

Interactive comment on “Estimation of aerosol particle number distribution with Kalman filtering – Part 2: Simultaneous use of DMPS, APS and nephelometer measurements” by T. Viskari et al.

T. Viskari et al.

toni.viskari@fmi.fi

Received and published: 7 November 2012

Dear anonymous referee,

At the very beginning we wish to apologize for the late response to your comments. We found your comments to be thoughtful and well-argued. We are also happy to hear that you consider the research scientifically important. We find that your comments have greatly improved our manuscript.

We begin with the major complaint, which is the lack of validation by independent observations. The reason for this is that we simply did not consider it to be feasible. It is not possible to know the true state and all other observations concerning the state

C9073

would contain their own uncertainties. For example, DMPS I and II measure same variable over partially same particle sizes, yet the two instruments commonly measure different values in their overlapping measurement range. Furthermore, as the state estimate is relatively close to the observed state, it would be difficult to determine whether differences between the state estimate and possible independent observations were due to flawed observations. With these limitations in mind we considered that our only option was to compare the state estimate x_{EKF} to the raw observations y and decided that the residual $r=y-Hx_{EKF}$ has to meet two conditions: i) the bias and standard deviation of r has to be in the equal or better than the residual computed from the mathematical inversion, ii) large values of r are either due to measurement noise or special circumstances (e.g., precipitation, change of air mass). This validation is subjective, but we argue that the set standards are reasonable for validating the estimate. We have added to the first part of the article further clarification of these standards.

As for the combination of the two articles, we would rather not do that. The first article serves as the theoretical introduction to the method and does, in our opinion, focus on the statistical challenges of the aerosol size distributions, which are independent of chosen measurement instruments. The second part is, in contrast, a more practical application of EKF, which focuses on the specific issues concerning the chosen instruments. Even if the articles were to be combined, both issues would still have to be presented and studied. Thus we feel that the resulting article would be simply too large and difficult to follow.

The specific comments are answered below:

“Page 18893, section 2.1: The observation error information of DMPS should be described here, not in Chapter 4.”

We added that information here. We also added a brief sentence at the beginning of section 2 about our assumptions concerning the observation error.

C9074

“Page 18893, section 2.2: The observation error information of APS should be described here, not in the result chapter.”

We added that information to this section. Looking the manuscript over again, it is clear that this improves the presentation.

“Page 18895, line 1-4: The raw measurement of APS is aerodynamic diameters, isn't it? If so, why does the observation operator for APS calculate geometric diameters from aerodynamic diameters? My understanding is that geometric diameters are model variables. Is this right? The input data of the observation operator is model state vector x . What variables are included in your model state vector x ?”

Your questions are understandable, as we present the equation in the form it is generally shown, but of course the observation operator calculated aerodynamic diameters from the geometric diameters. We changed the formula to show that the observation operator calculates the aerodynamic diameters. The model state vector x contains the particle number concentrations as a function of the geometric diameter.

“Page 18895, eq (3): N in this equation is explained nowhere in this paper. Even if it is trivial, all of the variables in equations should be explained in the text.”

You are correct in both matters and this was an oversight on our part. $N(d)$ represent the number concentration value as a function of particle size and we have added this to the paper.

“Page 18895, line 20 and 25: What is D_p ? Even if it is trivial, explain it in the text.”

Another oversight on our part. D_p is the same diameter than the d defined in the text. This has now been corrected.

“Page 18896, line 21-22: The observation error covariance is ignored in this paper, which is acceptable because this paper describes a preliminary study. However, this assumption is not realistic. I hope that the effect of the error covariance is investigated in your next study.”

C9075

You are correct that it is not realistic and our handling of the error covariances includes several simplifications and assumptions. It will be our next step to improve both the observation and the model error covariances as that would both further improve the state estimates and also allow us to better depict a realistic physical system.

“Page 18897, line 1-4: “The model error term is omitted, and thus the standard deviations are artificially kept at 20%.” This means that the B matrix at time k is not evolved to time $k+1$ at all? If so, this is not called “inflation” (it is just a fixed covariance). This description contradicts the description of B in the Part 1 paper. The B matrix was time-evolved in the Part 1 paper, wasn't it?”

The B -matrix does evolve over time and is determined similarly as in Part 1, we apologize about the confusing text here. We were referring the model error matrix Q presented in Part 1, but as we do not show the formulas here, it does appear a bit misdirecting. We removed the omitted part from the text and added that the inflation is in order to prevent the background state from being more reliable than the observations.

“Page 18897, section 3.2: Please describe clearly all of the forecast variables of UHMA model and the control variables of the data assimilation. It is very important information.”

You are correct. We added that both the control and forecast variables are the number concentration values. “Page 18898, line 8-9: I am sorry, but I do not understand the meaning of “the increase in computational cost is even larger due to the increased number of tangent-linear model evaluations”...”

Our apologies about that, it is a bit unclear. Essentially what we meant was that if the state vector has n components, then the error covariance matrix has $n \times n$ components. As the uncertainty propagation requires twice the number of vectors in calculations, it means that if the state vector has 50 components, then the error propagation requires 100 calculations. However, if the state vector has 60 components, then the error propagation requires 120 calculations.

C9076

We removed the increased number part and simply refer to the already increased computational cost of the EKF implementation.

“Page 18898, chapter 4: My understanding is that the authors defined only the relative errors but did not define the minimum errors (standard deviations) of each instrument. This seems unrealistic. Generally, even if an instrument observes zero values, its observation error never gets close to zero.”

You are correct in this, but unfortunately there are complications concerning that. EKF requires that the associated uncertainties can be considered to be Gaussian. The problem is that the number concentration value itself is limited as negative number concentration measurements are impossible and thus the uncertainties are actually Poissonian in distribution for very small number concentration values. For this reason we do include information from zero measurements in the current implementation. We are considering ways to find out for which number concentration values the uncertainties can no longer be considered Gaussian, but before that we have to limit the Poissonian nature of the very small number concentration measurements. This is partially done by not defining a minimum error.

“Page 18899, line 8: Please briefly describe how “inverted particle number size distribution” was inverted from the raw measurements. Of course, Virkkula et al. (2011) might describe it in detail, but this inversion process is very important information in this paper to compare with the EKF process.”

Good point. We included in the paper that the inverse size distribution is calculated with a least-square, non-negative pseudo-inverse method.

“Page 18901, line 5: Is “then” the erratum of “than”?”

Yes. Thank you for pointing that out.

“Page 18901, line 6: “This partially leads to the discontinuity over particle size at 10 μm visible.” <- “This” indicates what? Could you please explain the logic of the discontinuity

C9077

more carefully.”

We expanded the explanation that due to problems with the model, observations and error assumptions in those particle sizes, the state estimate is not reliable for particle sizes larger than 4 μm . This can be seen for instance in discontinuity over particle size at 10 μm .

“Page 18902, line 4-7: The large differences in the measurement values imply that both the instruments, or at least either one, have a large bias. In principle, it is impossible to assimilate biased data with non-biased data. Generally, largely biased data deteriorate the analysis, so that it is better not to use the biased data for data assimilation. Did you make consideration of the bias elimination for APS?”

For the measurements from DMPS II and APS, the large difference visible in the Figure 3 isn't because of bias. The measurements from DMPS II are for electrical mobility and the number concentration measured is a proportional value of absolute number concentration in those particle sizes. In contrast, the measurements from APS are the absolute number concentration values in those particle sizes. So naturally in this case it is very difficult directly compare the two, which was what we were trying to point out in the text. We have expanded the part slightly to make this clearer.

Furthermore, it is very difficult to be certain what bias causes large differences between measurements from different instruments. For example DMPS I and II often disagree in their overlapping measurement area and if examined statistically, it can be seen that there is a tendency for DMPS II to measure higher concentrations than DMPS I. In this case it is fair to say that it is probably due to calibration issues, but very difficult to say which if either instrument is correct. The fact that the overlap areas here are located at the edges of the measurement ranges does not help.

“Page 18903, line 26-28: I do not simply agree with your conclusion “the implementation of EKF to retrieve consistent results from a combination of number size distribution measurements and light scattering could result in improvements in data quality”. We

C9078

can see systematic errors (= bias) between observations and the EKF analysis in Figs 5a and 5c. Biased data are often harmful to data assimilation, so that careful validation is needed.

Before a discussion about the small difference in data assimilation results induced from nephelometer data, the authors should discuss a large bias between nephelometer's and DMPS's measurements."

You are correct that there is a systematic difference between observations and the calculated value for 450 and 700 nm wavelengths of radiation, although we would point out again that it is quite common for different measurement instruments to disagree in the overlapping measurement area. Also, while Virkkula et al. (2011) paper referenced in the paper only gives one refractive index, there should truthfully be a different refractive index for all the different wavelengths. This is partially visible in the scattering coefficients, as the refractive index for 450 nm should be slightly larger and for 700 nm slightly smaller than the used refractive index. This would result in estimated scattering coefficient for 450 nm wavelength of radiation to be slightly larger and for 700 nm slightly smaller. However, as we had already simplified our approach to the refractive index, we chose to keep the refractive index constant for all wavelengths. We have expanded the paper to include this information.

Additionally we think it important to note that the systematic difference for 450 nm wavelength of radiation is quite small and is likely the direct result of the error in the refractive index. For the radiation of 700 nm the situation is different, as the difference between the observed and estimated scattering coefficient value is too large to be explained solely by the refractive index. It is something we discussed during our testing. Based on our current understanding, the measurements for 700 nm radiation wavelength are more unreliable than for the other wavelengths and we considered giving it a higher unreliability value, but as our current uncertainties are already so simplified, we decided to keep it the same for this basic testing.

C9079

You are also correct that biased data does harm data assimilation estimates. However, it also important to note that removing observation data from data assimilation in the case it doesn't support the existing measurements is also questionable. Again, it is common for observations from different instruments to disagree with each other in the overlapping measurement range. Also, if we examine the increments shown in Figure 7 and 8, we can see that while nephelometer data does seem to disagree with the DMPS measurements, it actually overall supports the APS measurements. This is how information from integrating measurement instrument generally impacts the state estimate. It is especially worth of note that according to Fig. 5, the estimation is most affect for the wavelength of 550 nm, where the refractive index is most likely to be closest to the right one and where the estimated scattering coefficient is already close the observed values even without the inclusion of the nephelometer measurements. This indicates the size-segregating measurements greatly lessen the impact of the biased data.

To sum our response, it is important to remove biased observations, but before that can be done, there are several things that one must be certain of, for example how biased are the observations in question and that are the other observations completely without bias. This is a difficult subject which we think is very relevant in current aerosol research and to which we wish to return to in future research.

"Page 18903, line 19: What is "from 21:00LT to 20:40LT"? That means from 20:40 to 21:00? Or, another time? Is it shown in any figure?"

Here it means that we change the nephelometer time stamps here so that measurements from 21:00 LT were treated as if they had been made at 20:40 LT. We tried to change the text to reflect this better.

"Page 18905, line 2: According to Page 18904 line 20-23, the reason of the acceleration is the error of nephelometer's timestamps, and the reason of the deceleration is nephelometer's response delay. In either case, the nephelometer seems inadequate

C9080

to be assimilated with the status quo.”

Actually we did not say at Page 18904 lines 20-23 that we see both of those things happening, but that the difference is most likely due to one of those reasons. In either case, though, it does not show an issue with the nephelometer itself, but rather that there were difficulties with the specific measurements from that time. The implementation, on the other hand, had no difficulties in incorporating that information to the status quo and having impact on that state. This same difficulty could be faced with any measurement instrument and not just with the nephelometer. Rather, we were trying to point out the importance of being certain that the measurements are synchronized.

We included the nephelometer measurements here to show that EKF can also use information from integrating measurement instruments. There are problems with it, of course, but so there are with all instruments, even those that are similar. In the future we hope to better address the possible outlier through statistical means. We have already expanded in Part I on how we wish to achieve this.

“Page 18905, line 21-22: “We note that this approach will emphasize somewhat more the positive than negative increments when calculating the averages.” I am sorry, but I do not understand what this sentence means. This implies the existence of a large model-bias? If so, describe it explicitly.”

Our apologies for being somewhat misleading here. No, the scaling discussed here does not affect the state estimation, but rather the statistical results shown here. To give a very basic example, let us say that we have a background state value of 10 units and an observation value of 20 units, and both information sources are given equal weights. In this case increment would be 5 units and the state estimate value being 15 units, with the relative increment value of $5/10=0.5$. Now let us say during the next observation time the background state value would still be 15 units, but the observed state would now be 5 units. The difference would again be 10 units and the increment value would be -5 units, however this time the relative increment value would be $-5/15=$

C9081

0.33. Thus even though the increments are the same in value, the relative increments are different with the positive relative increment being larger than the negative relative increment.

This is a problem when calculating the statistical averages, as EKF itself would still in both situations change the value by 10 units. We changed the text slightly to stress that it only affects the statistical analysis of the results.

“Page 18907, line 14-15: What made you decide definitively that DMPS II is more unreliable than APS? Based on what?”

The measurements discussed here, in particle sizes 900-1000 nm, are from the last measurement size bins for DMPS II. It is also where the measurements are known to be very unreliable, which could be seen by going through the individual measurements. Of course we cannot decide this definitively and actually since we used the constant relative uncertainty for DMPS II, the implementation itself does trust these measurements the same as any other.

“Page 18907, line 17-18: The authors’ description “due to changes in the parameters, e.g. particle density” needs further explanation and reason. What made you think so?”

We have no way of knowing for certain, which is why did not expand upon it before. We know that it is a possibility as the parameters can be time-dependent and thus we mentioned it. We expanded on the latter by mentioning the time-dependency and that such parameters affect the observations operators.

“Figures 1a, 1b, 2a, 2b, 3, 4, 6a, 6b, 7, and 8: The unit of diameter should be unified between the text and figures as much as possible. Although nanometer or micrometer is used in the text, meter is used in figures.

Without tick marks, the visibility of logarithmic scale axes is very low. For example, it is extremely difficult to understand the location of 600-800 nm in Fig 1 at a glance. Please add the tick marks.”

C9082

We used the meters in the figures as we considered it to be clearer when using a logarithmic scale over such a large area of particle size. However in text it became easier to refer to 100 nanometers than to 10^{-7} m. In addition, as we have particles in the sizes of micrometers, it was difficult to settle for a single unit of diameter that was clear in all situations. That is why we chose, and still argue, for the use of meters in the figures and nano- or micrometers in the text.

As for the ticks, their lack is due to technical problems with the images. The ticks are very difficult to make out against the color scheme, a problem which is further increased by the fact that in the logarithmic scale most of the ticks are clustered close to each other. If the ticks were included in the figure, they were hard to distinguish from each other, especially when the image is reduced, and from the colors in the figure. Because of this we decided not to include them, as they actually hindered the figures.

“Caption of Fig 1a: #m-3 ?”

Thank you for pointing that out, corrected to $1/m^3$.

“Caption of Fig 2a: “Note that the particle number concentrations ...” is the erratum of “particle volume concentrations”?”

Yes. Thank you again for correction.

“Caption of Fig 3: “300-2000 nm” contradicts both the text description and the figure plot.”

Thank you again.

“Caption of Fig 6a: “100-5000 nm” means what? Is it described in the text?”

An unfortunate remnant from a previous image and another embarrassing error on our side. Thank you for pointing that out and it has been corrected to 300 nm -10 μ m.

Interactive comment on Atmos. Chem. Phys. Discuss., 12, 18889, 2012.

C9083