Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2019-372-RC2, 2019

Atmospheric Chemistry and Physics
Discussions

Interactive comment on "How should we aggregate data? Methods accounting for the numerical distributions, with an assessment of aerosol optical depth" *by* Andrew M. Sayer and Kirk D. Knobelspiesse

Anonymous Referee #1

Received and published: 17 September 2019

The authors discuss the nature of probability distributions of AOD, the aerosol optical depth. The authors start by analysing for a diverse collection of datasets (AERONET, MODIS, MISR, GEOS5 Nature Run), whether spatially or temporally grouped data is better described by a normal or log-normal distribution. They show that at short time-scales (day), the normal distribution is appropriate but that at longer time-scales, log-normal distributions are more realistic. They then continue to show that means derived from such datasets, using either arithmic or geomtric means, can be quite different. In particular, they show that trend estimate can differ significantly in magnitude (though

not in sign). As the data mostly exhibit log-normal distributions at longer time-scales, the authors conclude that the common use of arithmic means in trend analysis is inappropriate. The paper concludes with suggestions for improvements in aggregation methods. This is an interesting paper about a fundamental issue in Earth sciences and entirely appropriate to AMT.

The statistical analysis in this paper seems sound. However, I feel the authors may be overstating the importance of this issue. Often we don't look at AOD but at differences in AOD (satellite evaluation, model evaluation, changes between present day and pre-industrial, model sensitivity studies, etc). The resulting distributions are in my experience usually more normally than log-normally distributed.

Also, any log-normal distribution can be accurately defined by arithmic mean and standard deviation. Actually, there is a one-to-one transformation from the arithmic statistics to the geometric statistics, see e.g. https://en.wikipedia.org/wiki/Log-normal_distribution. My interpretation is that it is not important whether one uses arithmic or geometric statistics, as long as one is aware that their use does not imply (!) either a normal or log-normal distribution.

Another issue is physical conservation of the property under study. AOD is not a good example so let's consider column burdens of aerosol or trace gases. These may be expected to have log-normal distributions in time and space as well. Describing them with geometric means would cause loss of mass conservation! Consider a dataset at 10 km that is aggregated to 100 km: the arithmic mean preserves total mass in the 100 km grid-box while the geometric mean does not.

It seems one has to consider what is causing the log-normality: if it is due to log-normal retrieval errors, geometric means seem justifiable as they ameliorate the effect of outliers. If it is due to the nature of the property, conservation-laws may be more important and arithmic means are to be used. I am sure much more can be said about this.

That said, if arithmic means are used to describe log-normal distributions and then carried forward through non-linear analyses under the assumption of normality, significant problems may arise. The authors allude to this on p. 23, l 19 when they talk about parametrisations.

It would be great if the authors take the above into consideration when preparing their final manuscript. In all this is a worthwhile discussion.

Minor comments:

p 3, l 33: it would be good to state the relation between arithmic mean (and stddev) with geometric mean (and stddev) for a log-normal distribution. Such relation exists, see https://en.wikipedia.org/wiki/Log-normal_distribution

p 5, l 17: "will overstate the typical level of AOD observed and its variability" . While I understand the authors' intention, it seems to me this sentence suffers from the absence of what is "typical". It would appear that "typical" here refers to the geometric mean as a parameter that defines a log-normal distribution. However, there is a simple 1-on-1 mathematical relation between arithmic and geometric mean of a log-normal distribution. I.e. the arithmic mean defines a log-normal distribution as well as the geometric mean. Hence both arithmic and geometric mean can be used to define what is "typical".

p 6, l 25: "using Normal-appropriate statistics has systematic quantitative implications for the interpretation of the data." Only if the arithmic mean and stddev are interpreted as defining a normal distribution. It is perfectly possible to calculate both without reference to a normal distribution. Actually, they can define a log-normal as well.

p 6, l3: "these factors may include": I believe turbulence is an important factor in the creation of log-normal distributions?

p 8, l 13: "This quadratic formulation is more robust to calibration problems in individual channels" more robust than what? Maybe consider dropping "more"?

p 11, l 7: "tail-waited" tail-weighted ?

p 12, l 3: Shouldn't arithmic and geometric stddev be compared as well?

p 13, l 3: "results for temporal (from AERONET and G5NR) and spatial (from MISR, MODIS, and G5NR) frequency distributions of" This confused me as both Figures show spatial distributions of the WS test. The test, in all cases, was presumably done on time-series of data. The captions to the figures seem to say something different: either data was a temporal aggregate (which suggests G5NR results are at its native resolution) or spatially aggregate (which suggests each 30 min of G5NR data was used). Please clarify this?

p 14, l 5: "calculating an arithmetic mean when the underlying distribution is Lognormal (or vice-versa) introduces an error smaller than 0.01." I disagree with the use of the word 'error'. No error is incurred at all. It is always possible to calculate arithmic means. Any error is due to limited sample size. See also my previous comments.

p 17, l 7: "Note also that the near-universal choice of aggregating daily on a UTC calendar day basis, rather in terms of local solar time, can further complicate matters for locations far from the meridian." For another example, see Schutgens, Partridge & Stier ACP 2016, Fig. 13 & 14.

p 23 , l 11: "Even a small change in reported AOD, if systematic, can have important implications for calculations of climate forcing." But changes (differences) in AOD are far more likely to have a normal distribution.

p 23, l 20: "the same argument may apply if forcing parametrisations are developed from model simulations aggregated in certain ways" This is a fair point. A lot of studies point out the distorting impact of non-linear physics/chemistry when using just the mean to represent a distribution. One example from remote sensing is the plane-parallel bias noted in cloud retrievals of LWP. Note however that such biases exist not because of an arithmetic mean but the representation of any disribution by a single number.

p 23, l 6: "This implies that no simple scaling correction can be applied to existing data sets to transform between arithmetic and geometric estimates" . Assuming a normal or log-normal distribution, exact transformations exist between mean and stddev of arithmic and geometric statistics.

Sect 4.2 The analysis in this section seems sound and I have no problems with it. That arithmic means yield different trends than geometric means is no surprise, after all these are different means. However, there is the suggestion that geometric means are better simply because the underlying distribution is log-normal. Rather, geometric statistics make it easier to interpret changes in a log-normal distribution but they do not provide more information (or put differently: the arithmetic statistics are not "wrong"). Note also that trend analysis of changing log-normal distribution really requires geometric stddev to be analysed as well but this is seldom done.

p 26, l 6: "but quantitatively have a tendency to overestimate their magnitude." It may be good to repeat here that at these three sites the log-normal distribution is the more appropriate distribution to use (previous analysis, Sect 3). At least that seems to be the suggestion here?

p 27, l 20: "estimated trends in geometric mean AOD are smaller in magnitude" Trends in satellite data are over often calculated over regions, not like the point sources the authors have used in their example. I wonder how this will affect these conclusions? At some point the central-limit-theorem should kick in and turn any log-normal distribution into a normal one?

p 27, point 2: this point seems to imply it is ok to average AOD in time and compare satellites with satellites or models, as long as we use the proper mean. The authors know that different sampling of data sources often has a far bigger impact. Maybe it is good to state that here.

p 27, l 9: "root mean square error" This is a difference between two AOD and is likely to be normally distributed.

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-372, 2019.