Finding profiles of farm profitability using clustering of the self-organizing maps

Mika Sulkava
MTT Agrifood Research Finland
Economic Research

25.11.2013

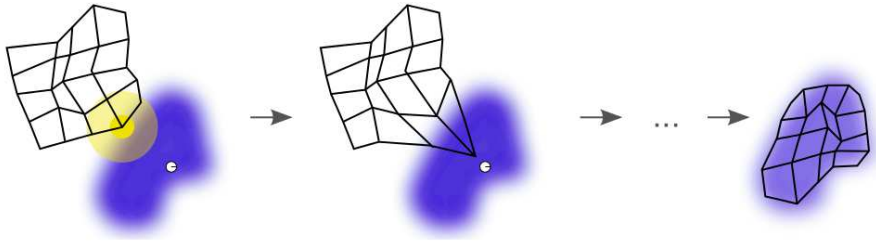
Contents

- Self-organizing map
- Clustering
- Farm bookkeeping data
- Clustering of farm bookkeeping data
- Summary

 25.11.2013 2

Self-organizing map (SOM)

- An artificial neural network
- Effective for analyzing large amounts of multidimensional data
- Projects complex data on a two-dimensional map
- Illustrative visualization of data
- Flexible net in data space

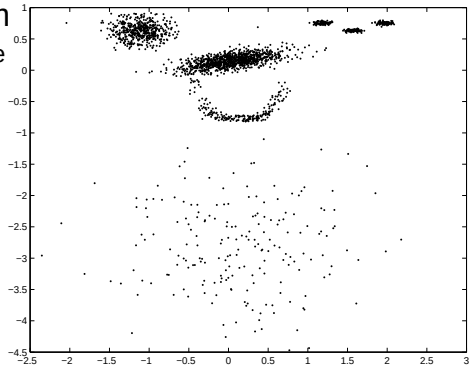


Self-organizing map (SOM)

- Exploratory data analysis
- Generate new hypotheses
- Find interconnections between variables
- Study data at different resolutions
- Find clusters
- Study changes in time

Clustering

- Divide the data into homogeneous groups so that the groups are different from each other
- Many algorithms: k-means, hierarchical methods, etc.
- We used the VS-algorithm
  - Clusters the data based on the self-organizing map
  - A hierarchical, density based approach

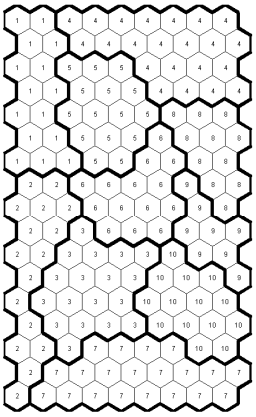


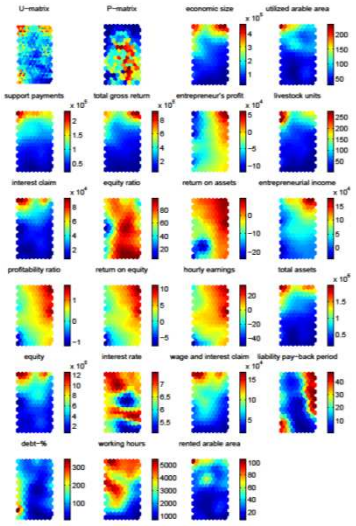
Farm bookkeeping data

- Data collected from Finland since 1912
- About 1000 farms
- Thousands of variables
- We used data from 2010
- 25 variables selected for analysis:
  - Ecomic size
  - Utilized arable area (UAA)
  - Support payments
  - Total gross return
  - Entrepreneur's profit
  - Livestock units
  - …

Clustering of farm bookkeeping data

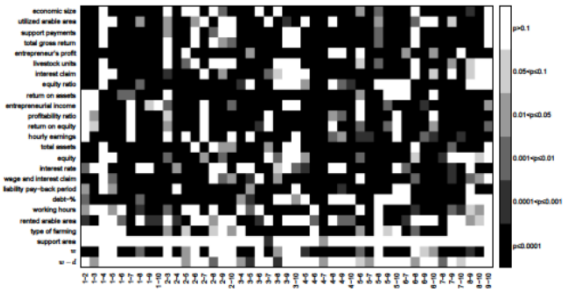
- Ten clusters were obtained

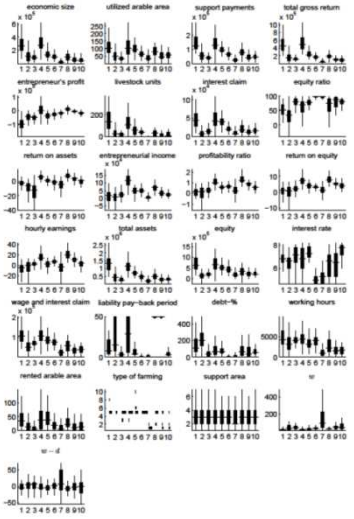




Clustering of farm bookkeeping data

- The clusters are mostly statistically significantly different from each other





Clustering of farm bookkeeping data

1. Large farms with large UAA, low return on equity, low profitability, and high debt-%
2. Mid-sized farms with very high debt-%, low equity ratio, and low profitability
3. Farms with very low UAA, low total assets, lot of working hours compared to economic size, very low return on assets, very low profitability. Relatively often indoor horticulture
4. Large farms with high UAA, high return on equity, high profitability, and medium debt-%

Clustering of farm bookkeeping data

5. Mid-sized, high-risk farms with large utilized area and especially large rented utilized area. Often cattle farms
6. Mid-sized, low-risk farms with small debt-%, very high equity ratio. Almost only dairy farms
7. Small, low-risk farms, which are mostly cereal farms and have very small UAA, very low debt-%, low total assets, very few working hours and livestock units, and very low profitability
8. Mid-sized high-profitability farms, which have larger utilized area and less working hours than other groups of farms of similar size

Clustering of farm bookkeeping data

9. Small farms with high profitability, which are rather seldom dairy farms but quite often cereal farms
10. Small high-risk farms with average profitability and short liability pay-back period

Summary

- Self-organizing map is useful for exploring complex, multidimensional data
- Suitable for analyzing farm bookkeping data
- Characterize Finnish farming using 10 clusters
- Complex relationships between profitability, debts, economic size and other variables
- Paper about SOM-analysis of 25-D farm data
  - M. Sulkava, A.-M. Sepponen, M. Yli-Heikkilä, A. Latukka. Clustering of the Self-Organizing Map Reveals Profiles of Farm Profitability and Upscaling Weights. Submitted to Neurocomputing.
- Future: Data from other years, study changes in time