Atmos. Chem. Phys. Discuss., 8, S141–S142, 2008
www.atmos-chem-phys-discuss.net/8/S141/2008/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Factor analytical modeling of C_2–C_7 hydrocarbon sources at an urban background site in Zurich (Switzerland): changes between 1993–1994 and 2005–2006" *by* V. A. Lanz et al.

V. A. Lanz et al.

Received and published: 8 February 2008

I would like to thank Stéphane Sauvage for his comment. It will certainly help to clarify some aspects of our manuscript.

We did not follow the regression approach he mentioned in his note. In our study, t-NMVOC was not regressed on the PMF factor scores, as we did not assume that PMF factors calculated for hydrocarbons are representative for all other VOC classes as well. The percentages provided in Sect. 3.2 characterize the explained variation of the 13 hydrocarbons that were measured during both campaigns (2005-06 and 1993-94, respectively) and included in the factor analytical model. The term "VOC concentra-

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

tion" (P918, L1) was not intended as a substitute for "t-NMVOC", but rather for "the concentration of the 13 hydrocarbons" and, obviously, can be misleading. Thus, in the revised paper we will avoid the term "VOC" in this context here.

The explained variability, also known as explained variation, EV, was determined for each measured hydrocarbon *j* in each factor *k*, EV(*k*,*j*), as described in Paatero (2007). By calculating the mean of EV(*j*) (93% and 96%, respectively; P917, L26), the species have equal weights, irrespective of their ambient concentrations. In contrast, on page 918, line 1 we provide an average ratio of modeled to measured total hydrocarbon concentrations (97% and 99%, respectively):

$$\overline{ratio} = (1/n) \sum_{i=1}^{n} (\sum_{k=1}^{p} \mathbf{G}_{ik} / \sum_{j=1}^{m} \mathbf{X}_{ij}), \quad \text{(AC1.1)}$$

where $\mathbf{X}_{ij}$ is the measurement matrix, $\mathbf{G}_{ik}$ represents the model including *k*=1...*p* factors (note that **F** was normalized to unity), and *n* is the number of samples consisting of *m*=13 species. This measure is only meaningful, if $\sum_{k=1}^{p} \mathbf{G}_{ik} - \sum_{j=1}^{m} \mathbf{X}_{ij} \leq 0$ is approximately fulfilled for virtually all samples and if there is little scatter in the data, which is the case here. We additionally calculated the R^2 of the linear regression, *lm*, data versus modeled values

$$lm(\sum_{j=1}^{m} \mathbf{X}_{ij} \sim \sum_{k=1}^{p} \mathbf{G}_{ik}), \quad \text{(AC1.2)}$$

yielding R^2=0.99 for both periods (2005-06 and 1993-94, respectively). Thus, almost all variance in the total concentration of the 13 hydrocarbons can be explained by PMF. We will highlight these differences accordingly in the revised manuscript and replace *p* by *k* in Eq. 2 (P913, L16).

Interactive comment on Atmos. Chem. Phys. Discuss., 8, 907, 2008.