Development and application of the STARS tool to optimise the schematisation for GeoPEARL

F. van den Berg, A. Tiktak, D. van Kraalingen and J.J.T.I Boesten

Background

In the past decade the interest in GIS model applications to calculate pesticide behaviour has increased. For the assessment of leaching of pesticides to groundwater in the Netherlands, the GeoPEARL model has been developed using a schematisation based on GIS-data on e.g. soil type and land. The first version of GeoPEARL has been released in 2004 and it has since then been used in the pesticide registration procedure in the Netherlands. In recent years efforts have been made to derive schematisations for other EU countries as well as schematisations for leaching assessments at the European scale.

Objective

Performing model calculations for every individual pixel would require too much computation time. Moreover, more detailed GIS data have become available, for instance the EFSA data (2012, see Figure 1), the LUCAS dataset (2013) and AGR4CAST (2014). Therefore there is a need for tool to optimize the schematization, so calculations can be done for a limited number of unique combinations.

Methods

The derivation of these unique combinations is preferably done using statistical methods. We therefore developed the STARS tool (Schematisation Tool for Areas at All Scales), which uses k-means clustering (Hartigan & Wong, 1979) to derive unique combinations based on a set of grid maps with the same spatial resolution. The STARS tool can handle continuous and categorical maps.

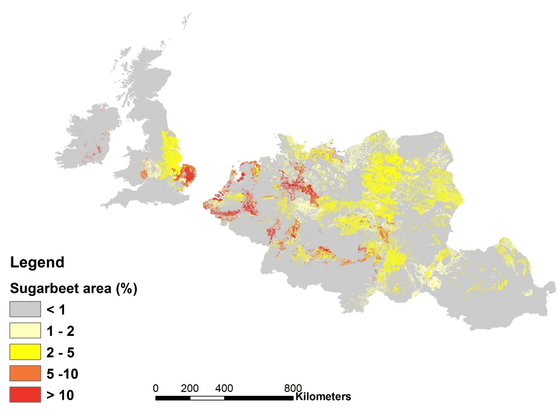

Figure 1. Sugarbeet area in the Centre Zone of the EU

The procedure for the use of STARS is as follows:

- Import data layer files for all properties for the cluster analysis
- Create a project with the data layers required
- Specify the number of clusters and weighting factors of the property in each numerical layer
- Execute cluster analysis (using R) and export results to csv or ascii files; different options for cluster output are available.

A first test of the STARS tool has been done using GIS data for the EU and limiting the area to be considered to the Netherlands. The properties included in the analysis were the organic matter content in the top soil (numerical), the annual mean temperature (numerical), the annual mean precipitation (numerical) and the topsoil texture (categorical, 5 classes). The weighting factor was set to 1 for all numerical properties. The clustering was done for a total of numerical clusters of 50, 500 and 2500 to study the effect of the number of clusters on the variability in organic matter content within a cluster.

Results

The results of the clustering to obtain 50, 500 or 2500 numerical clusters partitioned over all categorical clusters for the test set of data layers for the Netherlands are shown in Table 1.

Table 1. The partitioning of numerical clusters over the category clusters for a total of numerical clusters of 50, 500 and 2500.

Texture	Number of clusters N=50	N=500	N=2500	Number of pixels in category cluster
Medium	16	154	813	10562
Coarse	22	222	1175	15266
Medium fine	1	9	46	593
Fine	7	67	355	4607
Peat	6	55	291	3773

The homogeneity of the cluster properties increases substantially when increasing the number of clusters from 50 to 2500, as is shown in Figure 2 (left-hand side graph).

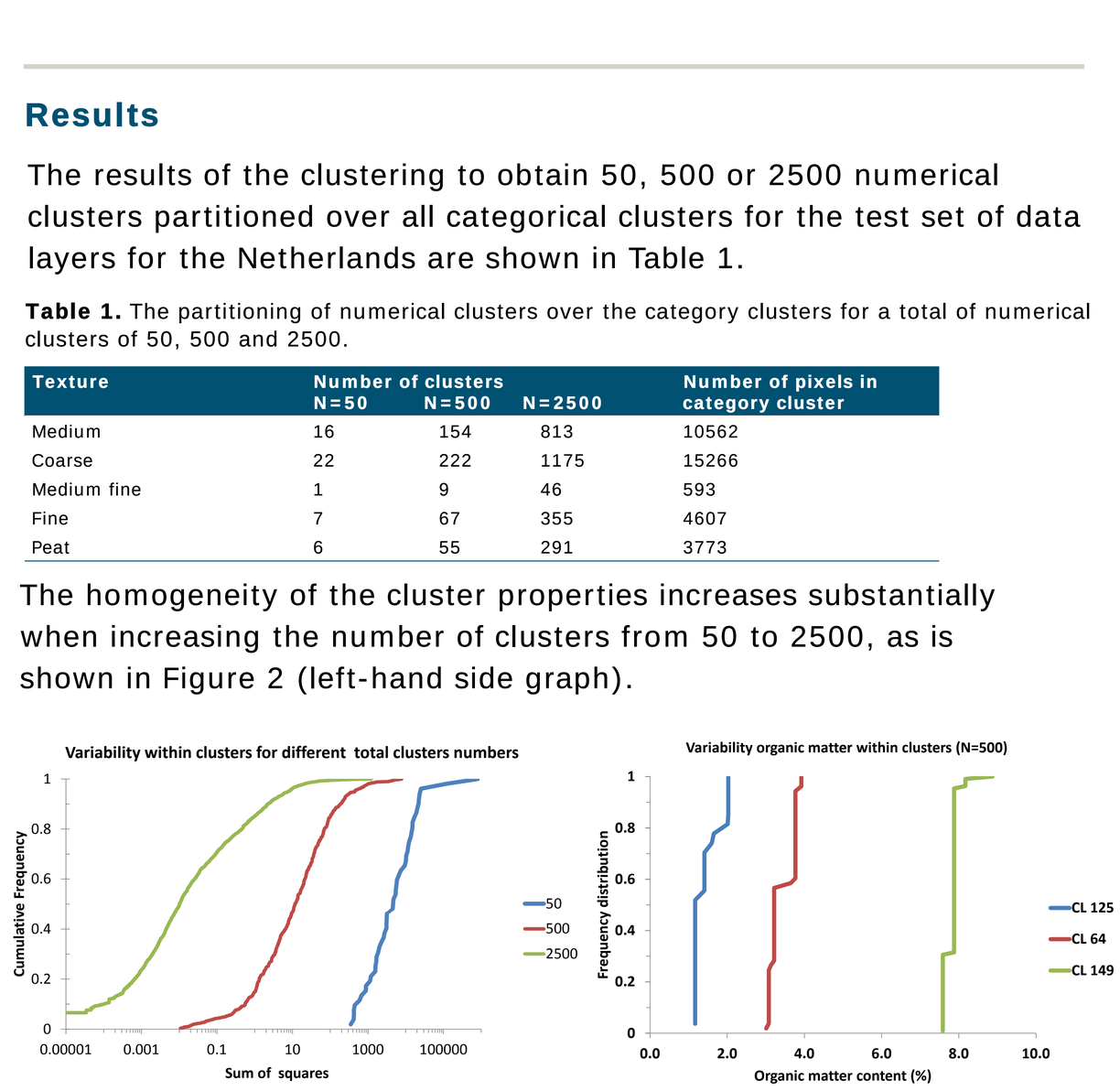

Figure 2. Frequency distribution of sum of squares for selected number of clusters (left) and that of organic matter content (right) for three clusters selected within one category cluster (i.e. soil with medium size texture). Total number of pixels 34800.

Limiting the number of clusters to 500 already results to a comparatively low variability in organic matter content, as shown for clusters 125 (27 pixels), 64 (53 pixels) and 149 (108 pixels) in the categorical cluster of pixels with medium-textured topsoil, see Figure 2 (right-hand side graph).

At present the STARS tool is being developed further to post-process the output of the cluster analysis to generate GeoPEARL input.

The schematisation can be optimized by preparing the schematisation files by changing the number of clusters and/or number of pixels per cluster, and comparing GeoPEARL output for different schematisations.

Conclusions

- STARS is a promising and very flexible tool to facilitate the process of deriving spatial schematisations at different scales
- The development of post processing procedures and their implementation in STARS would greatly enhance its usefulness to generate optimized spatial schematisations

Acknowledgements

We thank the Dutch Ministry of Economic Affairs for funding this research within the framework of research programme BO-20-002.

Wageningen University & Research
P.O. Box 123, 6700 AB Wageningen
Contact: erik.vandenberg@wur.nl
T + 31 (0)317 48 18 54
www.wur.nl