

Interactive comment on “Fine and ultrafine particles in the Zürich (Switzerland) area measured with a mobile laboratory. An assessment of the seasonal and regional variation throughout a year” by N. Bukowiecki et al.

Anonymous Referee #1

Received and published: 11 July 2003

General comments:

This paper is an extremely thorough multivariate exploratory statistical analysis of an aerosol dataset collected over a period of two years using an automobile platform in the Zurich area. The statistical techniques appear to have been properly applied, with considerable care such as transforming the data to make it compatible for the statistical technique chosen. There are two major findings to the work: that temporal variability in aerosol parameters exceeds spatial variability over the time scales investigated, and that a mobile platform such as that employed may be able to substitute for network of ground-based sites for pollution monitoring and support of epidemiological studies.

Analysis of this complex data set is important and relevant to the issues associated with human exposure to pollutants. The authors have used great care in their analysis and interpretation of the data. There are some issues regarding clarity that need to be addressed before final ACP publication. As detailed at the end of this review, I would also like to bring up for discussion the issue of the overall value of exploratory statistical analysis.

Specific comments:

This paper is quite long. There are several figures that support the statistical analyses by confirming the propriety of the application of these data to the chosen techniques; these are not necessary and serve to clutter the paper. Figures 3, 4 and 5 in particular support the ANOVA but are or can be described adequately by words in the text. I recommend removing these three figures.

Clarity of ANOVA section. The ANOVA section is confusing, even after several readings. The denominator in $F=v_1/v_2$ for the date levels is not clearly defined. This v_2 is defined as the "signal variation within the individual factor levels (i.e., within every measuring day)." Is this the variance between the morning and afternoon measurements (i.e., the replicate variance)? Please explain more clearly. Also, there needs to be a summary sentence when discussing rejection of the null hypothesis, explaining what this means (that the variance between the levels exceeds that within the levels?).

The ANOVA section is touted as being a more quantitative evaluation of the data. Yet the discussion of the ANOVA results is entirely qualitative. For example, on p. 2747 line 22 it is stated that, "for a true null hypothesis, F-values are close to 1 . . .", and, on p. 2747 line 4, "N_{bk}) sticks out to be the only value below 10 and and thus is not too far away from significance". Why is there an advantage to qualitative discussion of the processed statistical data compared to qualitative discussion of the raw data? Are there probabilities that can be quantified instead of the general statements?

I don't understand what is meant by p. 2748 lines 26-28, ". . . both factors have a

[Full Screen / Esc](#)[Print Version](#)[Interactive Discussion](#)[Discussion Paper](#)

similar effect . . . giving a more uniform distribution in the area throughout the year. . . .
." The distribution of this parameter is anything but uniform (Fig. 3).

PCA section. The principal components analysis is used to find factors that explain the variance of the data. In this case, PCA found two principal components that explain most of the variance. The first appears to be dominated by general "pollution" factors that includes everything but altitude, temperature, and radiation. The explanation of the second principal is less obvious and required cluster analysis. This analysis indicates that there are three meteorological and altitude-related clusters. The discussion of the differences in size distributions between these clusters is discussed in some detail. These differences discussed in the text are very hard to see in the figures because of the small size.

Ultrafine particles sourced from vehicles are described as primary, with the justification that these particles are rapidly formed in the exhaust plume of the engines. However, there is also discussion of lack of evidence of suppression of nucleation of these particles (p. 2758, line 27). It is a little confusing to talk about nucleation in the atmosphere of "primary" particles; other than soot, most of these particles are in fact secondary, and at the temporal and spatial scales of the mobile measurement it might be better to refer to them always as secondary.

Technical comments:

The Capaldo and Pandis reference is "grey" literature not readily accessible.

Many of the figures have lines, symbols, and text that are extraordinarily difficult to read because of their small size. The worst cases include Figs. 7, 8, 15 and 16—these should be twice the size, and the legends, symbols, and line thicknesses must be bigger yet. Also, in Fig. 7, it is not clear which data are from the ground sites and which are from the (parked) mobile lab, so clarification in the legend or figure caption is needed.

[Full Screen / Esc](#)[Print Version](#)[Interactive Discussion](#)[Discussion Paper](#)

The broken axes in figures 14 and 15 are annoying and should be removed.

Additional comments: General comment regarding multivariate exploratory analysis: Finally, I would like to offer for discussion the following question: is there much added value to evaluation of data such as these by the type of exploratory statistical analysis used in this paper—especially since the interpretation of the statistics is mainly qualitative? Many of the qualitative findings are immediately obvious by inspection of Figure 2, and the complexity of ANOVA and PCA may serve mainly to obscure the obvious. To evaluate the questions more quantitatively, could not one simply group the dataset using physically based arguments (e.g., taking using the Table 2 definitions and grouping the data by stability, temperature, season, radiation, wind) and compare the normal and percentile statistics of these subsets? What has been gained over this simpler approach?

Interactive comment on Atmos. Chem. Phys. Discuss., 3, 2739, 2003.

[Full Screen / Esc](#)[Print Version](#)[Interactive Discussion](#)[Discussion Paper](#)