Population Patterns in Switzerland 1850-2000

Martin Behnisch and Alfred Ultsch

Abstract Spatial planning and quantitative geography are far from adequate handling the growing amount of geospatial data and statistics. Techniques of Data Mining and Knowledge Discovery are therefore presented to examine by time intervals (=15 decades) the population development of 2896 Swiss communities. The key questions are how many patterns will occur and what are their characteristics? Relative difference (RelDiff) is proposed as an alternative to relative change calculation. Based on mixture models and posterior probabilities by decade the Bayesian theorem allows the detection of patterns. A procedure of information optimization aims to select relevant patterns for clustering. The use of a k-Nearest Neighbor Classifier is based on the assumption that similar relevant patterns are a good point of reference for the whole Swiss population development. The classification result is explained by significance with already existing classifications (e.g. central-periphery approach). Localisation leads to the verification in mind of the spatial analyst and provides the process of knowledge conversion.

1 Introduction

Urban planners and politicians have several impressions about recent problems of Swiss population losses in peripheral alpine regions as well as about the urban sprawl in the Midland [7]. The long-term development of all Swiss communities is not really quantified and not present in actual planning and decision processes [5]. Some temporal classifications [6], [1] have been established for the alpine regions based on hierarchical clustering. It is to emphasize that big and incomparable

Dr.-Ing. Martin Behnisch
Databionics Research Group, University of Marburg e-mail: Behnisch@informatik.uni-marburg.de

Prof. Dr. Alfred Ultsch
Databionics Research Group, University of Marburg e-mail: Ultsch@informatik.uni-marburg.de

variances dominate the calculation of Euclidean distances when values are not normally distributed. A deeper explanation of cluster is sometimes missing and such approaches did not take into account knowledge discovery techniques to trigger spatial abstractions. The aim of this contribition is to discover long-term developments using characteristics of 15 decades in Switzerland. Due to the lack of several long-term data dimensions the development of population between 1850 and 2000 is analyzed as a kind of overall indicator for such development of Swiss communities.

2 Data Inspection

The database of this contribution is characterized by data of official statistics (Swiss Federal Population Census, [9]). Several indexes are already in use to measure the change in population. As an alternative to relative change calculation the author suggest relative differences. The influence of extrem values on relative difference is alleviated due to an symmetric and limited range. This index is particularly suitable for normalisation and standardization [3]. Nevertheless data inspection leads to the awareness that there are 15 distributions of population change with difficult (wide) margins. A similarity measure is necessary, that allows to generate a typology of population dynamics. The idea is to model each of all 15 distributions as a mixture of three characteristic developments: losing communities (e.g. multiplicative process), typical communities (e.g. sum of many unobserved random population is acting independently, CLT theorem), winning communities (e.g. multiplicative process, growth). The mixture model is realized as a composite of a log-normal, normal, log-normal distribution (see Fig. 1). The expectation maximization (EM-) algorithm is used for parameter computation. 'Good' initial parameters are important as the algorithm only finds a local and not a global optimum. Several re-calculations are necessary to proof the intermediate results. In addition to other mixture models pareto density estimation (PDE, [8]) and probability density functions (PDF) are used to ensure the modeling process. The modeled distribution is proofed by Q-Q-plots. It is to mention that a two-stage modeling procedure is realized. It takes into account that the mixture distributions by decade are different and need to be comparable. Thus the distribution of typical communities is first modelled as a Gaussian (normal) distribution. The detection of the mean and standard deviation characterize the typical Swiss population change by decade. It is a kind of clinical thermometer (15 points in time). The second modelling stage deals with the whole mixture model. The mean and standard deviation of the first modeling process are the basis for the z-Transformation. The standardization provides the comparison of decades and clustering. In consideration of the distribution of typical communities as a normal one, it is further known that the z-transformed data belongs to a standard normal distribution. The mean and standard deviation of typical communities are therefore valuable to control the modelling process of the whole mixture model (M=0; S=1). Finally each distribution has its own characteristic parameters by decade: Mean, standard deviation and amount of communities.

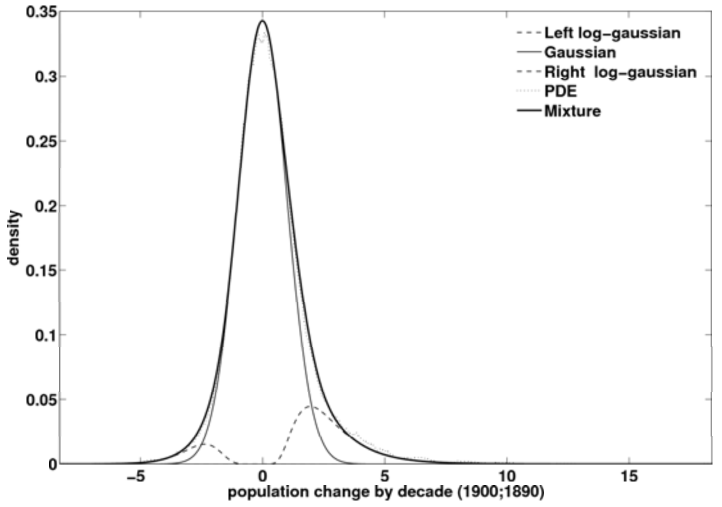

Fig. 1 Mixture model of population change by decade

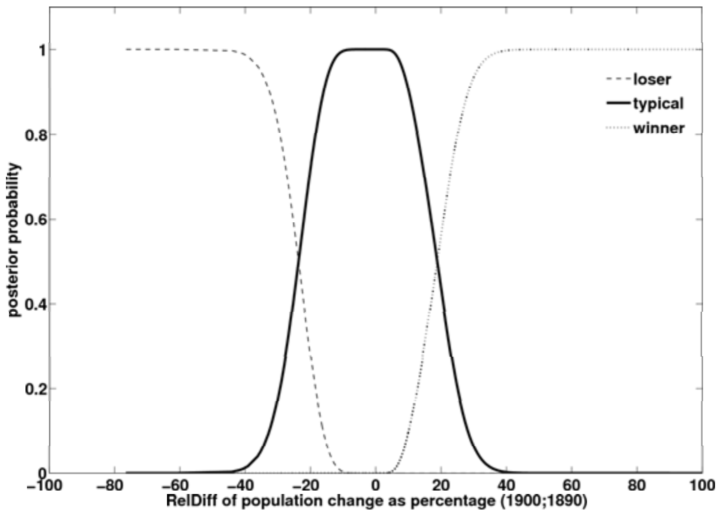

Fig. 2 Posterior probabilities based on population change and dynamic classes

3 Patterns of Population Change

The posterior probability and component probability density function are used to compute the posterior probabilities using Bayesian theorem. This theorem offers advantages through its ability to formally incorporate prior knowledge into model specification via prior distributions and allows considering the variability. A specific dynamic class (e.g. Winner, Typical, Loser) is therefore predominant observed for a given value of population change in a community (see Fig. 2). Abstract numerical values (-1,0,+1) are supplemented to represent one of the dynamic classes by decade. For a better characterization of communities the term pattern is defined to describe specific developments consisting of 15 unique dynamic classes (=15 decades). The Euclidean distance is appropriate for posteriors and now reasonable in view of a clear distinction. A complex variance estimation is not necessary. In view of 15 decades and 3 dynamic classes there are 880 different patterns finally observed. There is one large pattern (observed in 852 communities) clearly representing the "Typical" Swiss population development over time (15 x "Typical"). Many other patterns are characterized by severals "Typicals". There are about 775 patterns which are described by only 1 or 2 communities.

4 Relevance of Patterns

The authors aim to select relevant patterns for clustering. A pragmatic planning approach is to have a deeper look to the size of population. The impact of population on one pattern is therefore related to the communities and the sum of their average values of population (longterm mean value). A procedure of information optimization is used to select the patterns based on the Pareto principle [4]. A Lorenz curve presents the association of the number of patterns and the computed impact of population. From the ideal point 0% of population impact and 100% of knowledge of the patterns the distance to the real situations on the Lorenz curve is measured. The identification mark 'a' in Fig. 3 shows the shortest of such distances. From this it is concluded that in order to gain different patterns only about 14% of the patterns, the 14% relevant ones, should be examined in deep detail. The underlying assumption of the authors is that the minimal value of the impact on a pattern is within the range of 2000 to 10,000 due to observed number of communities per pattern and the average mean value of population. The observed value is finally 5000 per pattern and 122 patterns are selected and declared as a relevant pattern. In the presented context it is to remark that 65% of all 2896 communities belong to relevant patterns and about 85 % of the Swiss population.

5 Clustering and Classification of patterns

For the purpose of clustering of relevant patterns and in particular for the interpretation the authors have defined three periodical subdivisions. These periods are identified based on knowledge about the patterns by 15 decades and the three dynamic categories. The periods are as follows: period 1 (1850-1910, industrialization

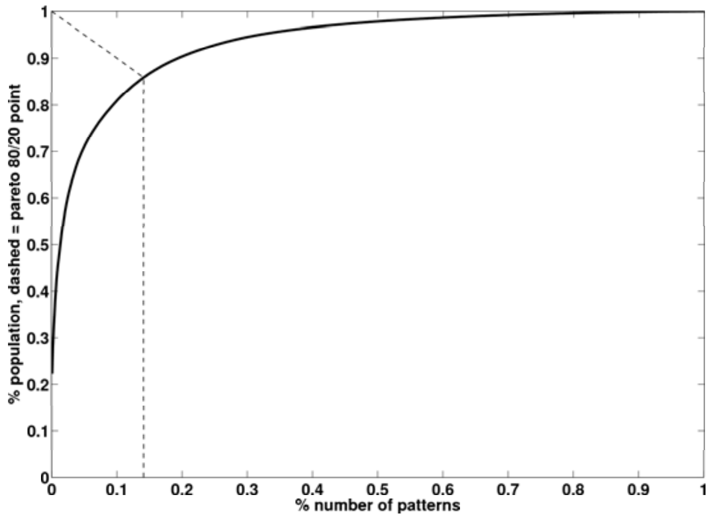

Fig. 3 Information optimization in view of relevant patterns in Switzerland

and urban growth), period 2 (1910-1950, World War I, II, subject of separation) and period 3 (1950-2000, urbanization, suburbanization, economic boom). For clustering growth indicators are defined and sum up the observed values by decade of a pattern for each period. Large positive values indicate a winner by period, values by zero indicate a typical by period and large negative values by period indicate a loser. A precise distinction of patterns is supported by the properties of the growth indicator (limited range) and related integer values (similar scale). Patterns are different from each other by a value of 1. Ward algorithm is used as a method for clustering. The information loss refers to the inner and outer cluster differences. It is conceivable when clustering 122 patterns based on information about population change that each cluster will contain about 10 patterns in average. The authors expect a solution of 6 to 12 clusters (e.g. each cluster will contain 10 patterns in average). The interpretation of the dendrogram (see Fig. 4) confirms the expectation and supports an eight cluster solution describing 1899 communities. The aim is subsequently to allocate all patterns to an observed cluster. This is the basis for a comparison with other existing spatial classifications. A k-Nearest-Neighbor classifier is constructed for this purpose. It is initially trained on the relevant pattern to identify the accuracy of allocation to the given cluster (accuracy=100%). A typical community is found for each of the 8 cluster based on the average dynamic property by period and the amount of population.

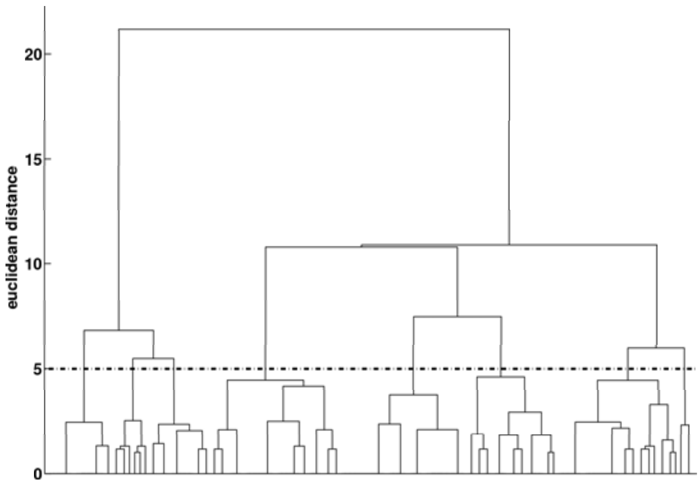

Fig. 4 Dendrogram of WARD clustering

6 Knowledge Discovery

The process of class explanation provides the transition from data to knowledge. The goal is to provide evidence-based insight through a deeper understanding of data (in the mind of the analyst) and to produce results that can be utilized at policy and strategy levels. At first localization of classes is used for spatial verification and spatial reasoning (see Fig. 5). Secondly the detected class partition and other well-known typologies [10] in Switzerland are compared using contingency tables in order to decide whether or not dependencies are significant. Relative differences [3] are here also appropriate to analyze the deviation of the expected and observed values. Third structure interpretation and validation in mind of the spatial analyst lead to knowledge about the Swiss communities (e.g. discovery of new patterns in data). Such knowledge comprises spatial abstractions of classes and generates hypothesis that might be valuable for further explanations and subsequent analysis. As an example the class "Working Suburbia" is presented as follows: Frequently winning communities (in particular before 1910 and after 1950), workplaces by significance, located in the region Lemanique, in the Midlands and significantly in the Zurich area. The expected value of Out-Commuters are not confirmed. Communities of this class are like "Dietikon" (Typical). Table 1 gives an overview of all classes and sugested spatial abstractions. Furthermore the specific population dynamic, the amount of communities and the impact of population (see sect. 4) are summarized. Each period is characterized by the growth indicator of the typical community.

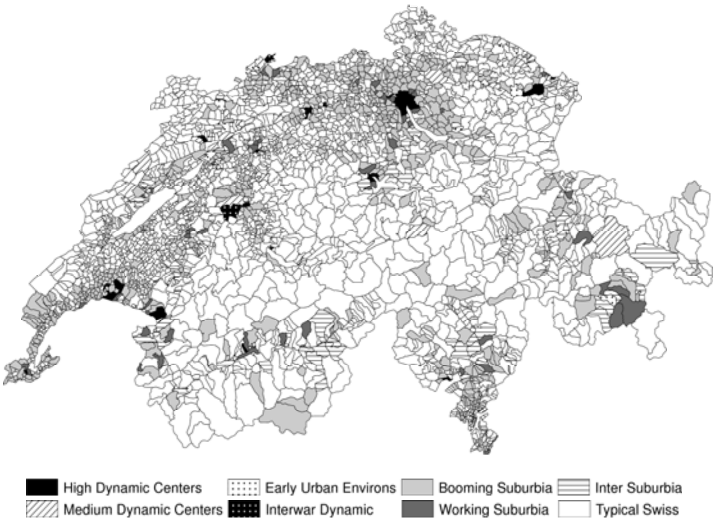

Fig. 5 Localization of the classification result

Table 1 Class properties in terms of spatial abstractions and population dynamics by period

class	size	typical	1850-1910	1910-1950	1950-2000	impact
Early urban environs	25	Gossau	Winner, +4	Typical, 0	Typical, +1	153773
High dynamic centers	10	St. Gallen	Winner, +5	Typical, 0	Typical, 0	469374
Medium dynamic centers	35	Chaux-de-Fonds	Winner, +3	Typical, 0	Typical, 0	436348
Booming Suburbia	868	Uster	Typical, 0	Typical, 0	Winner, +2	907280
Interwar Dynamic	8	Ascona	Typical, 0	Winner, 0	Typical, +2	40640
Inter Suburbia	224	Zug	Typical, +1	Typical, 0	Typical, +1	422423
Working Suburbia	55	Dietikon	Winner, +3	Winner, +1	Winner, +2	197778
Typical Swiss	1671	Solothurn	Typical, 0	Typical, 0	Typical, 0	1684991

7 Discussion

The authors are interested to examine the pool of data in depth. In particular the variance, the properties of distance measurements and the requirements of clustering algorithms are considered. In contrast to former approaches the observation of specific population dynamics is based on data properties and extracted patterns. Furthermore the authors are interested in class explanation by significance using contingency tables. The separation of high and low populated communities is observed as expected. One dynamic (15x "Typical") is clearly representing the typical Swiss population development. This dynamic is to understand as a relevant spatial extension to the well known alpine regions. Another interesting aspect deals with patterns with one or zero "Non-Typical" (=Winner or Loser). They are characterizing more than 50 percent of all Swiss communities. This group of observed specific patterns is much bigger than expected. Subsets of communities or specific patterns are valuable for detailed investigations using other structural and temporal parameters (e.g. age of the population, infrastructure, buildings, etc.).

8 Conclusion

The presented community data mining approach [2] provides the ability to identify 880 patterns within a large amount of Swiss communities. Eight typical Swiss population dynamics are finally extracted and valuable for further qualitative and quantitative studies in terms of spatial planning aspects. The 8 Cluster solution is explained by the population dynamic of three periods: 1850 to 1910, 1910 to 1950 and 1950 to 2000. The author want to suggest that knowledge about spatial dynamics should be formulated in two stages: 1) Useful for planning processes, 2) Valid and reasonable in terms of statistical tests (significance). In the future spatial analysis and explanation should be intensified (e.g.Spatial Autocorrelation, Local Indicators of Spatial Association, Spatial Regression) to optimize planning concepts and costs.

Acknowledgements Population statistics are available on the website of the Federal Office of Statistics (BfS). The geometry of communities is also there to find in a repository of open geodata.

References

1. Bätzing, W. and Dickhörner, Y.: Die Bevölkerungsentwicklung im Alpenraum 1870-1990 aus der Sicht von Längsschnittanalysen aller Alpengemeinden. Revue de Geographie Alpine. **89**, 11–20 (2001)
2. Behnisch, Martin. 2009. *Urban Data Mining.* Karlsruhe: KIT Scientific Press.
3. Ultsch, A. (2008). Is log ratio a good value for measuring return in stock investments? In A. Fink, B. Lausen, W. Seidel and A. Ultsch (Ed.), *Avances in Data Analysis, Data Handling and Business Intelligence* (pp. 203–210). Berlin: Springer.
4. Ultsch, A. (2001). Eine Begründung der Pareto 80/20 Regel und Grenzwerte für die ABC Analyse. In: Technical Reports No.30. Department of Mathematics and Computer Science, University of Marburg. Available via DIALOG. http://www.mathematik.uni-marburg.de/databionik/pdf/pubs/2001/ultsch01begruendung Cited 15 Aug 2010
5. Perlik, M., Messerli, P. and Bätzing, W.: Towns in the Alps. Mountain Research and Development. **21(3)**, 243–252 (2001)
6. Tappeiner, U., Tappeiner, G. Hilber, A. and Mattanovich, E. 2003. *The EU Agricultural Policy and the Environment. (Ed. European Academy EURAC)* Berlin: Blackwell.
7. Oswald, F., and P. Baccini. 2003. *Netzstadt* Basel: Birkhäuser.
8. Ultsch, A. (2003). Pareto Density Estimation: A Density Estimation for Knowledge Discovery. In D. Baier, K.D. Wernecke (Ed.), *Innovations in Classification, Data Science, and Information Systems* (pp. 91–100). Berlin: Springer.
9. Schuler, M.; Ullmann, D.; Haug, W. 2002. *Eidgenössische Volkszählung 2000 - Bevölkerungsentwicklung der Gemeinden 1850-2000* Neuchtel:Bundesamt für Statistik.
10. Schuler, M., P. Dessemontet, and D. Joye 2005. *Eidgenössische Volkszählung 2000 - Die Raumgliederungen der Schweiz* Neuchtel:Bundesamt für Statistik.