Atmos. Meas. Tech. Discuss., 7, C1924–C1926, 2014
www.atmos-meas-tech-discuss.net/7/C1924/2014/

Atmospheric Measurement Techniques Discussions

Interactive comment on "Using self organising maps to explore ozone profile validation results – SCIAMACHY limb compared to ground-based lidar observations" *by* J. A. E. van Gijsel et al.

Anonymous Referee #2

Received and published: 29 July 2014

General remarks:

The paper describes a methodology to apply the mathematical technics of self organizing maps (SOMs) to explore ozone profile validation results of SCIAMACHY limb compared to ground-based lidar observations of 7 different NDACC-stations. The aim is to study the influence of a set of different observational characteristics such as geographical location, solar zenith angle etc. In itself an interesting study and could be appropriate for publication in AMTD. Although the major content of the paper is well structurized and written, it still suffering that essential information and critical analysis of the data is missing. The use of SOMs in the comparison of different data sets is

not that new and since more than 10-20 years well known and also well established in atmospheric sciences, e.g http://www.cis.hut.fi/projects/somtoolbox/. This is also the SOMs-toolbox the authors have used for their study. In the present form the paper describes only the methodology of applying SOM's in mathematical terms, however, the compilation of mathematics into physical relevant terms is rather poor. Particularly in chapter 3 the physical base is almost completely missing. The paper should contain a more precise and thorough analysis based on interpretation in physical terms of the obtained results. In addition, the final results should also be (quantitatively) discussed in comparison with results obtained by other investigators using traditional validation methods. Further, it seems that the number of vertical profiles of each Lidar station applied in the analysis are not equally distributed in time and space and so introducing artefacts which are finally dominating the results of the SOMs-analysis such that consequently the conclusions made are neither adequate. This would mean that the analysis and the interpretation of the results has to be re-done. Therefore, I rate the paper as being only acceptable for publication after major revisions as also written in my specific comments below.

Specific Comments:

Below I have listed my major points of critics. Most essential is that the compilation of SOM-mathematics into physical terms is very poor and mostly missing and certainly not discussed adequately

Section 2.3: Physical base is missing

Section 3.1: Missing: number of lidar profiles per station finally used. In how far homogeneity of Lidar and Sciamachy vertical profiles is a pre-requisite?. Should number of difference profiles from each lidar station be the same and identical distributed in space and time? Page 4381, Line 14: What do you mean with matching meta data. Please explain in physical terms.

Section 3.2: Physical base is missing. P4382,L02: Why 46x75 and why hexagonal

AMTD
7, C1924–C1926, 2014

neurons? P4382,L03: What are “input vectors” in physical terms. Figure 5: What are the horizontal and vertical axis representing? P4382,L25: What are “codebook vectors” in physical terms.

Section 3.3: P4383,L12: The explanatory variables (EV) should be here discussed in more detail. What kind of meta-data information for each EV has been used in the mapping the EV-planes. This is essential in order to interprete the outcoming results. Figure 6: a.) What are the horizontal and vertical axes representing? b.) The pattern of the “Location” map is dominated by HOH, MAU and LAU which is also seen in “Latitude” and “Longitude” map. Less clearly but still visible a similar pattern is seen in the “Solar zenith angle” and “Solar azimuth angle”. All other maps show no significant pattern. Knowing that OHP is located very close to HOH I would expect both station represented more equally.

It seems that the number of vertical profiles of each lidar station applied in the analysis are not equally distributed in time and space and so introducing artefacts which are dominating the results of the SOMs-analysis. In general the question is raised: Does it make sense to have the seven lidar stations together with latitude and longitude as three independent EVs? Latitude and longitude are not really independent parameters but pre-dominated and directly linked to the geographical locations (incl. matching criteria of 800 km) of the 7 Lidar sites?

Section 3.5: P4385,L21: Explain k-means in physical terms? P4386: What is the physical meaning of the 3 clusters obtained and what they represent? Chapter 4: change title “Summary and Conclusions”

Interactive comment on Atmos. Meas. Tech. Discuss., 7, 4373, 2014.