Nonlin. Processes Geophys. Discuss., 2, C538–C541, 2015
www.nonlin-processes-geophys-discuss.net/2/C538/2015/

Nonlinear Processes in Geophysics
Discussions
Open Access

Interactive comment on "Identifying non-normal and lognormal characteristics of temperature, mixing ratio, surface pressure, and wind for data assimilation systems" *by* A. J. Kliewer et al.

Anonymous Referee #2

Received and published: 26 October 2015

The authors study the probability distribution of some parameters of the GFS model forecast (mixing ratio, temperature, wind field). They try to combine different tests to detect whether a probability distribution is non-normal and lognormal in particular. Finally they discuss where and when the GFS forecast follows a non-normal distribution. The main objective is to provide this information so that a data assimilation system can adapt and uses assimilation schemes specifically for lognormal distributions. Instead of using a single test, they combine the results of three statistical tests (Shapiro-Wilk, Jarque-Bera, a χ^2-test) together to decide whether a variable at a given location is lognormal distributed.

Unfortunately, I cannot recommend this manuscript for publication in Nonlinear Pro-

cesses Geophysics. I have the following major comments about this manuscript, essentially about the level of innovation, the robustness of the results and about the way the conclusions are supported by the statistical tests:

1. First of all, it is unclear to me if by combining the results of three standard statistical tests together in a single test, the authors reach a sufficient level of innovation to justify a publication in Nonlinear Processes Geophysics. The scientific advancement of this manuscript is not on a par with what I usually see in manuscripts published in this journal (and other journals with a similar scope for that matter).

2. It is also unclear how much better the new approach works. The authors do not attempt to determine the probability of the type I or type II error of the combined test. Further work could be done either on a theoretical level or with synthetic data with known distribution. This leaves the reader with the question, if it is really better to use 3 tests instead of one and by how much is it better?

3. The question where a variable follows a Gaussian distribution and how strong the non-Gaussian behavior is are actually two different questions. The statistical tests (in particular Shapiro-Wilk, Jarque-Bera) are designed to answer the first one. But for variables which are always larger than a given threshold value, we already know that these variables do not follow a Gaussian distribution. If the sampling size could be made arbitrarily large, then one would end up rejecting always the null hypothesis (rejecting the thesis that a variable follows a Gaussian distribution), for the mixing ratio in particular (since it is always positive). It should have been clear from the start that the mixing ratio does not follow a Gaussian distribution. The authors present the results of the test for a data set spanning one year. The presented maps are not robust if the data from increasingly long time spans were considered as the normality test would tend to reject more often

the null hypothesis (simply because the probability to have negative mixing ratio is exactly zero and the pdf of a Gaussian distribution is always strictly positive).

Another way to look at this is to assume that a variable follows a distribution which is close to a Gaussian distribution but has a very small skewness. Since the skewness is not exactly zero, for a sufficiently larger sampling size n (equation 8), the Jarque-Bera test will end up to rejecting the null hypothesis.

The question should rather be whether the deviation from Gaussianity is too large to apply a assimilation scheme like the Kalman Filter of 4D-Var.

4. Data assimilation is cited at several places as the application for this study where an assimilation scheme could change depending on the distribution of a variable at a given location and time. However, it is not clear to me if it is really feasible to change the assimilation schemes from normal to lognormal distributed variables from one grid point to the next. This would lead to unrealistic discontinuities at the transition zones. A clearer view on how the results of the detection method will be used later during the assimilation is necessary in my opinion in order to ensure that the results can be used.

5. When the null hypothesis cannot be rejected, one cannot conclude that the null hypothesis is true (see also below). This assumption is made at several places within the manuscript, especially when deciding if a variable is lognormally distributed or not.

Additional comments

- page 1399, line 26: the minimizing solution is usually noted $\mathbf{x}^a$ while $\mathbf{x}$ is the free parameter of the cost function.
- page 1367, line 22; Fletchera -> Fletcher (I suppose)

- page 1369, line 18: The validity of statistical tests relies critically on the assumption whether the samples are independent. I think that a paper focused on statistical tests should verify this assumption (even if it is not difficult to find other papers in the literature which do not verify this assumption).
- page 1371, line 14: "This design ensures that the data truly is lognormally distributed without a false positive." I guess (hope) that this is a typo here. A test without any false positive is of course not possible (unless α would be set to zero).
- page 1373, line 3: "If the distribution is normal, then asymptotically the Jarque–Bera (JB) test statistic has a chi2 distribution with two degrees of freedom...". Can you be more specific to what the term "asymptotically" refers to? Which parameter needs to tend to infinity? The sampling size?
- page 1375, line 10: "Therefore the composite test concludes that almost 29 % are lognormally distributed." The authors make often this shortcut in their reasoning. However, if the null hypothesis (variable follows a lognormal distribution) is not rejected, it does **not** mean that one can conclude that the variable follows a lognormal distribution (because the sampling size could also be too small to reveal deviations from the lognormal distribution).
- page 1381, line 1: "It has been shown in Fletcher and Jones (2014) that there is a negative impact on the performance of a normal distribution only incremental 4-D-VAR **when lognormal forecasts are assimilated**". I guess that the authors wanted to say "... when lognormally distributed observations are assimilated". Otherwise I do not understand this sentence.

Interactive comment on Nonlin. Processes Geophys. Discuss., 2, 1363, 2015.