Atmos. Chem. Phys. Discuss., 11, C3095–C3098, 2011
www.atmos-chem-phys-discuss.net/11/C3095/2011/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Comparing the effectiveness of recent algorithms to fill and smooth incomplete and noisy time series" *by* J. P. Musial et al.

Anonymous Referee #1

Received and published: 10 May 2011

This paper compares three method (one with two variant) to fill missing entries. The authors used synthetic time series as well as observed time series. Missing data are generated using three different methods covering a large range of what occur in real world (random, prolonged gap and seasonal gap). It is an interesting analysis, but I think that the authors could improve the post-processing of their analyses and they should include another score of "goodness of fit" different from MAE. I add that being not a native english-speaker, I can not really judge the quality of language.

Major issues In conclusion, the authors said hat "It must be realized that a single indicator of "goodness of fit" such as MAE cannot fully represent the range and diversity

of goals in any particular analysis". I fully agree with that, but why the authors did use a single statistical score to evaluate the goodness of fit ? They should at least use a different statistical score and it will be perhaps useful to choose a score that is very dissimilar to MAE (as a probability or a categorical score). It is impossible to get an "universal" view of the fit between observations and predictions but it is desirable to look at this fit at least through two different points of view.

In the evaluation of any prediction, we are usually interested with the residuals from a well-known regular cycle. For example, we could be more interested to know for example the difference between two consecutive winters rather than the difference between winter and next summer, because this signal is forced by the annual cycle of radiation and it is rather trival. For sunspot time series, it is perhaps desirable to simulate accurately the relative level of maximum and minimum of each cycle and perhaps the length of each cycle. Except to Dow Jones Index, most of the time series used in your paper contain a moderate to strong periodic component. Is it possible to remove this dominant periodic component and test the "goodness of fit" between observations and simulations for residuals ?

In the evaluation of MAE, nothing is provided to see when MAE is over what is expected by chance. Can you add a statistical test ?

Minor points

p.3 : About your third paragraph, I think that your use of "broad, slow variations" vs "random" , "signal" and "noise" is rather ambiguous since, at least in climatology, the "signal" could be strongly aperiodic.

p.8 : personnally, I don't know the MSL routine CSSMOTH : is a GCV method ? Did you use this method for each time series analyzed ? Does it increase the computational cost of this method ? Please clarify the end of this paragraph.

p.11 : a flow-chart will be useful to present each step used in SSA.

p.14 : I think, as said above, that your hypothesis that high frequency variations are random is too restrictive. I add that "high-frequency" and "low-frequency" is not precise enough (section 3.3).

p.14 : please refer to your equations in the first paragraph of section 3.3.

p.18 : what is the increment of % noise and % of missing entries, 1% ?, 5% ? 10% ?

p.19 : adapt the y-axis of fig. 1.

p.19 : "is more depend" on the last line sounds odd. Please check.

p.20 : about the winter-gap scenario, I think that your comparison between Lomb-Scargle (LS) and Kondrashov and Ghil (KS) methods on one hand and spline on the other hand is unfair since what happens for a summer-gap scenario ? I think that spline doesn't produce spurious peak as soon as you have enough data to reconstruct the autumn decrease and the spring increase, but this methods won't be superior to other ones if you don't get these tendencies. I understand that spurious peak appears because there is an intermittent "bi-annual" cycle which is simulated for each year by LS and SSA methods. But I think that the superiority of spline is mostly a matter of chance in your example. Your use of term "spurious" is perhaps ambiguous in that context and "intermittent" could be a better choice. This comment applies also to what is written in the second paragraph of section 5.3.

p.29 : I think that LS and KS methods have problem with intermittent signals.

p.30 : Can you give us an objective measure of the "considerable computational cost" for example as cpu for each method and how it increases with length of observed or simulated time series, and possibly the amount of missing entries to be filled. This would help to scale the advantage or disadvantage of each method in an objective way. The same comment applies to what is said at the end of the 2nd paragraph p.7 ("it will be seen that these methods ... ") and also page 11 ("on the other hand, the high computational requirements of the SSA gap-filling algorithm").

Interactive comment on Atmos. Chem. Phys. Discuss., 11, 14259, 2011.