Atmos. Chem. Phys. Discuss., 9, C5959–C5962, 2009
www.atmos-chem-phys-discuss.net/9/C5959/2009/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Exploiting the weekly cycle as observed over Europe to analyse aerosol indirect effects in two climate models" *by* J. Quaas et al.

J. Quaas et al.

johannes.quaas@zmaw.de

Received and published: 16 October 2009

[...] The statistical significance is assessed incorrectly and the widely varying time periods and spatial distributions of the different data sources make comparisons impossible.
It was indeed an unreasonable choice to have different time period lengths for the surface observations than for the other data. We have corrected for this now (please see the discussion on how this changes the results in additional reply). As for the different spatial distributions, we think that the distributions of model data and satellite data are readily comparable. Also for the EMEP data, the observations are averaged

over a large Europe-wide network, so that there should be no bias compared to the models in terms of weekly cycle. For the DWD data, it might be true that focusing on one central-European country introduces a bias towards a stronger weekly cycle. However, our conclusion is not that the models show as much variability as the data – thus, if Europe-wide data were available, the conclusions would very likely be the same.

Statistical significance is at the heart of any paper on weekly cycles. Any non-trivial time series has different, and hence always has some weekly cycle. The key is whether or not the differences are statistically significant.
We thank the reviewer for their advice on this, and we put a large effort into improving on this point. For each of the time-series, we computed a power spectrum, and estimated the significance of the contribution of the seven-day period. We tabulated the significance level for each of the quantities in Table 1. An explanation of this method to compute the significance has been added to the revised manuscript.

This manuscript uses a t-test to assess statistical significance, which has been shown to be invalid for this purpose (Barmet et al., 2009).
We agree that the t-test has several problems, in particular for the analysis of precipitation as discussed by Barmet et al. In the revised version, we refrain from testing the significance in this way at all, but we rather use a power spectrum analysis.

There are three problems with a t-test. First, it is intended for calculating the significance of a single threshold, not the 7 different possible weekly maxima.
While obsolete for the revised version, we would like to indicate for the records that we only compared two single distributions: distribution from the extreme day against each distribution of the other six days.

Second, the t-test assumes independent data points. Atmospheric data are strongly

autocorrelated in time and space so the number of independent data points is less than the total number of points.
We agree.

Finally, the t-test assumes normal distributions, and most atmospheric data are not normally distributed.
We also agree, in particular for the precipitation (the distributions of the other quantities might be better approximated by a normal distribution, though).

A full description of a proper way to assess statistical significance has not been published to my knowledge, but one method is a block bootstrap approach whereby weekly cycles are calculated for data randomly moved around in blocks and compared to the weekly cycle from the real time series. Calculating 6 and 8-day cycles is also helpful.
Thank you very much for this advice. Concerning for a pertinent method, we consulted colleagues who are experts in statistical analysis, and together developed our new approach, in which we now use a power spectrum analysis.
We also follow the reviewer's suggestion to calculate 6- and 8-day cycles. These are shown as supplementary material in the revised version, and a paragraph discussing this has been added to the manuscript.

The second major problem is the wide variation in the years covered by the data. [...]
As acknowledged above, we corrected for this.

In addition, the meteorological data are for 41 German stations whereas the satellite data are averaged over continental Europe.
We agree that it would be much better to have Europe-wide meteorological data.

However, as explained above, we think that this would not modify our conclusions.

The large collection of data sets in this work could be very useful when properly analyzed.
Thank you very much for this encouraging statement.

Interactive comment on Atmos. Chem. Phys. Discuss., 9, 11269, 2009.