

## ***Interactive comment on “Analysis of exceedances in the daily PM<sub>10</sub> mass concentration (50 µg m<sup>-3</sup>) at a roadside station in Leipzig, Germany” by C. Engler et al.***

**C. Engler et al.**

engler@tropos.de

Received and published: 7 November 2011

*Your comment: To me it is unclear on what basis the authors decided that clustering into 9 clusters is appropriate, the description on page 15839 (line 23 and following lines) and the first paragraph on page 15840 do not help very much. For example I do not understand the following two sentences "The deviation of the average PM<sub>10</sub> concentrations (and the other aerosol and meteorological data) between the clusters was calculated for each test run and used for the choice of the weighing parameters. Selecting a small number of clusters will generate larger, more representative subsets of the data.". What does "more representative" mean, how was the clustering*

C11547

*procedure exactly done? Please describe this more clearly. Otherwise it cannot be seen that the applied clustering is an objective method as stated. What is the effect of the choice of weights  $a_1$ - $a_4$  on the results?*

### **Cluster**

The number of synoptic weather situations or air mass types in Central Europe varies according to each investigator, ranging between a few and around 40. Our experience is that in back trajectory cluster analysis the scientist needs to make a balanced compromise between a simple display of the data containing a limited number of clusters, and the usefulness of the resulting clusters in terms of synoptic-scale weather situations, which clearly implies a higher cluster number. In our case, 9 turned out to be a viable number.

### **Weighing of parameters**

To determine the effects of the weight parameters  $a_1$  to  $a_4$ , we performed a sensitivity analysis, trying different combinations of  $a_1$  to  $a_4$  within wide pre-defined ranges. We also varied the total cluster number (3 to 14). In total 4200 different simulations were done, with the parameters varying between  $a_3=[1e-6$  to  $5e-4]$  m<sup>-1</sup>,  $a_4=[0.5$  to  $15]$  K<sup>-1</sup> and  $a_{1,2}=1$  °<sup>-1</sup>.

### **Inter-cluster variability**

The results of this sensitivity analysis are displayed in Fig. 1. The inter-cluster-variability is the standard deviation of the mean (e.g., meteorological or particulate) parameters (in this case air pressure, global radiation and SO<sub>2</sub> concentration) between the clusters: One cluster consists of a fraction of all considered days. The average values of all parameters of these days were calculated. This results in one average value for each considered parameters for each cluster. For example choosing 9 clusters means 9 average values for the temperature, PM<sub>10</sub> and so on, accordingly. The inter-cluster-variability is then the standard deviation of these 9 average values. It is a measure on the diversity of the clusters with respect to each (e.g. meteorological) parameter. The figure shows that the inter-cluster-variability increases with increasing

C11548

cluster number, but this increase is getting smaller to higher cluster numbers. So the increase of information is very small choosing more than 9 clusters. This is why we decided to work with 9 clusters. After that, the weighing parameters were chosen as the ones, used for the test calculation producing the highest inter-cluster-variability in Fig. 1 for these 9 clusters (inter-cluster-variability=2.1). The parameters were  $a_{1,2}=1 \text{ }^{\circ}\text{C}^{-1}$ ,  $a_3=1\text{e-}5 \text{ m}^{-1}$  and  $a_4=2 \text{ K}^{-1}$  (as described in the manuscript). The text in the manuscript was changed into: "A higher pre-defined cluster number will split the data set in smaller sub-sets, which correspond more specifically to certain synoptic-scale weather situations. A lower cluster number will, in contrast, produce less but larger sub-sets of data. Those larger sub-sets may be representative for larger parts of the entire observation period, but are less specific to individual synoptic-scale weather situations."

*Your comment: The manuscript is in some parts not carefully written, there are many typos and linguistic errors. An example is page 15849 lines 6, 9, 25 and 26. The text should be carefully revised accordingly.*

Our response: The text was revised with respect to typos and linguistic errors.

*Your comment: The methodical paragraph on the approach by Lenschow et al. (2001) should not appear in section 4.2.2, but in the statistical methods section 3.1. Are the given definitions for RT, UI and TI correct? If yes, then RT, UI and TI do not sum up to one. Please explain. Why not  $UI=(PM_{10,urban}-PM_{10,rural})/PM_{10,roadside}$  and  $TI=(PM_{10,roadside}-PM_{10,urban})/PM_{10,roadside}$ ? Then  $RT+UI+TI=1$  holds.*

Our response: You are right, the definition of UI and TI was wrong in the manuscript. The correct definition is as you mentioned. This was changed accordingly and the definitions are now given in section 3.1.

*Your comment: Section 4.2.2 appears lengthy, is hard to read and sometimes confusing. Many numbers are given and there are sentences that are hard to understand and irritating. I would very much appreciate if the authors could carefully revise this section focusing on readability.*

C11549

Our response: Thank you, we will take care in revising this section accordingly.

*Your comment: It is very hard to see any differences in the stacked bar charts (Fig. 6, 7, 9), as a consequence it is hard to follow the discussion. For a more clear illustration of the differences on exceedance days versus non-exceedance days, the authors might want to consider alternative representations. The concept of relative differences as used e.g. by Amato et al ACP 2011 could be a possibility.*

Our response: Figures 6, 7 and 9 were revised according to your suggestions. For Fig. 6 and 7, we tried out to show the absolute differences from the average over the entire period. We decided to add the new diagram to Fig. 6 in order to improve the content of information. For Fig. 7, the new diagram (Fig. 2) did not show the chemical composition any more, which is why we decided to keep it as it was in the manuscript. In Fig. 9,  $PM_{10}$  and  $PM_{coarse}$  were isolated for more clarity.

*Your comment: The summary and conclusions section is mostly a repetition of the discussion section. Please avoid this, I suggest to drastically shorten this section to a short conclusions section.*

Our response: Thank you for your suggestion, we agree that there is considerable repetition. We will consequently condense these sections accordingly.

*Minor comments: Abstract, line 17: What does "the latter factor was instrumental in generating ..." mean, please clarify.*

We noticed (in former studies as well) the stability being crucial for generating exceptionally high concentrations of particulate matter. As described in the paper, the trapping of pollutants under an inversion layer and missing horizontal or vertical dilution can lead to accumulation of pollutants close to the source.

*Abstract, line 24, there is a word missing: "Should be something like "...indicates that both contributions are equally..."*

The word "contributions" was added according to your suggestion.

C11550

Page 15833, line 7, abbreviations (TEOM, OPC) should be explained or avoided.

Explanations for the abbreviations were added.

Page 15835, line 1, typo "...with the analysis of the meteorological...", also line 22 "previously" can be deleted.

The words "the" and "previously" were deleted.

Page 15836, lines 17-19. Please give the spatial distance between the suburban and the roadside site.

A sentence was added: "The distance to the roadside site is about 7 km."

Page 15838: line 11. "The average chemical composition was calculated...", should be something like analyzed instead of "calculated".

The sentence was changed into: "The average chemical composition was estimated..."

Page 15838: line 13-14. Should be "It is used for testing whether two populations result from the same random distribution or not".

The sentence was change according to your suggestion.

Page 15838: line 25. What does "custom-made k-means cluster algorithm" mean. Please provide the details of the software that has been used.

Here, "custom-made" implies that we programmed the algorithm from scratch, in the programming language LabVIEW (National Instruments, Version 6.1). This is now mentioned in the text.

Page 15841, line 22. "with significant occurrence". What does significant mean here, how was significance determined?

"Significant" was probably the wrong word. It was deleted: "The highest values were found up to 100  $\mu\text{g m}^{-3}$ ."

C11551

Page 15842, line 22, "Figure 5 illustrates the seasonal differences in the occurrence of limit exceedances." Should be monthly instead of seasonal.

The sentence was changed accordingly.

Page 15842, line 28-29, "Another reason is middle-range transport from eastern neighboring countries playing an important role considering the  $\text{PM}_{10}$  mass concentrations in eastern Germany." It is unclear here how transport from eastern countries can be responsible for the seasonal dependence of exceedances. Do this transport events have clear seasonal dependence? Please provide more information here.

Usually, westerly air advection is dominating in Germany, but there is sometimes easterly advection as well. From a climatological point of view, there is no seasonal difference in the frequency of occurrence, but only in winter time, these weather conditions can lead to a problem in  $\text{PM}_{10}$  because of the reduced mixing layer height. Furthermore, also in neighboring countries more particles due to combustion processes are emitted in winter.

Page 15851, line 18. "alike a mixture of winter, spring and fall". What does this mean, please revise?

The formulation was obviously misleading. The text was changed into: "...cluster 9 shows similarities with winter, spring and fall."

Page 15853, line 1-3. "Due to the very low winds..." This seems not to be true. Figure 6 shows that  $\text{PM}_{\text{coarse}}$  is higher for exceedance days than for non-exceedance days. Or do you mean in relative terms - please correct or explain.

Yes, this was meant in relative terms. That is why we used the words "coarse particle fraction". We added "relative" to the text.

Legend of Table 2: I don't think that "significantly equal" exists. Should be changed, e.g to "not different".

C11552

The legend was changed into "...significantly different, otherwise not different".

Figure 4. Was regression line forced through zero? If yes, it should be mentioned.

Yes. This information was added to the text.

Interactive comment on Atmos. Chem. Phys. Discuss., 11, 15831, 2011.

C11553

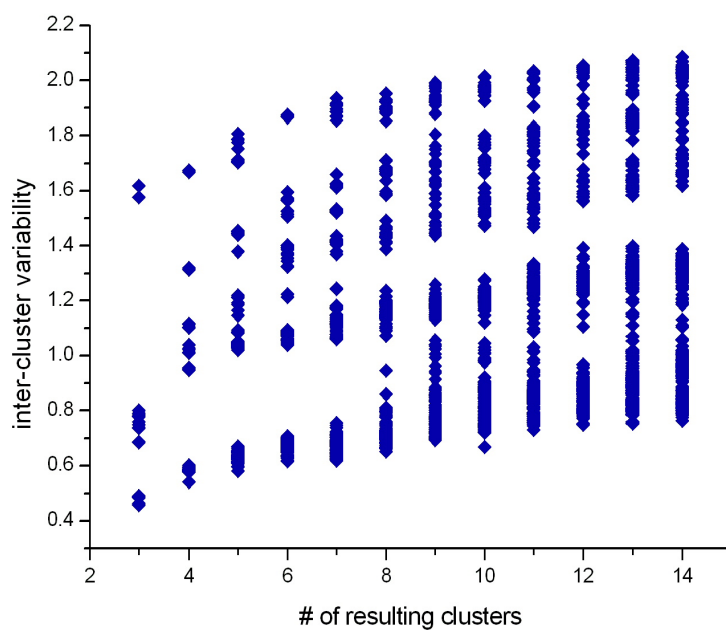
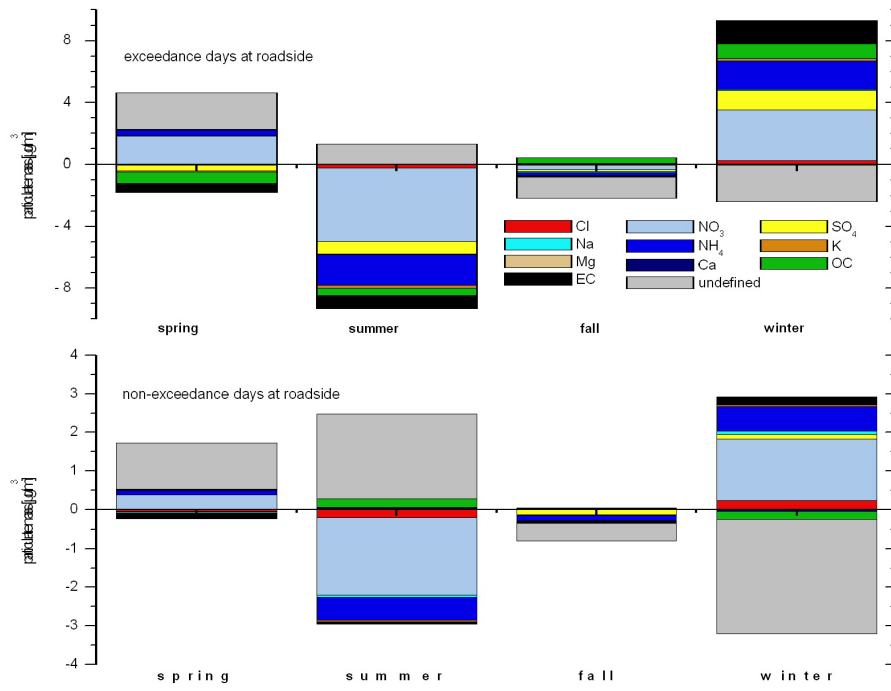


Fig. 1. Results of the sensitivity analysis.

C11554



**Fig. 2.** Absolute differences from the average chemical composition over the entire time period.

C11555