3
P 13981 L1. 'If needed only if' ! drop the first 'if'
P 13980, L18. 'more closer' ! closer

Reply from authors
These will be corrected with pleasure.

Reviewer no. 3
Section 2.4. Expand the description of the two models: the main relevant chemical reactions, the number of species modeled, emission inventories, etc.

Reply from authors
We will do it with pleasure in the revised version.

Reviewer no. 3
P 13984. "the standard deviation of the observation error including the representativeness error is believed to be higher than 5 ppbv". Please, state the typical mixing ratios in regions of high and low concentrations (industrial and rural). The same for PM2.5. This will help judging the validity of the Gaussian errors assumption, especially in areas of low concentrations.

Reply from authors
The reviewer seems to work on the premise that the Gaussian assumption is erroneous when the errors are large. As we explained earlier, the Gaussian behavior is not essential to derive the analysis equation. Nevertheless, we did observe a fairly Gaussian distributed O-P but with more heavy tails in the case of ozone. For PM2.5, a more leptokurtic behavior was observed. We cannot judge if this is attributable to observation error or model forecast errors. However, even with large biases (see Figure 3), the OA random and systematic error were reduced quite significantly for both ozone and PM2.5.

Reviewer no. 3
P13985 L17. Avoid using "etc". It's better to spell things out
P 13976, 25. A reference to Wikipedia? Some linear algebra textbook would be more suitable
P13974, L5 "analysis increments" ! "analysis increment" "could be view" ! "could be viewed"

Reply from authors
These specific comments will all be corrected accordingly.

Reviewer no. 3
P13986 L5 State what satellite

Reply from authors
The satellite image is a composite of satellites Aqua and Terra (the instrument is MODIS). This will be mentioned in the revised version.

Reviewer no. 3
Figure 7. Can you explain why there is an apparent seasonal dependence of the bias? Section 5.1.1

Reply from authors
We do not present any seasonal dependency of the bias because our study covers only the warm season with a focus on summer months (June-July-August). Note, however, that there is a strong diurnal behavior for ozone.

Reviewer no. 3
P13992, L14 The difference between two years of the analysis cannot be taken as evidence for a trend, even if these years are 'similar'. I suggest starting with actual trend analysis (Section 5.2) and then discussing the 2012 – 2005 difference as an illustration and without calling it itself a trend. That would mean swapping Sections 5.1

and 5.2 and the corresponding tables. It's important to get this right. The trend analysis and the regression on various predictors are the two key results of the study.

Reply from authors
We entirely agree with this idea. We will drop the reference to the word *trend* for the difference between 2012 and 2005 and correct accordingly by swapping section 5.1 and 5.2.

Reviewer no. 3
P13993 L8. 'positive trend becomes significant' – significant in what sense? Statistically? You can't claim that based on a two year difference.

Reply from authors
These are differences between two years and we wanted to say "meaningful" not significant (in the statistical sense). We will correct this mistake accordingly.

Reviewer no. 3
P13995, L19 onward. Since there are multiple predictors it is appropriate to do multiple regression analysis as it is done on PP13997+. I'm not sure if there is a point in analyzing pairwise correlations with meteorological and economic indices separately. I would start with multiple regression, base the entire analysis in this section on its results, and drop the individual correlations altogether. This would make the section more compact, easier to follow and less redundant. Also note that some indices are not independent (temperature and precipitation) as shown in Table 7. Again, There are some important findings there and they should not be buried under a lot of redundant numbers.

Reply from authors
We agree to drop table 7 and perhaps rather expand table 8 to present results for each individual region (Canada East, West and U.S. East and West). This would provide more information and diminish redundancy. Thank you for pointing that out to us.

Reviewer no. 3
P13997 , L10. 'A multiple regression model using a stepwise-like procedure'. Please be more specific and describe the procedure in some detail, if possible provide a reference, not just the name of the algorithm.

Reply from authors
We will do so in the revised version. The statistical package SAS (version 9.2) was used and the procedure used is STEPWISE which is a fairly standard procedure in the field of multiple regression analysis. However, we agree to give more details in the revised paper.

Reviewer no. 3
P14002, L6 'estimate of the two main components of smog' I would say 'two of the main components'. There are other main components such as NOx

Reply from authors
We will correct this with pleasure.

ANNEX A. Mathematical notes related to the adaptive scheme (Eq. 14 and 16)

The adaptive scheme of section 2.1 for the correlation length (Eq. 14) and the background error covariance (Eq. 16) have been developed based on trial and error. However, some connection to the theory could be done and are given here. In all generality, we can express the original correlation length obtained from the Hollingsworth and Lönnberg method (1986), L_c^n as:

$$L_c^n = f(X^2_n) \quad \text{(B.1)}$$

with $X^2 = \chi^2/p$ (B.2)

therefore, the iterative scheme (called adaptive scheme in the text, Eq. 14) can then be written as:

$$f(X^2_{n+1}) = f(X^2_n)/X^2 \quad \text{(B.3)}$$

A fixed-point of our iterative scheme is such that

$$\lim_{n \to \infty} X^2 = X^{*2}$$

Therefore, B.3 becomes

$$X^{*2} f(X^{*2}) = f(X^{*2}) \quad \text{(B.4)}$$

where X^{*2} is the fixed-point**.** One of the possible solution is $X^{*2} = 1$ and that was the solution often obtained experimentally in our study for intial $X^2 > 1$. The reader is referred to Collet and Eckmann (1980) for more information on iterative procedure for similar problems.

On the other hand, we now demonstrate that when the transformation is contracting, the solution is convergent. Let us introduce the transformation T

$$T(X^2_n)=X^2_{n+1} \quad \text{(B.5)}$$

where as in Eq. 14

$$T(X^2_n) = \frac{L_c^n}{\left(\frac{\chi_n^2}{p}\right)} \quad \text{(B.6)}$$

By elementary inspection, we can verify that transformation T is contracting whenever B.6 is used in situation when the initial value X^2_n (the denominator of B.6) is greater than 1. Therefore, we can write,

$$| T(x) - T(y) | \leq q \, | x - y | \quad (B.7)$$

where q is the Lipschitz constant.

Based on Banach fixed point theorem (Banach, 1922), the iterative procedure B.6 has therefore a unique fixed-point $\mathbf{X}^{*2}$**.** This fact proves convergence of the adaptive procedure of Eq. 14.

Finally, the fixed-point value does not always corresponds to values of

$$\mathbf{X}^{*2} = 1 \quad (B.8)$$

nor the initial value of X^2 is not always greater or equal than 1. In the first case, the procedure of Eq. 16 is used to scale the background error covariance (inflation procedure) to achieve the optimal condition $X^2=1$ by iterative convergence of Eq. 16. In the second case (initial chi-squared less than 1, which was less commonly found), the initial correlation length is inflated iteratively using Eq. 14 until $\mathbf{X}^{*2} = 1$.

Additional references to the manuscript

Banach, S : Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. *Fund. Math.* 3(1922), 133–181.

Collet, P., and Eckmann, J.-P.: Iterated maps on the interval as dynamical systems. Progress in Physics. A.Jaffe and D.Ruelle, Eds., Birkhäuser, 248-270.

Frydendall, J., Brandt, J., and Christensen, J.H.: Implementation and testing of a simple data assimilation algorithm in the regional air pollution forecast model, DEOM, Atmos. Chem. Phys., 9, 5475-5488, 2009.