Atmos. Chem. Phys. Discuss., 8, S1726–S1728, 2008
www.atmos-chem-phys-discuss.net/8/S1726/2008/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Gap filling and noise reduction of unevenly sampled data by means of the Lomb-Scargle periodogram" *by* K. Hocke and N. Kämpfer

Anonymous Referee #3

Received and published: 18 April 2008

This manuscript presents a well-developed mathematical tour de force using periodogram techniques to carry-out interpolation due to gaps in the original data series. As a statistical tool to recover missing geophysical data, the process thus developed has profound difficulties in having some built-in assumptions that would likely invalidate the process, and which the authors did not address.

In order to derive a power spectrum from a time series, this time series must be a realization of a stationary process of at least order 2. That is, it has the same mean and variance at all time points, and the covariance between the values at any two time points depends on the interval between these time points and not the location of the

points along the time axis (Priestley, 1981). By definition, this is correct in the synthetic series the authors generated as a test example, but this point was not addressed -or tested- for the real world data they use.

The number of independent frequencies that can be extracted with a periodogram is limited by the number of degrees of freedom available in the data series -in the limit by the Nyquist theorem. It was assumed by the authors that this was not a limitation, but was not tested. How this affects the validity of the authors deductions should be tested and discussed by them.

In addition, it is known that the Lomb-Scargle periodogram statistic is over-optimistic by at least a factor of 2 -in the limit for a very large sample- (Hernandez, 1999). Thus, the Lomb-Scargle significance test is likely to produce more 'meaningful' frequencies than are actually present.

There is a world of difference between irregularly sampled data, and data gaps. With the limiting case being the presence of periodic, or coherent, data gaps. In general, irregularly sampled data incur some noise in the resultant periodogram, while data gaps tend to change the spectral content found in the periodogram. An example of this behavior is illustrated in Figures A5 and A6 of Hernandez (1999), as the appearance of significant power at frequencies other than those originally present in the data, previous to the existence of the data gaps. The authors simply ignored this possibility.

Finally, the authors have not discussed the philosophical implications in interpolating geophysical data. Namely, that the interpolated data has no more information than the original data with gaps had. Thus, nothing new can be learned from the newly generated data. But one could say that this newly-generated data is more pleasing aesthetically.

In conclusion, the present manuscript is not complete. The authors should address the statistical limitations presently ignored in the text, as well as the statistical usefulness of interpolation.

ACPD

8, S1726–S1728, 2008

References.

Hernandez, G., Time series, periodograms and significance. J. Geophys. Res., 104, 10355 (1999)

Priestley, M. B., 'Spectral Analysis and Time Series'. Academic Press, London (1981)

Interactive comment on Atmos. Chem. Phys. Discuss., 8, 4603, 2008.

ACPD

8, S1726–S1728, 2008

Interactive Comment