

## ***Interactive comment on “Characterizing ecosystem-atmosphere interactions from short to interannual time scales” by M. D. Mahecha et al.***

### **Anonymous Referee #1**

Received and published: 9 June 2007

The present manuscript applies a novel technique from systems dynamics/ control theory, namely singular systems analysis (SSA), to a long-term eddy covariance time series from the Tharandt spruce forest in eastern Germany. The authors use SSA to determine that a low frequency signal that corresponds to interannual variability pervades both flux and climatic drivers, and advocate the future use of SSA for both analysis and processing (i.e. gapfilling) of eddy covariance flux data.

I believe that the analysis is strong and warrants publication given that the technique is novel, the analysis was carried out competently, and the results are highly interesting. However, a number of corrections - both relatively major and minor - must be carried out to improve the manuscript to increase its value to the flux community and readers of BGD.

In my opinion there are two major shortcomings:

1) It is unclear how the SSA is advantageous to or complimentary to other techniques such as wavelet analyses that are likewise able to deconvolve the dominant frequencies of complex, gappy, nonstationary time series. One of the major advantages is the lag-analysis (which wavelet techniques can also explore but to date have not). A major disadvantage is the computational expense of SSA. I agree that SSA would provide an improvement over Fourier decomposition for even more reasons than stated (but not quantified) by the authors, but I believe the discussion of the wavelet techniques was overly critical and not fair; nothing about the present paper suggested to me that SSA represents a substantial improvement over wavelet analysis outside the potential to fill gaps.

2) The discussion and analysis regarding the gaps is not straightforward. The authors discuss both how SSA can be used for gapfilling and it is not clear if the analysis was carried out on these gapfilled time series, maybe I missed something. A major concern is that Reco and T have different significant periods (none in the case of T); yet the Reco time series was largely determined by T. What are the significant modes of the Reco model without measured nighttime values?

In addition, most passages were poorly worded and should be improved for both readability and, in many cases, technical accuracy. As a small example, the second sentence in the abstract might be re-written: 'The respective time series currently comprise an observation period of up to one decade.' This is a minor case to make the sentence more accurate and readable, and may not present a major problem for comprehending the overall meaning. But the reader should not have to dissect a sentence to determine its meaning. As a side note, all sentences in the first paragraph have some type of English mistake. The second paragraph is better-written as a whole but not without mistakes.

Comments: How were the dimensions of P (the 'embedding dimension') chosen in the

[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)[Discussion Paper](#)

present analysis and are the results sensitive to its size (sizes)?

The variable 'A' is defined well after its first use in eq. 4 and the description of that equation is dominated by tangential asides. Please explain eq. 4 more clearly for those not familiar with principal components analysis.

I don't see the definitions of  $L_t$  and  $U_t$  (corrections for series boundaries). Any corrections must be thoroughly explained because, as a rule, corrections influence results.

I think that the technical description was well-written (outside of some language concerns), but was rather dense. A figure that summarizes the steps in the analysis and/or a brief appendix with pseudocode may make the algorithm easier for others to understand and apply.

I still do not completely understand the flux community's fixation with gapfilling. This will always be more art than science and, rather than making up more data (i.e. gapfilling), we should pay more attention to analyzing the wealth of data that we have. From this standpoint, wavelet analysis with the Haar basis function (e.g. Katul et al., 2001) makes perfect sense. It can decompose a gappy time series into the frequency domain and analyze the spectral dynamics of what is actually measured by the flux system. Consequently, a more thorough discussion of the benefits and drawbacks of SSA versus wavelets is necessary. Some good references for a comparison would be Braswell et al. (2005) GCB 11, 335-355; Stoy et al. (2005), Tree Phys, 25, 887-902; and especially with respect to the present manuscript Richardson et al. (2007) GCB, 13, 788-803.

If SSA carries a substantial computational burden (p. 1413, line 26; 1414,1), via the lag-covariance matrix, I question its general applicability as a gapfilling tool, as this implies it could be performed across the FLUXNET database. I would prefer that the gapfilling discussion be removed entirely as it is tangential to the present manuscript. The use of SSA to fill gaps is more appropriate for a different manuscript, e.g. one that might compare the preferred methods from statistics for gapfilling (imputation) with

[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)[Discussion Paper](#)

SSA and other approaches.

Why is the embedding dimension N/P set to 2.5? Discuss this in the methods section, not the results section.

Page 1415, line 2: I do not doubt that the flux time series contain modes of variability beyond the frequency of what the measurements and SSA technique can determine, but do these significant modes represent a real signal or edge effects in the analytic procedure?

I, too, am confused as to why T does not vary at low frequencies. Could the authors test this against a T time series that is known to have interannual oscillations to demonstrate the fidelity of the SSA approach?

1415, 27: Please describe the multiscaling behavior in P in more detail; can SSA quantify this behavior?

I question the exploration of Reco and GPP time series given that these are (usually) not directly measured products (especially the latter) and are instead the output of a model that is driven by a climatic variable which, in the case of GPP, is again filtered through the NEE time series. The model for Reco of course represents an assumption. Can the authors use SSA to determine the variability in measured (not modelled) Reco? How does this compare with the measured Reco?

1416, 2: How was it determined that high frequency components are essential to total variance for GPP, H, etc. and what is meant by 'essential'?

1417, 6; I like the NEE-T analysis. Would a residual analysis between these terms (and others) display complimentary or new information?

1418, 17: Pronounced hysteresis in the NEE-u relationship may not be causative, and may simply reflect the fact that both show some seasonal patter. I doubt this has anything to do with ecosystem 'memory'.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

Minor comments: Abstract, line 11. Listing abbreviations without defining them is understood by people in the field, but is jargon to others. Page 1407, line 6: This statement should be referenced. 1408, 20: I don't know what 'methodological chances' means. 1409, 21: 'eigenvectors and eigenvalues are themselves real valued' (or drop the word themselves entirely). 1411, 26: spelling of 'computationally'. Also, this is an extremely wordy sentence. 1412, 10: write 'shown to produce'. 1414, 15: 'spruced' is not a common adjective. 1414, 16: spelling of 'deciduous'. 1414, 23: 'gamut' is colloquial. 1414: 25: 'quantitatively summarized' is redundant. 1415, 13: 'frequency' instead of 'frequencies'. 1415, 27: Please expand on the multiscaling behaviour determined by Peters et al., (2002). 1415, 28: write 'Frequency' and 'Noticeably,' 1416, 7; write 'discernable' instead of 'decidable'. 1417, 14: write 'explaining' 1420, 3: do the authors mean T/P? 1420, 7: The eight year observation period is rather long for current eddy covariance time series. 1421, 10: spelling of 'rejected' 1421, 15: write 'optimally when only one of the leading modes'. 1425, 4: Spelling of 'Barford' 1434, write 'artificial' gap.

---

Interactive comment on Biogeosciences Discuss., 4, 1405, 2007.

**BGD**

4, S633–S637, 2007

---

Interactive  
Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper