Understanding Spatio-temporal Patterns: Visual Ordering of Space and Time

Menno-Jan Kraak & Daniël E. van de Vlag

ITC - International Institute of Geo-Information Science and Earth Observation
Department of Geo-Information Processing
PO Box 6, 7500 AA Enschede the Netherlands
kraak@itc.nl; vandevlag@itc.nl
Phone : +31 (0)53 4874463
Fax : +31 (0)53 4874335

Abstract

Deriving patterns and relations from large multivariate and multi-temporal datasets to acquire knowledge about real world processes is not trivial. To understand the content of such datasets current analytical tools do offer interesting solutions, but an integrated approach is lacking. Here we introduce a visual integrated solution that allows one to explore and analyze the data at hand. The approach introduced consists of a dynamically linked multi view environment that offers different interactive visual representations to 'look at and play with' the data. For the time component the temporal ordered space matrix (TOSM), which schematizes the temporal nature of the data set, is introduced. The rows of the matrix represent time and the columns the geographic units.

A preliminary usability test has been conducted to see how the multi-view approach in general performs when considering specific tasks oriented towards the understanding of spatio-temporal patterns. The TOSM functions well for naturally (linear) ordered phenomena such as rivers and coastlines. The paper also discusses the use of the TOSM for non linear ordered phenomena such as administrative units. The method is based on directional ordering and is compared with other ordering approaches, such as space filling curves, traveling salesman problem and plane sweeping algorithms.

Understanding Spatio-temporal Patterns: Visual Ordering of Space and Time

Menno-Jan Kraak & Daniël E. van de Vlag

Abstract

Deriving patterns and relations from large multivariate and multi-temporal datasets to acquire knowledge about real world processes is not trivial. To understand the content of such datasets current analytical tools do offer interesting solutions, but an integrated approach is lacking. Here we introduce a visual integrated solution that allows one to explore and analyze the data at hand. The approach introduced consists of a dynamically linked multi view environment that offers different interactive visual representations to 'look at and play with' the data. For the time component the temporal ordered space matrix (TOSM), which schematizes the temporal nature of the data set, is introduced. The rows of the matrix represent time and the columns the geographic units.

A preliminary usability test has been conducted to see how the multi-view approach in general performs when considering specific tasks oriented towards the understanding of spatio-temporal patterns. The TOSM functions well for naturally (linear) ordered phenomena such as rivers and coastlines. The paper also discusses the use of the TOSM for non linear ordered phenomena such as administrative units. The method is based on directional ordering and is compared with other ordering approaches, such as space filling curves, traveling salesman problem and plane sweeping algorithms.

1. Introduction

Monitoring of real world processes results in large multivariate and multi-temporal data sets. To be able to understand these complex datasets, one need options to explore these data sets to study trends, correlations and patterns, which should lead to knowledge discovery about the dynamic behaviour of geospatial phenomena. For analysts, it is a huge challenge to effectively and efficiently detect and understand interesting relationships, patterns, correlations and trends in such data sets (Guo et al. 2005).Visualization techniques are powerful tools to detect, analyze and study these changes. In particular, the field of geovisualization shows promising results, because geovisualization techniques offer the ability to dynamically analyze and interact with the data.

Representation of objects in space and time has attracted much attention of researchers during the last decade (among them (Andrienko and Andrienko 1999; Edsall 2005; Gahegan 1999; MacEachren et al. 1999; Peuquet 2002; Worboys 2005). Peuquet (Peuquet 1994) described how spatio-temporal data are characterized by three basic components: where (space), what (objects) and when (time). Based on the above basic components, (Andrienko, Andrienko, and Gatalsky 2003) classify spatio-temporal data according to the kind of changes occurring over time. They recognize existential changes, as appearance and disappearance, which can be related to the "when" component. Besides, they recognize changes in spatial properties, as location, shape and/or size, orientation, altitude, etc., and this can be related to the "where" component. Last, they identify changes of thematic properties that are expressed through values of the attributes and this can only be partly described by the "what" component. In fact, the "what" component not only indicates the object by itself, but also includes many important characteristics of the object, and these characteristics are described by several attributes. Depicting the temporal component of data simultaneously with its spatial information, in an integrated manner, would be a powerful method to support real data analysis tasks.

Here, an integrated method is suggested to visualize multivariate spatio-temporal data sets. The solution proposed is a prototype of three dynamically linked views, each representing one of spatial data's components. These are the interactive (animated) map view ("where" component), the parallel coordinate plot view ("what" component), and the temporal ordered space matrix view ("when" component). The Temporal Ordered Space Matrix (TOSM) is a new visualization technique and can be considered a grid 'map' in time space, whereby the rows in the matrix represent time, the columns represent geographic units, and the individual cell can be colour-coded according the value of user defined attributes. A direct link exists between the geographic units in the (animated) map view and the time view. The main advantage of the TOSM is that it offers a direct overview of spatial and temporal trends in an ordered space, while preserving (most) spatial relations. Therefore, geographic units are ordered according a linear principle using a directional ordering method. This approach is especially suitable for geographic features such as coastlines or rivers which have a natural 'linear-orientation'. Since the prototype of the TOSM proved to be useful it is a challenge to see if the approach can be extended to other geographic features as well. For this reason different ordering principles will be discussed.

2. Visualization of multivariate spatio-temporal datasets

Multivariate spatio-temporal datasets in a geographic context occur when geoscientists deal with geospatial data where each geographic space contains its own data array representing location, attribute and temporal dimensions. To provide greater insight in the multivariate and spatio-temporal characteristics of the dataset, there is a need for visualization technique(s). This would involve integration of spatio-temporal and multivariate visualization tools in a single visualization environment, with best of both worlds. Linking allows interaction between the views and can facilitate exploratory tasks. Recently, many examples of linking maps with graphic tools, such as dot plots, parallel coordinate plots, and scatter plots, can be found (Andrienko and Andrienko 1999; Anselin, Syabri, and Smirnov 2002; Buja, Cook, and F. 1996; Dykes 1997; Edsall, MacEachren, and Pickle 2001; MacEachren et al. 1999). Integration of different visual methods takes advantage of their inherent capabilities for visualizing spatial distribution of changing phenomena, as well as interrelationships with its change contributing factors, gaining greater insight to information from multivariate spatio-temporal data. As described earlier by DiBiase (DiBiase 1990) and recently by the Geovisualization community (Dykes, MacEachren, and Kraak 2005), visual methods can perform a range of functions in the visualization process, including exploration to reveal pertinent questions, confirmation of apparent relationships in the data, synthesizing or generalizing findings and presentation of the findings. The visualization process usually begins with exploration and confirmation of relationships in data from the private realm where specialists are engaged in research to public, where the research findings are released to the public for communicating the results. Hence, for implementation of the visualization prototype and its interface, it is essential that besides visualization of obvious relationships in multivariate spatio-temporal data, the prototype should also be able to explore non-obvious relationships.

The above mentioned multivariate spatio-temporal visualization concepts and products often represent multivariate data and spatio-temporal data separately. To bring those together a novel visualization technique is proposed, whereby spatial and temporal properties of spatio-temporal data sets stay preserved (Vlag and Kraak 2005, 2004). The Temporal Ordered Space Matrix (TOSM) representing time space plays a central role in these three view environment. In the matrix the locations of objects are ordered as linear elements and are plotted against the temporal variations of an attribute. A matrix of squares is constructed, where each column depicts a geographical unit, and each row represents time and each cell the value attribute. The TOSM is in particular effective for naturally ordered spatial datasets, such as coast and river compartments, traffic data, water level data, etc. For such datasets, as shown in Figure 1, a clear directional spatial ordering between the geographical units exists. For instance for flooding data, measuring point A is upstream of point B, and point C is downstream of point B, the directional spatial ordering from upstream to downstream will be A, B, C. The spatial ordering is inherited from the location of the measuring points, i.e. upstream and downstream. In the TOSM environment, the position of the geographical unit (i.e. the measuring points) can also be drawn using a linear line. Figure 1 is based on a flooding situation in 1995. Extreme rainfall in Germany, Belgium and the Netherlands caused flooding of the Maas and Rhine River, resulting in evacuation of 250.000 people. The case study involves 16 daily water height measurement points from 25 January to 11 February 1995 along the Maas River in The Netherlands. Water heights of the flooding are standardized to mean water levels. These values are normalized from 0 to 1 for optimal visualization. The Maas River flooding is a single-variate naturally ordered spatio-temporal application. Hence, there is a need for a temporal visualization tool to observe the evolution and movement of the peak flow in time. The TOSM can help for understanding these temporal trends. Visualizing this application in the TOSM, a clear initial orientation can be noticed based on the flow direction of the water.

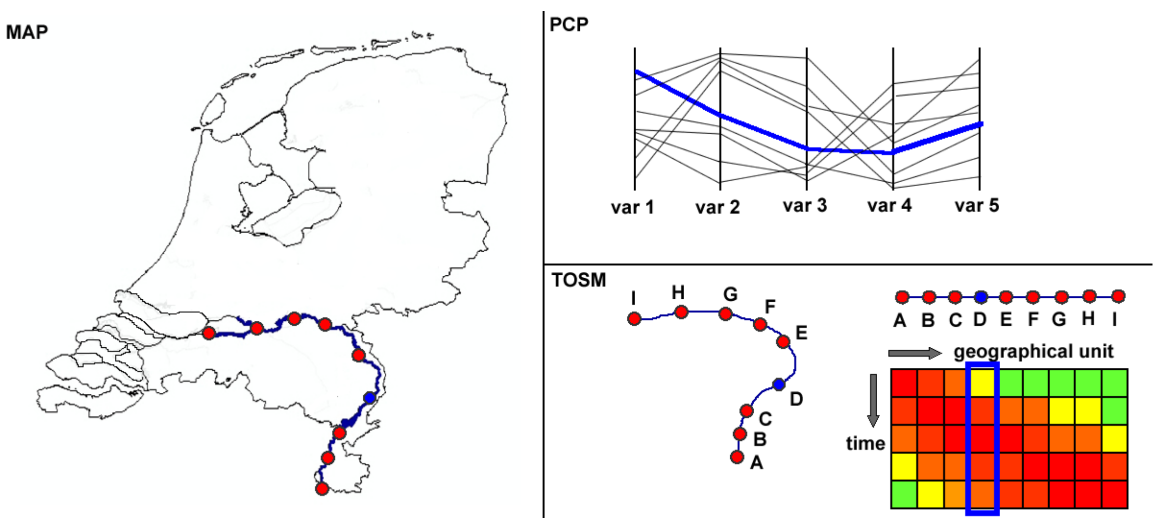


Figure 1. Concept of the TOSM environment for depicting multivariate spatio-temporal data, using a flooding application as illustration. Links exist between the (animated) map, the PCP and the TOSM. Selecting time in the TOSM automatically select corresponding elements on map and PCP.

3. The TOSM at work

The principles of the TOSM and its multiple view environment have applied in a beach management application in the northern part of the Netherlands. The application as been extensively described by (Van de Vlag et al. 2005). The prototype (see Figure 2) has been used to visualize trends and associations between beach compartments, changes in time and quality elements involved in the decision making.. As such it contributed to a better understand of these relationships and trends by coastal management engineers. The prototype for visualization and exploration of multivariate spatio-temporal datasets is developed in JAVA™ and implemented in UDIG (User-friendly Desktop Internet GIS), version 1.0.5, developed by Refraction Research Inc (UDIG 2004). It consist of three environments: a map environment (MAP) - already available in UDIG – and two new developed tools, a parallel coordinate plot (PCP) and the temporal ordered space matrix (TOSM).

The MAP environment displays the area map giving answer to the “where” component. By ticking boxes in the layer menu, the user can decide which map to draw in the MAP environment. Tools for pan, zoom, info, edit and selection are all available in the UDIG tools menu. Value attributes in the MAP environment can only be shown using the info button and selecting a particular object.

The PCP is a geometrically transformed visualization technique, suitable for representing multivariate data. Here, the “what” component is illustrated by the object and its attributes. The prototype is able to generate a PCP using the attribute table of the dataset, showing the attributes as vertical axes of the PCP and each object as a connected line between attribute values. Using a pull-down menu, a user can select the data layer to be seen in the PCP. Besides, a user can select the attribute values to be seen in the PCP by ticking the accompanying boxes. On the bottom of the attribute axis the minimum and maximum values are represented. On mouse-overs near the PCP axes, actual attribute values of objects are shown.

The TOSM is a representation of a matrix whereby ordered space is portrayed against time, illustrating the “when” component. It illustrates schematically location, time and theme of spatio-temporal objects. Here, the user has first to define the initial orientation of the dataset, by drawing a directional line. In the prototype, only the linear directional line has been implemented. Next, the user has to select the lookup attribute (i.e. the object ID) used to link the object to the spatial reordered location, and the attribute to display the theme of interest using pull-down menus. Subsequently, the TOSM is automatically generated. A colour table is added, calculated from the minimum and maximum values of the attribute of interest ranging from green (minimum) to yellow (intermediate) to red (maximum). On mouse-overs on each cell, the object ID number as well as its attribute value is shown.

A quantitative usability test has applied to judge the effectiveness of the prototype in improving understanding and exploration of trends, relationships and patterns of multivariate spatio-temporal data sets. The testing methods involved assessing the tool’s ability to meet user’s performance (effectiveness and efficiency) and satisfaction objectives looking at the prototype’s MAP, PCP and TOSM view separately and combined,. The assessment was conducted based on a number of representative user tasks for which a certain number of usability measures are measured. The kind of tasks given where: A summary of tasks participant has to execute during the test is given below:

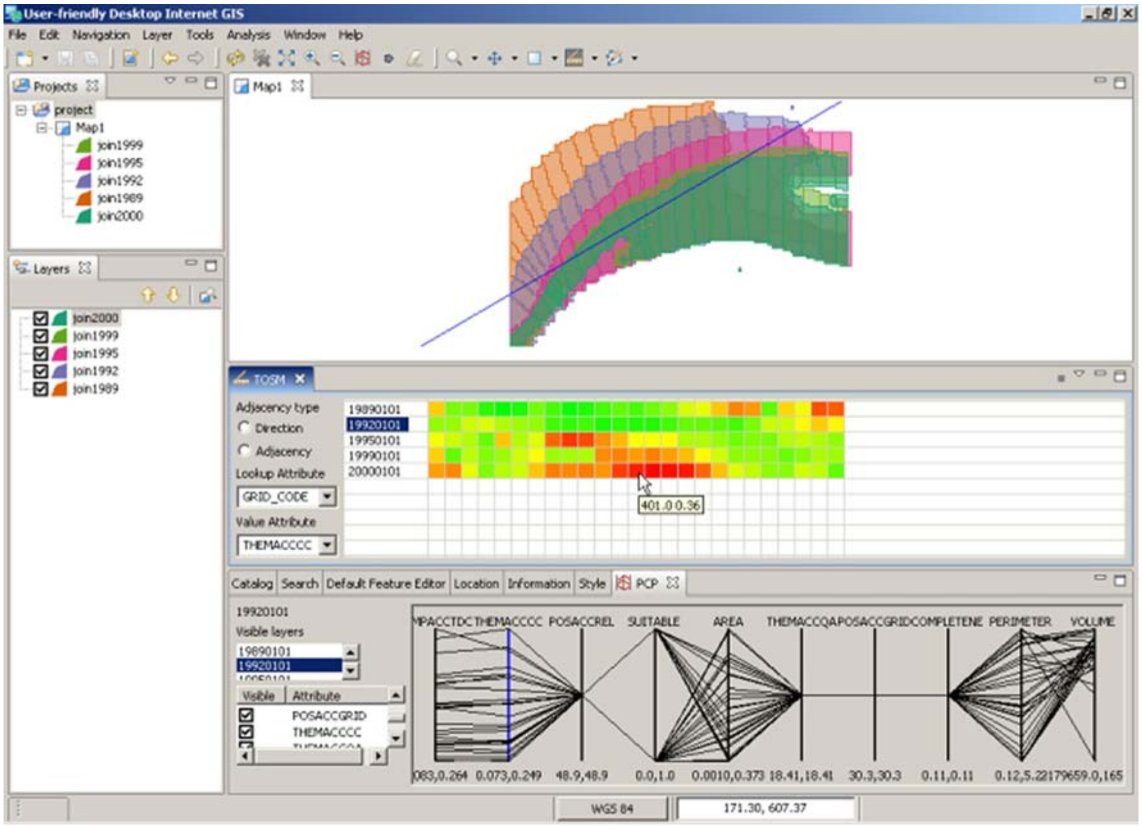


Figure 2. Screenshot of the prototype with beach management application data. Top view depicts the animated map environment; the middle view shows the TOSM; the bottom view illustrates the PCP.

Multivariate relationship	*Can you describe and explain the relation between variable X, variable Y and variable Z?*
Temporal trend	*Can you describe and explain the trend of variable X in time?*
Spatio-temporal pattern	*Can you describe and explain the pattern between location and X in time?*
Spatio-temporal pattern	*Can you describe and explain the pattern between location and Y in time?*
Spatio-temporal pattern	*Can you describe and explain the pattern between location and Z in time?*

Results of the usability testing can be found in (Van de Vlag 2006), and show that the prototype performed well for the natural ordered geographic phenomena.

4. Reordering of spatial data

To depict time and space in the TOSM space is collapsed into one-dimension. This results in a line along which the geographic units are ordered. With natural phenomena such as coastlines or rivers this does not give much trouble and a directional linearization approach can be applied, whereby the orientation can be inherited from directional relations of the geographical units. The approach results in relevant spatio-temporal patterns in the TOSM. However could this method also be applied to other geographic data such as for instance census data based on administrative units? How can this type of data be linearized? Methods to extract dimensions are strictly related to the spatial neighbourhood of the geographical units. After all, the explicit location and extension of geographical units define implicit relations of its spatial neighbourhood. These spatial relations have been classified in several types (Frank 1996; Papadias and Sellis 1994)), including directional relations that describe order in space (Frank 1996; Freksa 1992), distance relations that describe proximity in space (Hernandez, Clementini, and Di Felice 1995)and topological relations that describe neighbourhood and incidence (Egenhofer 1994). When datasets are indeed randomly ordered other (non-linear) methods may be more useful. Here, the directional linearization approach is discussed first, followed by other linearization approaches.

Spatial ordering by directional linearization

For the TOSM, reordering of geographic objects – while preserving most of its spatial relations – is the strength of this visualization technique. Usually, to refer to geospatial locations, people use qualitative describers based on their spatial perception and to denote relative directions among geographical units, such as East, West, North East and South West. But this kind of directional reference is imprecise and makes it difficult to identify these objects. There are different spatial models for handling directional terms. A simple one is to divide an area into eight directions based on an angular division of an area into eight equal sectors (N, NE, E, SE, S, SW, W, NW). More precise directions can be derived using compass directions in degrees or sketching the initial orientation of the objects.

In case of the TOSM, the centroid of a geographical unit is chosen to locate the geographic object, as shown by an example of reordering 46 municipalities in Overijssel, The Netherlands (see figure 3). For complex (irregular) thematic data like most census data, spatial reordering of the objects may not perform effectively. The linearization is executed by sketching a directional line with the initial orientation. The initial orientation can be deduced from the structure of the data set, e.g. the SE-NW orientation in map window of figure 2, or it can be derived from the task at hand, e.g. if you want to explore the trend from feature X from North to South. Also, the user can freely rotate the line in each desirable direction, which may lead to new discoveries in pattern and trend recognition. In the final stage, the centroids of the objects are perpendicular projected along the directional line and depicts the order of the objects in the horizontal matrix cells (figure 3).

Obviously, the selection of initial direction is critical for the spatial behaviour of the TOSM. After all, when one draws the directional line not from West to East (figure 3a), but from North to South (figure 3c), geographical units are reordered differently and various other spatial relations and patterns are or may be illuminated. This however can be related to the task at hand e.g. if one is interested in finding spatial or spatio-temporal patterns from West to East, North to South, etc.

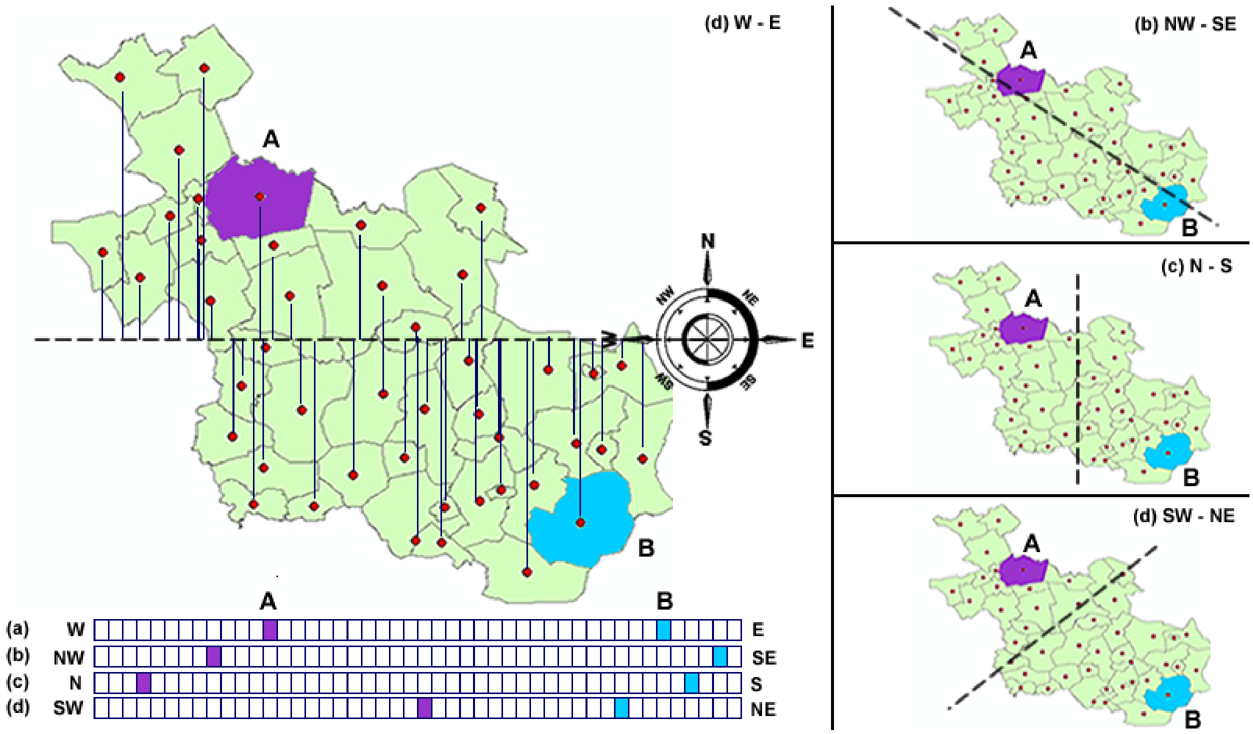
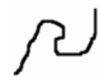
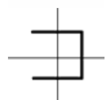
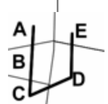
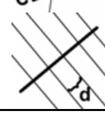


Figure 3 Implementation of initial orientation using a directional line for reordering municipalities in Overijssel, The Netherlands. The red dots are the centroids of the individual geographic objects. For two municipalities the results for four different directional lines are illustrated: (a) West-East, (b) Northwest – Southeast, (c) North – South, (d) Southwest – Northeast.

Alternative linearization methods

Linearity between administrative units is unlikely to exist. Accordingly, alternative ordering for the linearization of the dataset have to be found to reduce a 2D location on the map into a 1D location in the TOSM. Some alternatives are discussed and illustrated in table 1. A method most closely related to the directional line is freeform drawing of the line, for instance based on particular phenomena of the geography. Other alternative techniques related to ordering of spatial data are space filling curves, traveling salesman problem and plane sweeping algorithms (Shekhar et al. 2004) (see table 1). Space filling curves are optimization techniques, passing through all centroids in particular order, where they are at an unique distance from each other. Hence, they are effective for spatial ordering of raster data sets. The traveling salesman problem is also an optimization technique that relates the question to find a roundtrip of minimal total length between the centroids, visiting each centroid exactly once. The traveling salesman problem is efficient for transportation and logistics applications; however for detecting spatial patterns in geospatial dynamics they seem too rigid, as ordering of attributes is done only by distance. Alternatively one could work with weighted distances based on selected attribute values. Plane sweeping algorithms reorder data according to the closest distance *d* between centroids, using an initial direction with a perpendicular plane of distance *d*. Plane sweeping algorithms, space filling curves and traveling salesman problem, are useful for reordering spatial data sets and could greatly improve the general applicability of the TOSM for other spatio-temporal applications.

Table 1 Illustration and applications of spatial ordering methods

Linearization method	**Illustration**	**Applications**
Linear directional line		Linear data, such as coast and river compartments, road paths (straight roads), networks. (applied in TOSM environment for all datasets)
Freehand directional line		Non-linear data, such as rivers, stream channels, road paths (curved roads)
Space filling curves		Raster data
Traveling Salesman Problem		Transportation, logistics data
Plane sweeping algorithms		Linear and non-linear data

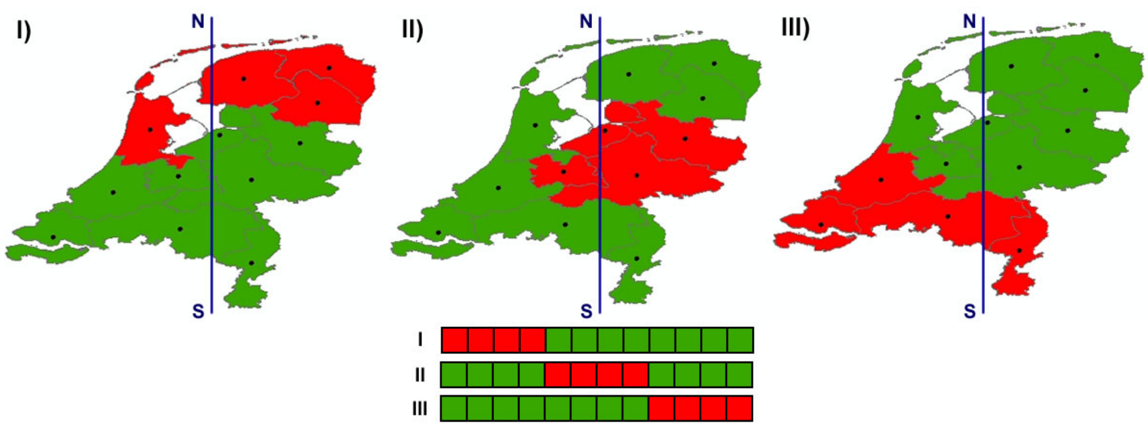


Figure 4 - Time series of an simulated event of a phenomenon moving from north to south through the provinces of the Netherlands. The resulting TOSM is based on the linear directional line.

To judge the effect of each off the above ordering methods as input to the TOSM a simple experiment has been set up. Starting point is a time-series of the administrative areas of the Netherlands simulating how a particular phenomenon moved from north to south trough the country. This controlled event is represented in figure 4 by maps I, I and III. The linear directional line – from north to south – is applied to construct the TOSM. Not surprisingly it results in a spatio-temporal pattern ‘moving’ diagonally from left to right through the matrix.

The four other methods given in table 1 have been applied on this data set. It has to be mentioned that it is very much a preliminary experiment. The freehand directional line has been applied according an order based on how school children learn the provinces from north to south. The space filling curve has been applied on rectangle gird. Only when a grid cell had a centroid in it was added to the TOSM. The travelling sales man algorithm as been applied without special conditions and the plain sweep algorithm has been applied from north to south and from left to right. The results of this exercise are displayed in Figure 5. Looking at the different TOSM’s and comparing these with the one in figure 4 shows that that all are different, but somehow the diagonal trend is visible in all of them. However, this is certainly no argument to claim that all methods could be applied, because the data set was artificially prepared. It is obvious from the experiment that the location of the centroid, e.g. the shape of the geographic units is of great influence on the final result. If one looks at figure 5a a change to a path along the capitals of the provinces instead of the centroid would create a different TOSM. In Figure 5b and 5d starting at the upper left instead of the upper right would result in different TOSMs’. Defining a sales man not to return but stay in the southern most province might have improved the TOSM in 5c.

5. Conclusions

An integrated method to visualize multivariate spatio-temporal data sets is introduced. A prototype is developed consisting of three dynamically linked views, each representing one of spatial data’s components. The MAP environment gives answer to the “where” question, the PCP environment to the “what” question and the TOSM environment to the “when” question. In the prototype, spatial reordering of geographical units is currently done by directional linearization.

A controlled basic experiment with some other spatial ordering methods has been executed showing resulting in different TOSM. From the experiment it is not yet clear if these methods can be used since it appears, despite the fact that the pattern are not completely of the controlled pattern, that the nature of the data set alone already influences the result. And one has to wonder if indeed the patterns that appear have a real meaning. Further experimentation with more different, larger and more complex data set is needed, while the influence of the location of the centroids has to be determined as well.

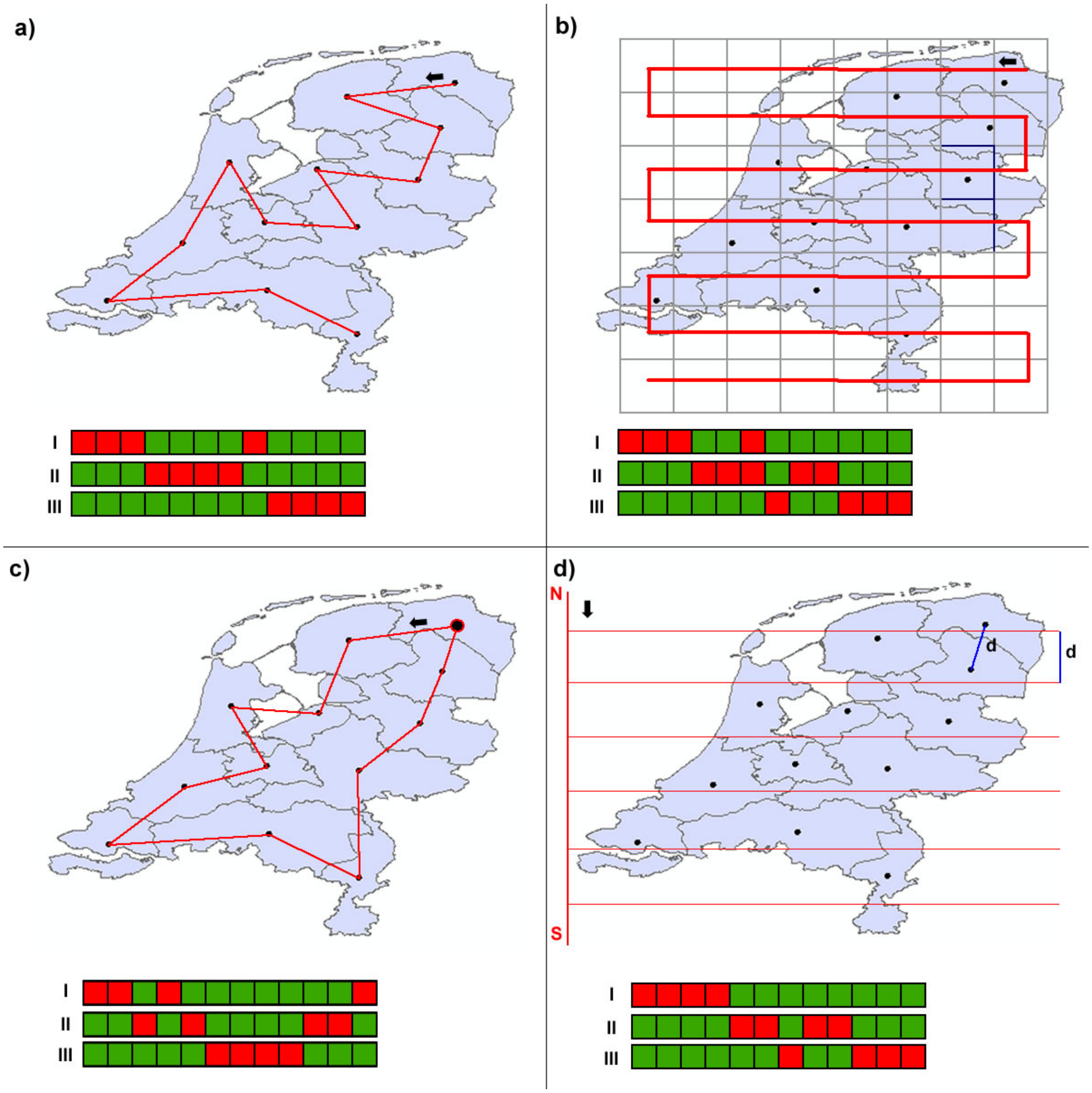


Figure 5. The application of diffent spatial ordering method on the controlled data set presented in figure 4 as input for a TOSM. a) the freehand directional line; b) the space filling curve; c) the travelling sales man algorithm; d) the plane sweep algorithm.

References

Andrienko, G.L., and N.V. Andrienko. 1999. Interactive maps for visual data exploration. *International Journal for Geographic Information Sciences* 13 (4):355-374.

Andrienko, N., G.L. Andrienko, and P. Gatalsky. 2003. Exploratory spatio-temporal visualization: an analytical review. *Journal of Visual Languages and Computing* 14:503-541.

Anselin, L. , I. Syabri, and O. Smirnov. 2002. Visualizing multivariate spatial correlation with dynamically linked windows. Paper read at Proceedings of the SCISS Specialist Meeting 'New tools for spatial data analysis', at Santa Barbara, USA.

Buja, A., D. Cook, and Swayne D. F. 1996. Interactive high-dimensional data visualization. *Journal of Computational and Graphical Statistics* (5):78-99.

DiBiase, D. 1990. Visualization in earth sciences. *Earth & Mineral Sciences, Bulletin of the College of Earth and Mineral Sciences* 59 (2):13-18.

Dykes, J. 1997. cdv: A Flexible Approach to ESDA with Free Demonstration Software. Paper read at British Cartographic Society Meeting.

Dykes, J., A. M. MacEachren, and M. J. Kraak, eds. 2005. *Exploring geovisualization*. Amsterdam: Elsevier.

Edsall, R. M., A. M. MacEachren, and L. J. Pickle. 2001. Case study: design and assessment of an enhanced geographic information system for exploration of multivariate health statistics. Paper read at IEEE Symposium on Information Visualization 2001 (INFOVIS'01), at San Diego, USA.

Edsall, R.M. & Sidney, L.R. 2005. Applications of a Cognitively Informed Framework for the Design of Interactive Spatio-temporal Representations. In *Exploring geovisualization*, edited by J. Dykes, A. M. MacEachren and M. J. Kraak. Amsterdam: Elsevier.

Egenhofer, M. J. 1994. Deriving the composition of binary topological relations. *Journal of Visual Languages and Computing* (5):133-149.

Frank, A. U. 1996. Qualitative Spatial Reasoning: Cardinal Directions as an Example. *International Journal of Geographical Information Science* 10 (3):269-290.

Freksa, C. . 1992. Using Orientation Information for Qualitative Spatial Reasoning. In *Theories and Methods of Spatial-Temporal Reasoning in Geographic Space*, edited by A. U. Frank, I. Campari and U. Formentini. Berlin: Springer.

Gahegan, M. 1999. Four barriers to the development of effective exploratory visualisation tools for the geosciences. *International Journal of Geographical Information Science* 13 (4):289-309.

Guo, D., M. Gahegan, A. M. MacEachren, and B. Zhou. 2005. Multivariate analysis and geovisualization with an integrated geographic knowledge discovery approach. *Cartography and Geographic Information Science* 32 (2):113-132.

Hernandez, D., E. Clementini, and P. Di Felice. 1995. Qualitative Distances, Austria, Lecture Notes in Computer Science 988. In *Spatial Information Theory-A Theoretical Basis for GIS. International Conference COSIT '95 Semmering*, edited by A. U. Frank and W. Kuhn. Berlin: Springer-Verlag.

MacEachren, A. M., M. Wachowicz, R. Edsall, D. Haug, and R. Masters. 1999. Constructing knowledge from multivariate spatio-temporal data: integrating geographical visualization with knowledge discovery in database methods. *International Journal of Geographic Information Science* 13 (4):311-334.

Papadias, D. , and D. Sellis. 1994. Qualitative Representation of Spatial Knowledge in Two Dimensional Space. *VLDB Journal, Special Issue on Spatial Databases* 3 (4):479-516.

Peuquet, D. 1994. It's About Time: A Conceptual Framework for the Representation of Temporal Dynamics in Geographic Information Systems. *Annals of the American Association of Geographers* 84 (3):441-461.

Peuquet, D. J. 2002. *Representations of space and time*. New York: The Guilford Press.

Shekhar, S., P. Zhang, Y. Huang, and R. Vatsavai. 2004. Trends in Spatial Data Mining In *Data Mining: Next Generation Challenges and Future Directions*, edited by H. Kargupta: AAAI Press.

Van de Vlag, D. E. 2006. *Modeling and visualizing dynamic landscape object and their qualities*. Vol. 132, *ITC Disertation*. Enschede: ITC.

Van de Vlag, D. E., B. Vasseur, A. Stein, and R. Jeansoulin. 2005. An Application of Problem and Product Ontologies for the Revision of Beach Nourishments. *International Journal of Geographical Information Science* 19 (10):1113-1130.

Vlag, D. van der, and M. J. Kraak. 2004. Mutlivariate Visualization of Data Quality Issues for Coastal Zone Monitoring. Paper read at ISPRS Conference Proceedings Commission IV, at Istanbul.

———. 2005. Temporal ordered space matrix: representation of multivariate spatiotemporal data. In GISPLANET'05. Paper read at GIS Planet'05 Conference Proceedings, at Lisboa.

Worboys, M. F. 2005. Event-oriented approaches to geographic phenomena. *International Journal of Geographical Information Science* 19 (1):1-28.