

***Interactive comment on “Exploratory data analysis and clustering of multivariate spatial hydrogeological data by means of GEO3DSOM, a variant of Kohonen’s Self-Organizing Map” by L. Peeters et al.***

**L. Peeters et al.**

Received and published: 10 November 2006

The authors would like to thank the editor for the valuable contribution in the reviewing process of the submitted paper.

The suggested minor editorial comments will be incorporated in the revised manuscript.

With regards to the suggestion to stress the advantages and disadvantages of

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

K-means and fuzzy C-means clustering if compared to SOM, following paragraph will be added in the revised manuscript:

The Self-Organizing Map algorithm is closely related to K-means clustering. A SOM with the number of units equal to the number of clusters in the dataset and a neighborhood equal to zero will act as a traditional clustering technique (Kaski, 1997). A SOM may, however, be used in two very distinct ways: a large SOM, also known as emergent SOM, with many units, used for exploratory data analysis and cluster detection (Ultsch and Hermann, 2005), and a small SOM for cluster centroid determination (Bacao et al., 2005). In this study, both SOM and GEO3DSOM are used for exploratory detection of clusters. When using K-means and fuzzy C-means-clustering, cluster centroids will always be detected based on the objective criterium of sum-of-squared distances. In emergent SOM's on the other hand, clearly separated groups of units may or may not be detected. Small SOM's used for centroid determination will act as a robust K-means initialization in the first training iterations and due to the decrease of learning radius and neighborhood during training, the SOM will perform exactly as a K-means clustering in the final steps of the learning process (Bacao et al., 2005).

Compared to K-means clustering and fuzzy c-means clustering, SOM has, in addition to the ability of SOM to directly visualize the results of the clusters in terms of the original variables, the advantage that the number of clusters does not need to be specified a priori. The advantage of K-means clustering and fuzzy c-means clustering on the other hand, is the ease of implementation since there are less parameters to be chosen.

On the performance of clustering of SOM compared to other techniques a debate still exists in literature (overviews can be found in Bacao et al., 2005 and Mingoti et al., 2006). Provided the SOM is parameterized correctly, SOM will outperform K-means clustering since SOM is less sensitive to local optima compared to K-means (Bacao et al., 2005).

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

## References:

Bacao, F., Lobo, V. and Painho, M., 2005. Self-organizing maps as substitute for K-means clustering. In: V. Sunderarm, G. van Albada, P. Sloot and J. Dongarra (Editors), International conference on computational science. Lecture Notes in Computer Science. Springer-Verlag Berlin, Heidelberg, pp. 476-483.

Kaski S, (1997) Data exploration using Self-Organizing Maps. Acta Polytechnica Scandinavica: Mathematics, computing and management in engineering, Series No 82 57

Mingoti SA and Lima JO, (2006) Comparing SOM neural network with Fuzzy c-means, K-means and traditional hierarchical clustering algorithms. European Journal of Operational Research 174:1742-1759

Ultsch, A. and Herrmann, L., 2005. The architecture of emergent self-organizing maps to reduce projection errors, ESANN2005: 13th European Symposium on Artificial Neural Networks, Bruges, Belgium, pp. 1-6.

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 3, 1487, 2006.

**HESSD**

3, S1432–S1434, 2006

---

Interactive  
Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper