Atmos. Chem. Phys. Discuss., 13, C12988–C12992, 2014
www.atmos-chem-phys-discuss.net/13/C12988/2014/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Primary and secondary biomass burning aerosols determined by proton nuclear magnetic resonance (H-NMR) spectroscopy during the 2008 EUCAARI campaign in the Po Valley (Italy)" *by* M. Paglione et al.

M. Paglione et al.

m.paglione@isac.cnr.it

Received and published: 19 March 2014

The authors would like to thank Anonymous Referee #2 for his/her comments.

The Referee's comments followed by our replies are listed below.

Comment: I have one comment on factor analysis of NMR. The number of samples is 17, which appears too few for factor analysis. The high Q/Qexpected (≈20, Fig. 5) for 5-factor solution indicates a large underestimation of error which should be addressed

in the text.

Authors Reply. The definition of a minimum dataset size in factor analysis is an issue not yet resolved univocally: even if there are a lot of different theoretical rules (Arrindell & van der Ende (1985), Velicer & Fava (1998), and MacCallum et al. (1999) have reviewed many of these recommendations), many studies have demonstrated that the general rules of thumb of the minimum sample size are not always valid and useful (MacCallum et al., 1999; Preacher & MacCallum, 2002). The minimum level of N (dataset size) was object of a very high number of studies that demonstrated its dependency on other aspects of design, such as: the communality of the variables (percent of variance in a given variable explained by all the factors jointly and interpreted as the reliability of the indicator) (Hogarty et al., 2005; MacCallum et al., 2001; Costello & Osborne, 2005); the degree of overdetermination of the factor (or number of factors/number of variables) (Preacher & MacCallum, 2002; MacCallum et al., 1999); etc. For the above reasons, we chosen the most appropriate factor number on the comparison between the outcomes of distinct factor analysis methods. The best agreement was achieved by far for the 5-factor solution, and this remains the most realistic NMR spectral deconvolution in this experiment. In fact, the present study is based on a short timeline of samples (# = 17), but the variability in NMR composition was high, and this is witnessed by the poor correlations between the spectral profiles (except between the two "polysubstituted aliphatics"). In turn, the variability in the chemical composition was influenced by the great variability in the weather and atmospheric dynamic conditions which characterized this field campaign. In respect to the meaning of the diagnostic Q/Qexp we copy here our reply to one of the Refereee#1's criticism: It should be noted that Figure 5 of the original manuscript reported the Q/Qexp for the MCR and NMF methods, not PMF. These algorithms do not account for the measurement uncertainty and the function (Q) to minimize is the sum of squares:

(see Fig. 1 for the equation)

The fact that the above sum is not normalized by the measurement uncertainties

ACPD

13, C12988–C12992, 2014

means that Q is not adimensional, and its actual value depends on the units used in the sample concentration matrix. Therefore, in general the ratio Q/Qexp, where Qexp is defined here analogously than for PMF, will not converge to 1. We acknowledge the fact that the use of the ratio Q/Qexp as a diagnostic for methods other than PMF can be misleading. For this reason, we now simply report the variation of Q with the number of factors as a diagnostic for the factor analysis performed using the MCR and NMF techniques.

Additional references

Arrindell, W. A., & van der Ende. J.: An empirical test of the utility of the observations-to-variables ratio in factor and components analysis. Applied Psychological Measurement, 9, 165 – 178, 1985.

Costello, A. B., & Osborne, J. W.: Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. Practical Assessment Research & Evaluation, 10(7), 2005. from http://pareonline.net/pdf/v10n7a.pdf

Hogarty, K. Y., Hines, C. V., Kromrey, J. D., Ferron, J. M., & Mumford K. R.: The quality of factor solutions in exploratory factor analysis: The influence of sample size, communality, and overdetermination. Educational and Psychological Measurement, 65, 202-226, 2005.

MacCallum, R. C., Widaman, K. F., Preacher, K. J., & Hong S.: Sample size in factor analysis: The role of model error. Multivariate Behavioral Research, 36, 611-637, 2001.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong S.: Sample size in factor analysis. Psychological Methods, 4, 84-99, 1999.

Preacher, K. J., & MacCallum, R. C.: Exploratory Factor Analysis in Behavior Genetics Research: Factor Recovery with Small Sample Sizes. Behavior Genetics, 32, 153-161, 2002.

ACPD

13, C12988–C12992, 2014

Velicer, W. F., & Fava, J. L.: Effects of variable and subject sampling on factor pattern recovery. Psychological Methods, 3, 231-251, 1998.

Interactive comment on Atmos. Chem. Phys. Discuss., 13, 33343, 2013.

ACPD

13, C12988–C12992, 2014

$$Q^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} (x_{i,j} - g_{i,k} * f_{k,j})^2$$

Fig. 1. Object function Q representing the sum of squared residuals

ACPD
13, C12988–C12992, 2014