

## ***Interactive comment on “Statistical uncertainty of top of atmosphere cloud-free shortwave Aerosol Radiative Effect” by T. A. Jones and S. A. Christopher***

**T. A. Jones and S. A. Christopher**

Received and published: 17 May 2007

This paper studies the relevance of the arithmetic mean and standard deviation used to describe a dataset of aerosol optical thicknesses or radiative effects. The authors argue that these quantities are misleading in many cases. Since the published mean and standard deviation are often used when comparing different studies, and are the numbers reported in high profile works such as the IPCC assessment reports, questioning their validity is important. However, while I agree with the points the authors make, I think the paper overstates the problem. I recommend major revisions.

We thank the reviewer for the comments provided, and have revised the manuscript substantially based on his and other reviewer comments. Responses to specific com-

S1865

ments are listed below.

General comments. 1. What is meant by “statistical uncertainty”? This expression is not defined in the paper and I am unsure whether it is actually an uncertainty. Isn't it more a question of whether the arithmetic mean is able to adequately represent a non-normally distributed dataset? The authors put the skewness forward, yet this is an unfamiliar statistics for many scientists. Why not promote the use of the median, which is more robust than the mean for non-normal distributions? They also show probability density functions but fail to mention that it involves a binning of the data, which is often hazardous.

The term “statistical uncertainty” was removed and replaced with the term “statistical variability”. Statistical variability is defined as the robustness of globally averaged statistics relative to data distribution. This is now stated explicitly in the abstract. We chose to analyze skewness for the exact reason that is it not well known by many scientists. Part of the purpose of this work is to show AOT and SWRE statistics in ways previously not published. Median was not used, since it does not provide any distribution information. You are correct that it is more resilient to non-normal distributions and its use is listed as an alternative to what is done here.

Bin sizes were kept small (0.01 for AOT and 0.2 Wm<sup>-2</sup>) such that small scale change in probability were not lost. For this work PDFs primarily used as a tool to visualize data distributions. No actual AOT or SWRE statistics are calculated from binned data.

2. Uncertainty and variability. In section 4.1, the authors write that “the large sample deviations [...] exceed all known uncertainties present within the data”. Here, they are comparing standard deviations that do not quantify the same things. Standard deviations may measure: - The variability within a dataset. For aerosols, the variability is large since both spatial and temporal distributions are very inhomogeneous. - The uncertainty in a retrieved value. The inhomogeneous aerosol distributions are due to the short aerosol lifetime. The fact that the variability is larger than the uncertainty is

S1866

good news. An artificial increase in the variability due to the poor sampling of satellite retrievals would be bad news, but the test on random sampling made by the authors shows it is not the case.

Comparisons of “statistical variability” to observed uncertainties within the data have been toned down or removed from the manuscript. We agree that a direct comparison cannot be made as the 2 things are wholly independent of each other. Just as a note, some research reports “uncertainty” as the standard deviation of values from various works; however, we chose not to compare with these values (e.g. Anderson et al. 2005).

3. Bias due to clear-sky-only retrievals. I reckon this is the most important issue and I am disappointed by the authors’ analysis. There is a clear positive correlation between the distributions of the shortwave radiative effect (Figure 2), its standard deviation (variability, Figure 3) and the pixel count (Figure 4) and a negative correlation with the cloud fraction (Figure 5). Is it surprising? Is it really a problem when working on clear sky only? What are the consequences when scaling the clear-sky estimates to all skies? A more in-depth analysis is really needed here, but it is complicated since aerosols can have an effect on cloud cover, reducing it through the semi-direct effect and increasing it through the indirect effects.

Additional discussion concerning the implications of using only clear sky CERES-SSF pixels to derive globally averaged AOT and SWRE had been added. Essentially, the large size of a CERES-SSF pixel (400 km<sup>2</sup>) combined with the requirement that it be almost entirely cloud free, bias the resulting data to regions that are predominately cloud-free. Often cloudy areas are under-sampled, and these regions also often include significant aerosol concentrations. As a result, global aerosols are underestimated and SWRE must be adjusted to compensate. Indirect and semi-direct effect of aerosols on clouds are considered outside the scope of this research.

4. Inclusion or non-inclusion of missing aerosol components. Due to the relatively short

S1867

lifetime of aerosols, a given aerosol species cannot be present everywhere across the globe. The authors rightly state that one may therefore work only on those scenes where the aerosol species exists, or set its optical thickness and radiative effects to zero where it does not. Obviously, the resulting statistics are very different, since the datasets are different. Again, the authors make that sound like a problem before finally writing that “the matter really depends on what one is trying to show” (section 3.5). Indeed. There is no problem there and I fail to see why it takes so long for the authors to reach that conclusion. When working on a given aerosol species, only include those scenes where that species is present. When comparing different aerosol species, make sure the area taken into account in the study is the same for all species by adding zeros - and those zeros are physically meaningful.

The discussion concerning missing aerosol components has been shorted to some degree, but remains a key point of the results. One of the primary reasons this issue is discussed in such detail is that many other works report anthropogenic (or dust) statistics without mentioning which assumption they use. As a result, we decided to quantify its importance as a guide to future researchers.

5. Structure of the paper. My opinion is that sections 3 and 4 should be merged, in order to go straight to the results and avoid unnecessary repetitions of the methods.

Both sections were revised to remove duplication, but the authors feel that the manuscript as a whole reads better with Sections 3 and 4 remaining separate.

Specific comments

1. Introduction, third paragraph: Christopher and Zhang (2002) is cited as (2004).

2004 changed to 2002

2. Same paragraph: “the number of pixels can vary from a single pixel to over a thousand”. Shouldn’t a threshold on the pixel count be defined, in order to ensure that the grid cell average is representative of the aerosol variability? Could the authors offer

S1868

guidance on the choice of that threshold?

A threshold was not used since even a models threshold number (~50) would eliminate grid-cells from relatively large regions around the globe, mainly located around regions with high cloud cover. The resulting gridded data would then not be (relatively) spatially uniform and again be biased towards cloud-free regions. This issue is now discussed in greater detail

3. Introduction, fifth paragraph: “mathematical limitations result in cases where one or two aerosol types may not exist” and “this is perfectly reasonable”. There are no limitations, then, since in reality some aerosol types may not be present everywhere.

Under some circumstances, the Kaufman algorithm can produce negative values for either dust or anthropogenic AOT, which is clearly a non-physical result. Thus, the assumption is made that all negative AOT values are equal to 0. This point has been clarified in the manuscript.

4. Section 2.1, second paragraph: “clear-sky is defined as cloud and aerosol-free regions” followed by a criterion on cloud cover only. How is it ensured that those regions are also aerosol-free?

Since no such “aerosol free” regions actually exist in the data set. The aerosol-free SW flux background is calculated by regressing SW flux against AOT and taking the intercept, where AOT=0, and the Fclr value. This is now more explicitly defined in the text.

5. Section 3.1, last paragraph: “An aerosol pixel of a particular type with a certain reported AOT should have the same effect [...] no matter when or where the aerosols are located, assuming a uniform background such as the ocean.” This is not true. The direct effect has a strong dependence in the solar zenith angle (which depends on the location, day of the year and hour in the day). Typically, the same component aerosol loading will have a larger direct effect at mid-latitudes than in the Tropics. The ocean

S1869

reflectance is also very dependent on the solar zenith angle. See sections 3.4 and 3.5 of Boucher et al. (1998).

The authors regret the implication that radiative efficiency should always be the same. We recognize that it is a function of not only solar zenith angle, but even aerosol type. This section has been revised accordingly.

6. Section 3.3., first paragraph: The authors write that gridding the data forces a “spatially homogeneous dataset”. I can’t see how. The pixel count will still vary between grid boxes. Am I misunderstanding the authors’ definition of “spatially homogeneous”?

“Spatially homogenous” has been defined as “data as data that has an equal number of data points for equal spatial domains.” This has been added to the manuscript.

7. Section 3.4: Instead of taking only a small, fixed number of pixels to create the grid-box averages, couldn’t the authors take a fixed percentage of pixels? They would avoid the increase in variability for cells made out of many pixels.

Yes, that would be another way of doing this portion of the analysis. However, the random sampling testing for gridded data was, in part, designed to test the effect of the greater variability of data in grid-cell is many pixels. We want to know if this variability significantly affects the overall average. (It does not).

8. Section 4.1, first paragraph: “AOT are biased towards smaller values”. It is not really a bias in the statistical sense, since the real AOT distribution is very likely to include more small AOTs than very large one. This is because those events associated with very large AOTs (some mineral dust and biomass-burning events, typically) happen less often and span a smaller area than the background optical thickness of sea-salt, for example. It is also important to state that those smaller optical thicknesses are associated with larger retrieval uncertainties.

The reference to “AOT biased towards lower values” has been removed.

9. Same paragraph: It resembles a “Gamma distribution” with a small shape parameter.

S1870

Done

10. Section 4.1, second paragraph: "Figure 2 shows gridded data". I'm confused: Grid data will only be studied in section 4.2.

Discussion of Figures 4-7 has been moved to Section 4.2.

11. Same paragraph: "globally-averaged MODIS cloud fraction for all (clear and cloudy) data". The cloud fraction is obviously given for all scenes. They are classified as clear or cloudy using that cloud fraction.

The cloud fraction map was created using all available valid data, with no cloud fraction thresholds applied. The goal was to show the relationship between number of cloud-free pixels and cloud fraction.

12. Section 4.1, last paragraph: When looking at Figure 7, it is hard to tell whether a linear relationship exists between AOT and SWRE. Plotting the point density would show whether most of the data points are indeed along the fitting line.

Figure 7 is now Figure 3. From our perspective, it is important to show the relationship between AOT and SWRE for all available data. While some spread does exist in this relationship, its linearity (for  $AOT < 1.0$ ) seems clear on this figure.

13. Section 4.2, last paragraph: "Still, it is the latter that best agrees with previous research". That does not prove that the other value is wrong.

The statement has been removed

14. Section 4.3, last paragraph: It is very unclear. Do the authors mean that the Kaufman et al. (2005) method is not internally consistent? How exactly?

This has been reworded. The algorithms employed by Kaufman et al. (2005a, b) to calculate dust and anthropogenic AOT components respectively are not quite consistent with each other. Due to difference in assumptions and thresholds applied, dust and anthropogenic AOT calculated from 2005a or 2005b equations is going to be slightly

S1871

different.

15. References: Bellouin et al. (2005), Fan et al. (2005a), Li et al. (2004) are referenced but not cited. Kaufman et al. (2005) is cited throughout the paper as 2005, 2005a or 2005b. Which of the two Zhang et al. (2005) is 2005a?

References fixed

16. Figure 1, 6, 8 and 9: What do the numbers and the vertical lines represent? They do not seem to match.

Numbers represent mean values. The discrepancy was a result of generating a new set of figures during a previous revision which used an improved data set, which has a better Fclr definition and improved spatial coverage. However, we failed to change the text overlies on the figures. This has been corrected.

17. Figure 1: The caption should state that symbols are used for the PDF and the solid line represents the ideal normal distribution.

Fixed

18. Figure 10: In the text we are told the figure shows the distribution of the anthropogenic optical thickness, and there is indeed one plot. However, the caption reads "Dust and anthropogenic component AOT and SWRE"?

Reference to dust AOT was remove

Interactive comment on Atmos. Chem. Phys. Discuss., 7, 3557, 2007.

S1872