Hierarchical self-organizing maps for clustering spatiotemporal data

Julian Hagenauer & Marco Helbich

Institute of Geography, University of Heidelberg, Heidelberg, Germany


To cite this article: Julian Hagenauer & Marco Helbich (2013): Hierarchical self-organizing maps for clustering spatiotemporal data, International Journal of Geographical Information Science, DOI:10.1080/13658816.2013.788249

To link to this article: http://dx.doi.org/10.1080/13658816.2013.788249

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Hierarchical self-organizing maps for clustering spatiotemporal data

Julian Hagenauer* and Marco Helbich

Institute of Geography, University of Heidelberg, Heidelberg, Germany

(Received 29 September 2012; final version received 17 March 2013)

Spatial sciences are confronted with increasing amounts of high-dimensional data. These data commonly exhibit spatial and temporal dimensions. To explore, extract, and generalize inherent patterns in large spatiotemporal data sets, clustering algorithms are indispensable. These clustering algorithms must account for the distinct special properties of space and time to outline meaningful clusters in such data sets. Therefore, this research develops a hierarchical method based on self-organizing maps. The hierarchical architecture permits independent modeling of spatial and temporal dependence. To exemplify the utility of the method, this research uses an artificial data set and a socio-economic data set of the Ostregion, Austria, from the years 1961 to 2001. The results for the artificial data set demonstrate that the proposed method produces meaningful clusters that cannot be achieved when disregarding differences in spatial and temporal dependence. The results for the socio-economic data set show that the proposed method is an effective and powerful tool for analyzing spatiotemporal patterns in a regional context.

Keywords: spatiotemporal data mining; self-organizing maps; dependence

1. Introduction

Technological progress and an emerging awareness about the importance of spatiotemporal data throughout spatial disciplines have resulted in an enhanced capability to collect data (Miller 2010). As a consequence, the availability of large heterogeneous databases for different spatial granularities and multiple temporal states has increased considerably (for census data, see, e.g., Martin 2006). Such databases typically contain hidden and unexpected information, which cannot be discovered using traditional statistical methods that require a priori hypothesis and cannot handle large amounts of data (Miller and Han 2009). This deficiency led to the emergence of spatiotemporal data mining, which is dedicated to the revelation of hidden information in spatiotemporal databases and its subsequent transformation into new and potentially useful knowledge. An important task of spatiotemporal data mining is spatiotemporal clustering, the aim of which is to organize a collection of patterns occurring in nearby areas at a similar time and with akin attributes. In this way, it reveals similarities at a higher level of abstraction and thus facilitates analysis and interpretation of the data.

Space and time are distinct dimensions with different properties (Peuquet 2002) that add substantial complexity to the mining process (Kechadi et al. 2009, Miller and Han

*Corresponding author. Email: hagenauer@uni-heidelberg.de

© 2013 Taylor & Francis
One fundamental property that attracted widespread attention in spatial sciences is spatial dependence (Tobler 1970, Goodchild 1986), which states that spatially nearby entities are likely to share more similarities than those that are far apart. The principle of spatial dependence is closely related to temporal dependence, although the latter is less complex (Miller and Han 2009). Temporal dependence refers to dependencies across time, meaning that the state of a process is related to its previous states (Chatfield 2003). Both spatial and temporal dependence hamper the use of conventional methods of statistical analysis, which typically assume independence between observations but require techniques that take spatial and temporal dependence into account (Andrienko et al. 2010a). Nevertheless, they are not merely confounding properties; they can also be a valuable source of information in spatiotemporal data mining (Miller and Han 2009).

Research on clustering has developed a broad set of algorithms and methodologies for assigning similar observations to clusters (see, e.g., Jain et al. 1999, Hand et al. 2001). However, none of these deal explicitly with spatiotemporal processes. An obvious approach to cluster spatiotemporal data is to use conventional clustering algorithms and to measure similarities between observations as a weighted mean of similarity in feature space and spatial and temporal proximity. However, the weighting of location and time assumes comparability between them, which cannot be expected because of their distinct nature (Openshaw 1999, Peuquet 2002). Consequently, the resulting clusters may not properly depict the inherent spatial or temporal patterns (Openshaw and Wymer 1995).

Several performance studies have demonstrated the advantages of the self-organizing map (SOM; Kohonen 1982, 2001) over other common clustering approaches (see, e.g., Ultsch and Vetter 1994, Openshaw and Wymer 1995, Watts and Worner 2009). The SOM is a flexible, unsupervised neural network for data analysis and clustering (Flexer 2001). It maps input data to neurons in such a way that the distance relationships between input signals are mostly preserved. Therefore, two-dimensional SOMs are especially valuable for visualizing high-dimensional data sets (Vesanto 1999); they are of great interest for GIScience (e.g., Openshaw 1992, Hewitson and Crane 2002, Skupin and Fabrikant 2003, Bação et al. 2005, Hagenauer et al. 2011).

A first approach to model spatial dependence based on SOMs is the GeoSOM (Bação et al. 2005). Conceptually, the GeoSOM is an adaption of the Kangas Map (KM; Kangas 1992), which was originally designed for temporal sequence processing, to spatial observations. The GeoSOM maps spatially close observations to map regions that are spatially close with respect to the distance on the map. Since the GeoSOM only takes spatial proximity into consideration, it neglects temporal dependence.

SOMs have also been previously used for spatiotemporal modeling. Skupin and Hagelman (2005) mapped trajectories of individual observations onto an SOM, which is useful for analyzing their temporal development. However, this approach neither provides a comprehensive spatiotemporal model of the data nor accounts for spatial or temporal dependence. Other approaches of spatiotemporal analysis utilize SOMs within an interactive and visual analytical framework (e.g., Guo et al. 2006, Andrienko et al. 2010b). Such frameworks are not strictly comparable to computational clustering approaches, since they rely heavily on the perception and analytical capabilities of humans to detect patterns.

Although, as discussed earlier, the limitations of current methods are obvious, a spatiotemporal clustering algorithm based on the principles of self-organization has been lacking thus far. Therefore, the objective of this research is to develop a hierarchical method that combines the concepts of the KM and GeoSOM for spatiotemporal clustering, taking both spatial and temporal dependence into account independently.
The rest of the article is organized as follows: Section 2 presents the basic principles of the SOM algorithm and two of its extensions, the KM and the GeoSOM. Section 3 introduces the method for clustering spatiotemporal data based on a hierarchical architecture of different SOMs. The efficacy of this method is tested in Section 4. Section 4.1 uses an artificial data set with known properties to evaluate the method’s performance and to demonstrate its advantages in comparison to a variant of the KM that regards time as an additional dimension. Section 4.2 applies the method to a socio-economic data set to show its usefulness for regional planning and analysis tasks in a real-world setting. Finally, Section 5 concludes the article with important remarks and identifies future research.

2. Background

2.1. Self-organizing maps

The SOM (Kohonen 1982, 2001) consists of an arbitrary number of neurons, which are connected to adjacent neurons by a neighborhood relation, defining the topology of the map. Although the dimension of the SOM’s topology is arbitrary, two-dimensional topologies are commonly used because of their ease of visualization. A prototype vector with the same dimension as the input space is associated with each neuron. During each training step, the Euclidean distances of all neurons’ prototype vectors to the training input vector are computed, and the neuron with the smallest distance, commonly termed the best matching unit (BMU), is selected. Then the prototype vector of the BMU and that of its surrounding neurons within a specified neighborhood on the map are moved toward the presented input vector. The magnitude of movement depends on the distance of the neurons to the BMU on the map and the learning rate. The neighborhood’s size and adaptation strength decrease monotonically during the training phase. Thus, the learning process gradually shifts from an initial rough phase, where the neurons are coarsely arranged on the map, to a fine-tuning phase, where only small changes to the map are permitted. As a result, the input pattern is iteratively learned by the SOM, where each neuron represents some portion of the input vector space. For a thorough discussion about the mathematical details of the SOM algorithm, see Kohonen (2001).

For visualizing SOMs, U-matrices (Ultsch and Siemon 1990) are convenient. The U-matrix plots the differences of neighboring neurons’ prototype vectors by means of a color scale. If the distance between neighboring neurons is small, these neurons then belong to a cluster with similar characteristics. If the distance between the neighboring neurons is large, the neurons do not share many characteristics and belong to different clusters. In the U-matrix, clusters become visible by distinct outlines of their cluster boundaries. Hence, the U-matrix shows both the present cluster structure and the quality of the clustering. If the actual class assignment of the input data is known, this information can be charted using pie charts (Mayer et al. 2007). The pie-chart map draws a pie chart for every neuron, which graphically represents the class proportions of observations mapped onto the neuron.

The basic SOM algorithm does not consider the temporal context of input patterns. However, there are a variety of SOM variants that deal with temporal data. A detailed review of temporal extensions is given by Guimarães et al. (2003). One extension that is particularly relevant for the subsequently developed method is the KM (Kangas 1992).

2.2. Kangas map and GeoSOM

The KM can be considered a specialization of the hypermap concept (Kohonen 2001). In the hypermap approach, the input vector is composed of two distinct parts: the context
J. Hagenauer and M. Helbich

and the pattern vectors. The context vector is used to select a subset of candidates. The BMU is then chosen from the established subset using the pattern vector. In general, the context vector is used exclusively for limiting the search space for the BMU, while the prototype vectors are adapted by using either the pattern vector or, alternatively, both parts (Guimarães et al. 2003). Conversely, in the KM, a subset of neurons surrounding the previously found BMU gives the context. The actual BMU is then chosen from this subset. Consequently, the mapping of the KM is determined by the temporal sequence of the input vectors presented to the map. Because temporally close observations are, in this way, naturally mapped to close neurons on the map, the KM takes into account temporal dependence within the input data.

The idea of the KM for incorporating temporal dependence into the mapping process was brought to spatial science by Bação et al. (2005). They applied the KM to spatial observations by considering spatial proximity instead of temporal closeness. Just like in KMs, the search for a BMU is made in two steps. First, the neuron with the prototype vector spatially closest to the input vector is determined. Second, the final BMU is searched within a fixed radius \( k \) around this neuron, commonly with respect to similarity of non-geographic attributes. This approach is termed GeoSOM. It can be used for determining spatially coherent and homogeneous regions. The radius \( k \) determines the strength of spatial dependence that is incorporated into the mapping process. Since geographic observations are typically spatially dependent (Goodchild 1986), the GeoSOM produces meaningful clusters, which provide interesting insights into spatial processes (Bação et al. 2005). For an elaborated technical discussion of these algorithms, see Hagenauer and Helbich (2012).

In general, if \( k \) is set to 0, input vectors are mapped solely to the neuron with the spatially closest prototype vector, neglecting the nonspatial components of the input vectors. If \( k \) is increased, the subset for the BMU search that is around the neuron with the spatially closest prototype vector increases, enabling the mapping of observations to more neurons. Because distances on the map relate to spatial distances, less dependence is incorporated. If \( k \) is chosen greater or equal to the maximum distance between neurons in the map, spatial locations are neglected because the complete map is considered for the final BMU search. However, the actual degree of dependence incorporated in the GeoSOM depends not solely on the radius, but also on the topology and size of the map (Hagenauer and Helbich 2012).

3. Method

The concept of the KM has been applied to account for temporal (Kangas 1992) and spatial dependence (Bação et al. 2005). The flexibility of the KM’s concept and its straightforwardness make the KM also promising for applying it to spatiotemporal clustering. However, several possible approaches exist for this purpose. A pragmatic approach is to modify the BMU search procedure of the KM so that it searches the final BMU within a radius \( k \) around the neuron whose prototype vectors are spatially and temporally closest. Thus, time is conceptually considered an additional spatial dimension. The single parameter \( k \) determines the degree of spatial and temporal dependence incorporated into the mapping process at the same time. In the following, this approach is termed spatiotemporal KM (STKM).

A more elaborated approach that is proposed here is to train two KMs in parallel, whereat each KM is trained independently for spatial and temporal dependence, respectively. The different properties of space and time can be taken into account by setting the parameters of the KMs accordingly, that is, the radii, dimensionality, and size of the maps. After the training, the two KMs independently model spatial and temporal patterns of the
input space. To get a comprehensive model of spatiotemporal patterns, it is necessary to merge the KMs subsequently. Because of their visualization capabilities for exploratory spatial data analysis, SOMs are particularly appealing for such a model (e.g., Takatsuka and Gahegan 2002, Guo et al. 2006, Andrienko et al. 2010b). In the following, this architecture, consisting of a spatial and temporal KM in the upper layer and a basic SOM in the lower layer, is termed hierarchical spatiotemporal SOM (HSTSOM). Figure 1 illustrates the architecture of HSTSOM.

To train a basic SOM in the lower layer with the results of the two KMs in the upper layer, it is necessary to build an appropriate training data set. Two conceptually different approaches exist for this purpose.

An obvious approach is to form training vectors by concatenating the prototype vectors of the KMs’ BMUs for each input vector of the data set. However, the basic SOM algorithm heavily depends on the calculation of Euclidean distances between input vectors and prototypes; it does not consider any spatial and temporal distance relationships explicitly. Consequently, training a basic SOM with this data set results in a confounded model, which essentially neglects the results of the KMs.

A better approach followed here is to form training vectors by concatenating the positions of the KMs’ BMUs for each input vector of the data set (see Figure 1). Because SOMs preserve distance relationships of the input data, the Euclidean distances between neurons of the KMs also reflect their spatial and temporal distances, respectively. Thus, training a basic SOM in the lower layer with this training data set results in a comprehensive spatiotemporal model, which takes the spatial and temporal distance relationships of the original data set into account and allows exploring spatiotemporal clusters.

4. Case studies

This section demonstrates the usefulness and advantages of the HSTSOM by applying it to two different case studies. For both case studies, the subsequent parameter settings of
the SOMs are used: the prototype vectors are randomly initialized and the training time is 100,000 iterations. The map topology is a two-dimensional hexagonal grid. The learning rate decreases linearly from 0.5. The neighborhood function is a Gaussian kernel with an initial neighborhood radius spanning the entire map (Skupin and Agarwal 2008).

4.1. Artificial data set

This case study uses an artificial data set whose spatiotemporal properties and settings are a priori known to compare the performance of the HST-SOM to the STKM and to outline its distinct features.

4.1.1. Data generation

The artificial data set adapts the idea of Bação et al. (2005) for an artificial spatial data set to a space–time cube, where time is considered an additional dimension. A set of 100,000 evenly spaced data points \((x, y, t)\) with \(x \in [0, 1]\), \(y \in [0, 1]\) is created, where \(x, y\) represent the spatial location of a point in time \(t\). Each point is also associated with a single component \(v\), which is 1 when \(x < 0.4\) and \(z < 0.48\), \(x > 0.6\) and \(z < 0.48\), \(x < 0.4\) and \(z > 0.52\), or \(x > 0.6\) and \(z > 0.52\), and 0 in all other cases. Since the spatial and temporal coordinates of the sample are randomly drawn from continuous intervals, each sample is likely to have a unique spatial location and time stamp. The resulting data set comprises four non-overlapping cuboids. The spatial distance between the cuboids at the same time is five times the temporal distance of the cuboids at the same locations. Figure 2 illustrates the cuboids, which define the data points of the data set. The different spacing of the cuboids in the temporal and spatial dimension reflects the different dependent properties of the input data set.

![Figure 2. Data cuboids. The \(x\) and \(y\) axes represent the spatial dimensions and \(t\) represents the temporal dimension of the input data space. The remaining space between the cuboids \((v = 0)\) is colored green (not shown).](image)
4.1.2. Spatiotemporal Kanga map

To cluster the data set with an STKM, it is necessary to determine a reasonable map size and radius for the STKM. For this purpose, 100 STKMs are trained with different settings, that is, map sizes 8 × 8 to 24 × 24 in steps of 4 and radii 0 to 35. For each resulting map, the U-matrix is segmented using the watershed algorithm of Vincent and Soille (1991). The algorithm regards the U-matrix as a topographic landscape flooded with water. Basins represent clusters, whereas watersheds mark cluster borders. The clustering results are then evaluated with regard to the actual class distributions by calculating their mean normalized mutual information (NMI). NMI is a symmetric measure that assesses the similarity between two different clustering results by considering their mutual dependency (Strehl and Ghosh 2003). If two clustering results are identical, the NMI takes the value of 1. Accordingly, if two clustering results have no information in common and are thus totally independent, the NMI takes the value of 0.

The results are shown in Figure 3. Only radii 0 to 15 are shown because from radius 15 onward, the mean NMIs are almost constant. For most of the radii, the mean NMI is moderate and varies according to map size and radii. The mean NMIs range from 0.457 to 0.618, whereas an initial increase is noticeable between \( k = 0 \) and \( k = 1 \). The best mean NMI was obtained for a KM consisting of 24 × 24 neurons with a radius of 1.

Figure 4 shows the U-matrix and pie charts for an exemplary STKM of size 24 × 24 and with a radius of 1. The U-matrix reveals a complex structure of the map, even though clear-cut separations between different map regions are apparent. However, the U-matrix shows more clusters than are actually present in the input data. Furthermore, the spatially distant clusters in the input data (red and blue, yellow and cyan) are wrongly assigned to different clusters in the map, whereas temporally close clusters in the input data (yellow and red, cyan and blue) are partially mapped to the same clusters in the map. Thus, it is evident that the STKM failed to clearly outline the clusters in the input data.

![Figure 3. Mean NMI versus radius for different-sized STKMs.](image-url)
4.1.3. Hierarchical spatiotemporal self-organizing map

The HSTSOM uses two upper layer KMs: one for the temporal (temporal KM) and one for the spatial aspects of the data (spatial KM). To alleviate the clustering of the lower-layer map, only one-dimensional KMs are deployed in the upper layer of the HSTSOM. Thus, the dimension of the input space is reduced. The results of the HSTSOM depend on the deployed upper-layer maps. To find reasonable parameter settings for both the spatial and temporal KM, each one is trained 100 times for different map sizes and radii, that is, for map sizes $8 \times 1$ to $24 \times 1$ in steps of 4 and radii 0 to 23. As outlined earlier, for each setting, the trained KMs are clustered by computing the watersheds of the U-matrices and, subsequently, the mean NMI.

The results of the spatial KM are shown in the left panel of Figure 5, whereas the results of the temporal KM are shown in the right panel. From radius 14 onward, the mean NMIs are constant for all map sizes. Therefore, only radii 0 to 14 are shown.

![Figure 5. Mean NMI versus radius for spatial (left) and temporal KMs (right) of different sizes.](image-url)
The radius for which a maximum mean NMI is achieved increases with map size, illustrating the dependence between map size and radius. On average, the spatial KM achieves the best NMI with radius 7 and map size $16 \times 1$, whereas the temporal KM achieves the best NMI with a map size of $16 \times 1$ and a radius of 4. Figure 6 shows the U-matrices and pie charts of an exemplary spatial and temporal KM trained with these parameters. It can be seen that both KMs succeed in separating the temporal and spatial clusters, respectively.

The spatial and temporal KMs constitute the upper layer of the HSTSOM. The lower-layer map of the HSTSOM combines the results of the upper-layer KMs into a comprehensive representation. Since the input data has only five inherent clusters, a lower-layer map consisting of five neurons is sufficient. However, if no information about present clusters in the input data is available in the input data, it is reasonable to use larger maps. Figure 7, for example, depicts a $10 \times 10$ SOM, trained with the prototype vector positions of the two KMs. The five clusters (blue regions) of the input data are clearly separated. The NMI for the HSTSOM is 1; hence, the HSTSOM has separated the spatiotemporal clusters perfectly.

4.2. Spatiotemporal clustering of the Ostregion, Austria

The exploration of spatiotemporal clusters is substantial for spatial planning to formulate policies and decision-making. It is the aim of this case study to identify such clusters in a socio-economic context for the Austrian Ostregion. However, the validation of the identified clusters is difficult because there exist no correct solution to the problem in a formal sense. Therefore, this case study evaluates the clustering results by linking them to existing geographic knowledge. Note that a comprehensive in-depth analysis of these regions is beyond the scope of this article.
4.2.1. Study site and data set

The planning region Ostregion is a joint initiative of the federal states of Vienna, Lower Austria, and Burgenland. It consists of 744 municipalities and 23 Viennese districts. The Ostregion is characterized by a highly dynamic urban center, the city of Vienna and its urban fringe, and an outer area. It receives large attention in the planning context because of its location within the prospering trans-national Central European Region. Austrian census data of the Ostregion for this case study is obtained from Kranabether et al. (2012). It consists of four variables that have been identified by Kranabether et al. (2012) as being important for analyzing spatial planning aspects over time: (a) population per km² (POP), (b) non-agricultural work places per km² (NAWP), (c) percentages of commuters (COM), and (d) percentages of employees (EE). These variables are present for five equidistant time stamps, spanning the years 1961 to 2001. The location of each municipality is represented by its centroid coordinates. Table 1 lists the descriptive statistics of the data set exemplarily for the year 2001. To make meaningful and comparable distance calculations, the attributes POP and NAWP are logarithmically transformed to normalize their skewed distributions. Furthermore, all attributes are scaled to be in the range of 0 to 1. The aspect ratio of the spatial coordinates is maintained.

4.2.2. Hierarchical spatiotemporal self-organizing map

The HSTSOM consists of a one-dimensional temporal KM and a two-dimensional spatial KM. Thus, the dimensionalities of the KMs match the ones of time and space. Moreover, it is necessary to determine appropriate map sizes and radii. The sizes of the maps involved must be chosen to achieve a compromise between mapping accuracy, generality, and computational demands. Based on preliminary experiments and empirical considerations, the size of the temporal KM is set to 12 × 1 neurons and the spatial KM is set to 14 × 12 neurons. The rectangular shape of the spatial KM roughly matches the spatial distribution of the input data, which can stabilize the map during the training process (Kohonen 2001). The lower-layer SOM consists of nine neurons, arranged in a 3 × 3 grid, each representing a distinct cluster. Preliminary analysis confirmed that nine clusters are a fair compromise between generality and precision.

To find adequate radius settings, 100 KMs are trained with different radii, that is, 0 to 11 for the temporal KM and 0 to 18, and the resulting quantization errors (QEs) are averaged. The QE is calculated by determining the average distance between each observation and the prototype vector of its BMU (see Pöllzbauer 2004). To analyze the effect of different radii for the different aspects on the data, the QE is further divided into demographic, spatial, and temporal QE. The demographic QE (DQE) solely considers the distance of socio-economic attribute values. The temporal QE (TQE) investigates the time

Table 1. Descriptive statistics of the data set for the year 2001.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>POP</td>
<td>4.342</td>
<td>40.650</td>
<td>64.480</td>
<td>391.700</td>
<td>121.484</td>
<td>24,687.565</td>
</tr>
<tr>
<td>NAWP</td>
<td>0.080</td>
<td>1.301</td>
<td>2.187</td>
<td>29.940</td>
<td>4.828</td>
<td>2,897.000</td>
</tr>
<tr>
<td>COM</td>
<td>0.262</td>
<td>0.659</td>
<td>0.740</td>
<td>0.716</td>
<td>0.796</td>
<td>0.915</td>
</tr>
<tr>
<td>EE</td>
<td>0.359</td>
<td>0.432</td>
<td>0.450</td>
<td>0.449</td>
<td>0.467</td>
<td>0.601</td>
</tr>
</tbody>
</table>
periods between observations, whereas the spatial QE (SQE) is solely focused on the spatial distances. The different QEs’ means of the spatial and temporal KM with different radii are plotted in Figure 8. Both panels show that the DQE decreases with an increasing radius, whereas the TQE of the temporal KM and the SQE of the spatial KM increases. The QE of the temporal KM does not change notably with radius, indicating that the formation of temporal coherent clusters does not significantly affect their spatial configuration. In contrast, the TQE of the spatial KM changes notably with radius size.

The SQE and DQE of the spatial KM, and the TQE and DQE of the temporal KM, are almost equally low at a radius of 3. Therefore, a HSTSOM consisting of a temporal KM with a radius of 3 and a spatial KM with a radius of 3 seems most appropriate and is used in the following.

Every neuron of the HSTSOM’s lower-layer map represents a portion of the input space. These portions are ultimately the clusters formed by the HSTSOM. The quality of this mapping can be evaluated by measuring the QE of these clusters with respect to their centroids. Table 2 compares the DQE, TQE, and SQE of the HSTSOM’s clustering results to the ones of a STKM of size 3 × 3 and with radii from 0 to 3. The STKM with radius 3 equals a basic SOM that disregards the time stamps and locations of the observations in the process of mapping, since the radius equals the maximum distance between neurons in the map. The table reveals that the TQE of the STKM with radius 3 is considerably lower than the SQE, indicating that the similarity of the demographic attributes relates more to temporal closeness than to spatial proximity. This supports the results of the sensitivity analysis (see Figure 8). Even though the radii of the spatial and temporal KM are equally set to achieve a sound compromise between the QEs, the spatial KM incorporates a higher

![Figure 8. Mean spatial, temporal, and demographic QE versus radius for a spatial 14 × 12 KM (left) and a temporal 12 × 1 KM (right).](image)

<table>
<thead>
<tr>
<th>HSTSOM3,3</th>
<th>STKM0</th>
<th>STKM1</th>
<th>STKM2</th>
<th>STKM3</th>
</tr>
</thead>
<tbody>
<tr>
<td>SQE</td>
<td>0.156</td>
<td>0.082</td>
<td>0.196</td>
<td>0.236</td>
</tr>
<tr>
<td>TQE</td>
<td>0.165</td>
<td>0.301</td>
<td>0.218</td>
<td>0.204</td>
</tr>
<tr>
<td>DQE</td>
<td>0.206</td>
<td>0.262</td>
<td>0.154</td>
<td>0.134</td>
</tr>
<tr>
<td>Mean</td>
<td>0.176</td>
<td>0.215</td>
<td>0.189</td>
<td>0.191</td>
</tr>
<tr>
<td>SD</td>
<td>0.027</td>
<td>0.117</td>
<td>0.032</td>
<td>0.052</td>
</tr>
</tbody>
</table>

Note: The subscripts represent the applied radii.
degree of dependence because it consists of significantly more neurons (see Hagenauer and Helbich 2012). Moreover, the HSTSOM enables a more balanced clustering and also lower mean QEs than the STKM for all reasonable radii.

To analyze the HSTSOM’s clusters in more detail, it is necessary to visualize the temporal, spatial, and demographic characteristics of the clusters. Figure 9 shows box plots of the demographic attributes for each cluster, while Figure 10 shows a geographic map of the clusters and the histograms of their time stamps. The clusters in the geographic map, the histograms, and the box plots are arranged in accordance with the positions of the neurons which map these clusters.

The analysis of the clustering results allows drawing conclusions about the study region’s socio-economic characteristics and their temporal development. For instance, cluster 5 in Figure 10 comprises the whole northern area of the study region in the years 1961 and 1971. It is characterized by a moderate employment rate and a very low rate of commuters (see Figure 9). Cluster 5 does not persist in this form for more recent years. In fact, in the years 1981 to 2001, the western regions of cluster 5 are present in

![Box plots of the clusters' demographic attributes. Observations outside 3 × the interquartile range are marked as outliers.](image)

Figure 9.
Figure 10. Spatial distributions of the clusters and the histograms of their time stamps. Black lines represent district borders, while municipality borders are not shown.

Clusters 4 and 6, whereas the eastern regions are primarily located in clusters 7 and 8 (see Figure 10). This splitting can be attributed to a diverging socio-economic development of both regions. In particular, comparing the box plots of the clusters 4 and 8 (Figure 9)
reveals that the rate of commuters is slightly higher for the eastern region than for the western region in the middle of the study period. This difference increases at the end of the study period (see box plots of clusters 6 and 7 in Figure 9).

Furthermore, cluster 2 is clearly distinct to cluster 5, both geographically and also with regard to its attributes. However, this distinction is not clear for more recent years. In particular, the western regions of both clusters are part of the recent cluster 6 (Figure 10), indicating that both regions have undergone a development that has led to similar socio-economic characteristics.

These analytical findings can be explained by the urbanization process taking place in Vienna and its close surroundings. In particular, in recent years, the socio-economic characteristics of the study area seem to be more and more determined by the evolving urbanization and prospering economy of Vienna. Thus, regions close to Vienna converged to typical urban characteristics like high rates of non-agricultural working places and high population (e.g., cluster 7), while regions far apart feature more rural characteristics like lower employment rates and low population (e.g., cluster 6).

5. Discussion and Conclusions

This research presented the HSTSOM, a method for analyzing spatiotemporal data based on SOMs. The HSTSOM considers the distinct nature of space and time by representing both aspects independently by means of individual KMs in the upper layer of its hierarchical structure. The KMs permit the independent adjustment of spatial and temporal coherence and can be individually charted to investigate spatial and temporal patterns separately. Finally, the lower-layer SOM combines both KMs and thus provides a comprehensive representation of the input data, which facilitates the exploration of spatiotemporal clusters. The usefulness of this approach was demonstrated on an artificial data set and in an empirical case study.

The dimensionality of the upper-layer maps of the HSTSOM is principally arbitrary. A distinct advantage of using maps with different dimensionalities is the ability to take into account the different dimensionalities of space and time. However, the choice of the upper-layer maps’ dimensions has severe consequences for the HSTSOM. Low-dimensional maps reduce the input space of the lower-layer map and thus facilitate the final mapping. However, they are hardly able to preserve the topology of high-dimensional input spaces. Furthermore, the visualization capabilities of maps with dimensions other than two are severely limited.

Baçao et al. (2005) draw the positions of neurons’ prototype vectors directly onto a geographic map to illustrate the effect of the radius on the prototype vectors’ spatial distribution. For the HSTSOM, the prototype vectors of the lower-layer map have no spatial or temporal components. Thus, it is not possible to directly map the neuron of the lower-layer map in a spatial or temporal context. However, the neurons of the lower-layer map represent subsets of the upper-layer maps’ neurons, which represent partitions of the input space that can, as demonstrated in Section 4.2, easily be depicted on geographic maps. Moreover, the neurons’ prototype vectors of the temporal and spatial KM can be depicted on a geographic map in a straightforward manner.

The mapping of KMs depends on spatial or temporal contexts. Hence, post-processing of a KM using algorithms and visualizations, which are solely based on the distance relationships between prototype vectors, for example, conventional clustering algorithms (Vesanto and Alhoniemi 2000), is inappropriate. This shortcoming does not apply to the HSTSOM because it employs a basic SOM at the lower layer. Therefore, it can be used for
further analysis and processing by means of basic methods, that is, conventional clustering algorithms, which generally do not consider contextual information.

The HSTSOM makes two simplifying assumptions: first, since the algorithm uses two different KMs in the upper layer to model spatial and temporal dependence, it assumes separability of spatiotemporal dependence. This corresponds to the common geostatistical approach to decompose a covariance model into a product of a purely spatial and purely temporal component (see Kyriakidis and Journel 1999, Heuvelink and Griffith 2010). In physically justifiable situations, separable covariance models allow computationally efficient estimation and inference (Mateu et al. 2008). Accordingly, the separate models for spatial and temporal dependence improve the computational efficiency of the HSTSOM: first, the KMs of the upper layer can be trained in parallel, and the dimensionality of the input space for the lower-layer map is reduced to the sum of the upper-layer’s dimensionalities, facilitating the training of the lower-layer map. Second, the HSTSOM assumes stationarity of spatial and temporal dependence. This is in accordance with the GeoSOM and KM. Currently, there exists no computationally efficient spatial or temporal SOM algorithm that incorporates nonstationary of dependence.

In general, the basic SOM is a parameter-sensitive algorithm. The selection of suitable parameters is subjective and guided through empirical analysis. This pragmatic approach is supported by a lack of a profound mathematical analysis of the SOM’s properties (Cottrell et al. 1998). In practice, the results of the SOM are robust for a wide range of reasonable parameter settings (see, e.g., Smith and Ng 2003). However, the problem of finding adequate parameters is more difficult for hierarchical SOMs (e.g., Lampinen and Oja 1992, Dittenbach et al. 2002, Hung and Tsai 2008). Therefore, future research aims to achieve a more robust parametrization of the HSTSOM (i.e., choice of the radii and map sizes).

Note
1. Concatenation refers here to the joining of two vectors to a single vector end-to-end. For example, concatenating the vectors \((1, 2)\) and \((3, 4)\) results in the vector \((1, 2, 3, 4)\).

References


