Extracting information from regional data using Self-**Organizing Maps**

forum stat

Legend:

age ratio

com p.c. i

population densit

(per

ate

K-means clustering of SOM

mapping (12 clusters)

hotel density 1000

ousiness enterprises

holdings (per ' agricult. a

Markus Casper¹, Marcus Herbst¹

...Bar plane"

الموبا للجود للجود للجود للجود الجود الجود الجود الجود الأجل

ألخوه ألخره الخرف الخرف الجرف الجرفي الجرفي ألجرم الجرمي الجرفي الجرفي

ألتهم أبتوه التؤه التوه التوهر التوهر التوم التوم التوم التوق عتوق عنوق

اللؤم اللؤم اللؤم اللؤم اللؤم اللؤم اللؤم اللؤم اللؤم المؤم المؤق المؤق المؤق

ألجه الجه الجه الجه الخهم الخيم الخيم الخيم الخيم الخيم الخيم الخيم

فجاه غجاه غجه عجه عجه عجه عجه عجم عجم عجم عجم عجم عجم عجم

أنجب أنجد لنجد لنجد لتجد تتجد تتجد لنجد لنجد انجد انجد عجل عجل

ألؤه ألؤه الؤه الؤه الجه الجه تجه بنهم بنهم بنهم عيم عيش عيش

ألؤم ألؤم للؤم للؤم للؤم للزم للزم بلزم بلزم بلزم بلزم عزش عزش

غاره غاره غاره عاره عاره عاره عاره عاره مارد مارد مارد مارد

اللبأيها للبارة البارة البارة البارة البارة البارة البارة البارة البارة ا

المأو المأو المارة المارة المارة المارة المارة المارة المارة المارة

اللج للج فلم فلم فلم فلم فلم ملم ملم ملم ملم

اللبلا للبلا خليه خليه خليه عليه عليه عليه عليه مليه مليه مليه

حاط حارم خارم حارم حالم حالم عارم عارم عارم عارم عارم

مطط مطط فطع فطع حظم حظم عظم علام علام علم علام عرم

أحفظ أحفار أحفار أحفار حافر حافر خازم خازم عازم عازار عازار

النفط النفط النفعر فنقفر خطفر خطفر خلفم خلافر خازها عازق عاراز وأرزار

Mapping of "similar" communities in Rhineland-

Conclusions

e.g. "correlation hunting"

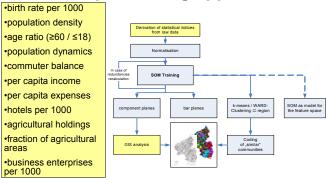
SOM for subsequent mapping in space

Palatinate

Problem

The extraction of information from regional statistical data involves processing high-dimensional data sets (i.e. with many variables) in such a way that similarities between the individual spatial units and relationships in their characteristics become accessible to our cognitive abilities.

Data and processing approach

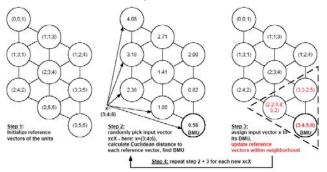


Training of a Self-Organizing Map

A Self-Organizing Map (SOM), consisting of nodes *i=1...k* which are characterized by a vector

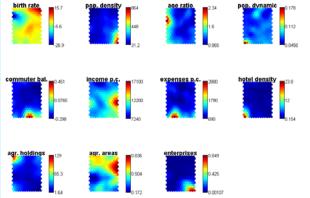
 $\boldsymbol{m}_i = \left[\boldsymbol{\mu}_{i1}, \boldsymbol{\mu}_{i2}, \dots, \boldsymbol{\mu}_{in}\right]^T \in \Re^n$

with the same dimensionality as the data sets $\mathbf{x} \in X$ is trained:



Results

"Component planes" (i.e. distribution for individual variables)



similarity of socio-economic characteristics

· the SOM also allows extrapolating new data sets

regions, rural areas, peripheral communities etc.

· SOM allows topologically ordered projection of highdimensional data by similarity (i.e. similarity in data characteristics translates into distance on the SOM)

· this allows an intuitive exploration of regional statistical data,

· k-means clustering helps delineating "similar" regions on the

• the clusters can be interpreted as urbanized regions, forested

• it can be seen that spatial vicinity does not always imply

Further Information on Self-Organizing Maps and their application

This work was carried out using the SOM-Toolbox for Matlab by the "SOM Toolbox Team", Helsinki University of Technology (http://www.cis.hut.fi/projects/somtoolbox) Kohonen, T.: Self-Organizing Maps,3rd ed., Berlin, Heidelberg, New York, 501pp., 2001. Herbst, M., Casper, M.C.: Towards model evaluation using Self-Organizing Maps, Hydrol. Earth Syst. Sci., 12, 657-667, 2008. Herbst, M., Gupta, H.V.: and Casper, M.C.: Mapping model behaviour using Self-Organizing Maps, Hydrol. Earth Syst. Sci., 13, 395-409, 2009. Herbst, M., Casper, M.C.; Grundmann, J. and Buchholz, O.: Comparative analysis of model behaviour for flood prediction purposes using Self-Organizing Maps, Nat. Hazards Earth Syst. Sci., 9, 373-392, 2009.



¹Physical Geography, University of Trier, 54286 Trier, Germany e-mail: casper@uni-trier.de, Tel.: +49 651 2014518, Fax: +49 651 2013976

