



Exploring the spatio-temporal dynamics of geographical processes with geographically weighted regression and geovisual analytics

Urška Demšar¹
A. Stewart Fotheringham¹
Martin Charlton¹

¹National Centre for Geocomputation, National University of Ireland, Maynooth, Co. Kildare, Ireland.

Correspondence:
Urška Demšar, National Centre for Geocomputation, National University of Ireland, Maynooth, Co. Kildare, Ireland.
Tel: +353 1 7086178;
Fax: +353 1 7086455;
E-mail: urska.demsar@nuim.ie

Abstract

The paper examines the potential for combining a spatial statistical methodology – Geographically Weighted Regression (GWR) – with geovisual analytical exploration to help understand complex spatio-temporal processes. This is done by applying the combined statistical – exploratory methodology to a simulated data set in which the behaviour of regression parameters was controlled across space and time. A variety of complex spatio-temporal processes was captured through space-time (i.e. as spatio-temporal) varying parameters whose values were known. The task was to see if the proposed methodology could uncover these complex processes from the data alone. The results of the experiment confirm that the combined methodology can successfully identify spatio-temporal patterns in the local GWR parameter estimates that correspond to the controlled behaviour of the original parameters. *Information Visualization* (2008) 7, 181–197. doi:10.1057/palgrave.ivs.9500187

Keywords: Geographically Weighted Regression (GWR); Geovisual Analytics; visual data exploration; spatio-temporal dynamics; spatio-temporal patterns; spatio-temporal processes

Introduction

In regression analysis a dependent variable is typically modelled as a function of a set of independent variables. This function is usually (but not necessarily) a linear combination of the independent variables, where the coefficients are considered to be stationary over space and produce a global model that is assumed to be representative of the processes operating at every location in the study area. Since relationships in spatial data are often intrinsically different across space, such an assumption can be misleading – the problem of identifying local variation in spatial processes is the concern of several spatial statistical methods, such as Geographically Weighted Regression (GWR).¹

GWR extends the regression framework by dropping the stationarity assumption and considers the parameters of the model to be continuous functions of location. One outcome from a GWR analysis is a set of continuous localised parameter estimate surfaces that describe the processes that generated the data.¹ These results effectively form a large and highly-dimensional spatial data set and can be difficult to interpret – a step that is necessary for proper understanding of the spatial non-stationarity of the geographical processes. This task is even more complicated if the focus is on the spatio-temporal dynamics of the processes and not only on the spatial dimension. GWR run on spatio-temporal data can produce results with several hundreds of dimensions which need to be interpreted. In this paper, we suggest using a Geovisual Analytical approach for the

Received: 17 April 2008
Revised: 30 May 2008
Accepted: 9 June 2008
Online publication date: 31 July 2008

identification and interpretation of patterns in the spatio-temporal GWR parameter estimates. Motivation for this work was a spatio-temporal extension of GWR for hedonic price modelling on house price data in London.² To be able to efficiently and correctly interpret the results of the spatio-temporal GWR for hedonic modelling, we decided to perform an experiment in which we use the combined methodology of GWR and Geovisual Analytics on a simulated data set where the behaviour of the model parameters could be controlled over space and time. This behaviour describes the spatio-temporal processes generating the data and the aim was to see if the combined statistical-exploratory methodology could reproduce patterns that describe this known behaviour. If the method proves successful in this controlled experiment, we can have confidence in its use on real data for spatio-temporal modelling.

The paper is structured as follows: the second section introduces the two methodologies, GWR and Geovisual Analytics, in more detail. The simulated data set and the visual exploration environment are described in the third section. The fourth section presents the exploration results. The fifth section lists some conclusions and discusses the implications of the results.

GWR and geovisual analytics

Global regression models a dependent variable Y as a weighted combination of independent variables:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_n X_{in} + \varepsilon_i \quad (1)$$

In this equation X_{ij} are independent variables at a given location i , ε_i is the residual and β_j are unknown parameters which are assumed to be constant over space (stationary), that is, independent of location i . The model in Eq. (1) is linear, but this is not necessarily always the case. GWR extends this global model by dropping the stationarity assumption. A typical GWR model is represented as:

$$Y_i = \beta_0(i) + \beta_1(i)X_{i1} + \beta_2(i)X_{i2} + \dots + \beta_n(i)X_{in} + \varepsilon_i \quad (2)$$

where the unknown parameters $\beta_j(i)$ are now dependent on location i . These parameters are estimated from the sample data using a geographically weighted ordinary least squares method. Geographical weighting means that instead of taking into account the entire data set when parameters $\beta_j(i)$ are estimated, the estimates $\hat{\beta}_j(i)$ are calculated only from a subset of neighbouring data points, where closer points are assigned higher weights during the calculation. More details about the method and its calibration are given in Fotheringham *et al.*¹ – here we are interested in the structure of the result space of the method and potential for its visual exploration.

The GWR result space is a set of continuous localised parameter estimate surfaces, each of which describes the spatial process linking the respective independent variable with the dependent variable. These surfaces

are typically visualised as a set of separate univariate choropleth maps that are then used to examine the plausibility of the stationarity assumption of the traditional regression and different possible causes of non-stationarity for each separate parameter.^{1,3} The downside of these separate univariate visualisations is that they cannot be used for identification of multivariate spatial and non-spatial relationships and patterns in the parameter space. In an attempt to counter this inadequacy, we previously suggested treating the result space of one single GWR analysis as a multivariate data set and visually exploring it.⁴ The goal was to identify spatial and multivariate patterns that the separate univariate mapping cannot recognise. In this paper, we extend this approach by including the temporal dimension: we combine GWR analysis with Geovisual Analytics by visually exploring the results of a temporal series of GWR analyses.

In exploratory spatial data analysis,⁵ spatial data are explored visually in order to identify patterns that serve as a basis for knowledge generation. Complex data are projected onto the two dimensions of the computer screen using various multivariate visualisation techniques (geographical as well as non-spatial). These visualisations are interactively connected to each other to let the analyst explore and look for patterns, relationships and structure in the data. Results are then used to analytically reason about the data and infer knowledge not only about the characteristics of the data but also about spatial processes that generated the data.⁶ This process is the focus of Geovisual Analytics, ‘the science of analytical reasoning and decision-making with geospatial information, facilitated by interactive visual interfaces, computational methods, knowledge construction representations and management strategies’.⁷ Geovisual Analytics is a part of Visual Analytics,⁸ which is a new discipline in Information Visualisation that has received much attention. Geovisual Analytics is based on geovisualisation^{9,10} and its methods are used to derive knowledge from large highly-dimensional spatial and spatio-temporal data sets.

Here we use a combination of GWR and Geovisual Analytics on an artificially created spatio-temporal data set. We designed an experiment to investigate whether known complex spatio-temporal patterns describing the processes that generate the data can be visually identified in the GWR parameter estimate surfaces. In the experiment, we use GWR to model the spatial processes with known controlled behaviour at the same locations but at different points in time. This is achieved by running a series of GWR analyses on a data set where the regression parameters $\beta_j(i)$ are mathematical functions with known behaviour in space and time. GWR result spaces are then merged into one single highly-dimensional spatio-temporal data set, which is visually explored using a Geovisual Analytical exploratory environment in an attempt to see if the combined statistical-exploratory methodology can reproduce the original patterns. The design of the experiment is presented in the next section.

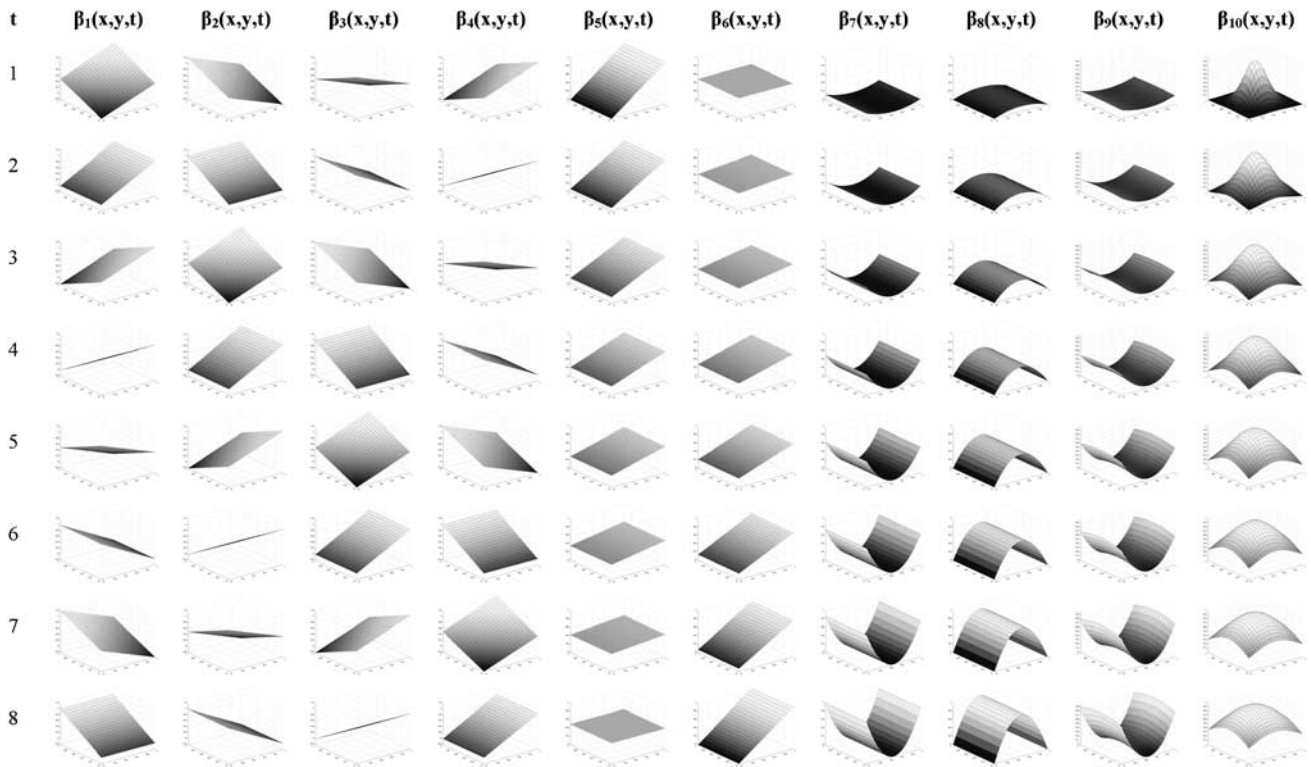


Figure 1 Spatio-temporal behaviour of regression parameters $\beta_j(x, y, t)$, $j = 1, \dots, 10$ in the controlled data set for time steps $t = 1, \dots, 8$. The first six parameters, $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ and β_6 , are planes with rotational behaviour. The second group of parameters, $\beta_7, \beta_8, \beta_9$ and β_{10} , are non-linear functions with increasing/decreasing temporal behaviour.

Analysing spatio-temporal dynamics with GWR and Geovisual Analytics

The simulated data consist of 2500 points spaced evenly in the centres of a 50×50 grid inside a two-dimensional unit square $[0, 1] \times [0, 1]$. We decided to work with a linear GWR model with 10 independent variables X_j , where time was counted in eight discrete steps. The dependent variable Y was created as follows:

$$Y(x, y, t) = \beta_1(x, y, t)X_1(x, y, t) + \dots + \beta_{10}(x, y, t)X_{10}(x, y, t) \tag{3}$$

In this equation, x and y are the coordinates of the location of each point in the unit square, t is a discrete time step (which runs from 1 to 8) and $\beta_j(x, y, t)$ are known parameter functions with controlled behaviour for each location (x, y) and time t . Independent variables, $X_j(x, y, t)$, $j = 1, \dots, 10$, were constructed as random data sequences with mean 0 and standard deviation 1 for each variable and time step – this was done so that the intercept term in the regression model would be constant and equal to 0. Figure 1 shows the behaviour of all parameter surfaces $\beta_j(x, y, t)$ through space and time.

The first four parameter functions, $\beta_1, \beta_2, \beta_3$ and β_4 , were designed as identical rotating planes, with a delay in rotation. β_1 is a plane at 45° inclination, rotating around

the vertical axis that intersects the two-dimensional unit square in its centre $(x, y) = (0.5, 0.5)$. β_2, β_3 and β_4 have the same rotating behaviour as β_1 but each has a two-time-steps delay after the previous one. Parameters β_5 and β_6 are planes rotating around a horizontal axis and mirror each other's behaviour over the eight time steps. β_5 starts at a 45° inclination and slowly rotates downwards around the central axis ($x=0.5$) to reach the constant plane at 0.5 height in the last time step. β_6 has the same rotational behaviour, but in reverse. Parameters $\beta_7, \beta_8, \beta_9$ and β_{10} are non-linear surfaces with specific increasing/decreasing behaviour in time: β_7 and β_8 are parabolic cylinders, β_9 a hyperbolic paraboloid and β_{10} an expanding Gaussian kernel.

Once the dependent variable Y was constructed from these parameter surfaces according to Eq. (3), we ran one GWR analysis for each of the eight time steps and used the results to build a spatio-temporal data set for visual exploration as described below. Mathematical details about the construction of the parameter surfaces and the performance of GWR on this data set are beyond the scope of this paper and are presented elsewhere.¹¹

Each GWR result space consisted of parameter estimate surfaces for each of the 10 dependent variables, that is, functions $\hat{\beta}_j(x, y, t)$, $j = 1, \dots, 10$, plus the estimate for the

regression intercept, $\hat{\beta}_0(x, y, t)$. When merged together, these results form an 88-dimensional spatio-temporal data set with 2,500 data elements – referring to the point locations in the centres of the 50×50 grid in the unit square. This data set can also be seen as consisting of 11 sets of time-dependent parameter estimate surfaces, one for each of the 10 dependent variables and the intercept. These data were then spatially joined with an area representation of grid cells for a clearer geographical visualisation, that is, central points were replaced with their respective grid cells as basic data elements.

To explore these data visually we adapted the exploration system that we previously used for investigation of one single GWR result space.⁴ We skipped the bivariate matrix that did not prove to be very informative in the previous experiment. The system we used consisted of a Self-Organising Map (SOM), two parallel coordinates plots (PCP) – one linked to the SOM and another in its temporal version – and a bivariate geoMap. These visualisations are all briefly described below. The system was built using GeoVISTA Studio, a collection of geographical and multivariate visualisations and computational data mining methods.^{12,13}

A SOM is an unsupervised neural network. It projects multidimensional data onto a two-dimensional space which is usually represented as a lattice of hexagonal cells. This projection preserves the topology and the probability density of the multidimensional input space meaning that it also preserves patterns from the input space.¹⁴ These can be identified easily since the result is two-dimensional and can therefore be nicely visualised in a number of ways.¹⁵ Because of all these properties, a SOM functions as a popular method for knowledge discovery from spatial and spatio-temporal data, either as the original Kohonen SOM operating exclusively on the attribute space or with incorporated geographical location as a GeoSOM.¹⁶ Some recent applications of SOM on spatial and spatio-temporal data can be found in Koua and Kraak,¹⁷ Guo *et al.*¹⁸ Skupin and Hagelman,¹⁹ Demšar²⁰ and Špatenková *et al.*²¹

The SOM visualisation used here is the one implemented in GeoVISTA Studio¹⁸ and is a hexagonal U-matrix. This matrix consists of two types of cells: node cells that represent the nodes of the SOM and distance cells, which are located between node cells and show the level of dissimilarity between their respective neighbouring node cells. This dissimilarity is shown by grey shade: the darker the cell, the more dissimilar are its neighbours. Groups of light cells in the lattice therefore indicate areas with similar cells and as such represent clusters. Dark areas in the lattice indicate borders between clusters. The GeoVISTA Studio implementation of the SOM visualisation shows the distribution of the data in the SOM through circles that are superimposed on the node cells. The larger the circle, the more data elements have been assigned to the respective node cell. The colour of the circles is defined through a smooth 2D colour map and is used in visual brushing to connect the SOM to all

other visualisations, that is, the map, the temporal PCP and the SOM PCP. Another connection to the SOM PCP is through the size of the circles, which defines the width of the polygonal lines in the SOM PCP as described below. More details about this particular SOM visualisation can be found in Guo *et al.*¹⁸

The map in our exploratory environment is the geoMap from GeoVISTA Studio. We use it as a thematic map showing the spatial distribution of the SOM results through a colour ramp that has been transferred to the map from the SOM via visual brushing.^{12,18}

There are two PCPs in this environment. One is directly linked to the SOM. In this SOM PCP, each vertical axis represents one dimension of the input space and each polygonal line represents a node cell from the SOM,^{22,18} – a property that is valuable for cluster analysis. In the second PCP, each polygonal line represents one data element and intersects each of the axes at the value that corresponds to the data element value in that particular dimension.²³ The two PCPs differ also in the way their axes are scaled. In the SOM PCP, the axes are scaled using nested means scaling,²² the purpose of which is to reduce overprinting. While this scaling distorts the statistical distribution of data at each axis, it is useful for anomaly detection. We used a temporal version of the second PCP to investigate the time series of parameter estimates for each regression variable. In a temporal PCP all the axes are scaled linearly over the same range – the minimum and maximum values across all time intervals. This scaling variation is required in a PCP to successfully visualise temporal trends.²⁴

We divided the exploration of the spatio-temporal dynamics of the simulated data set into three approaches. Approach I was to look at separate regression variable-defined subspaces (i.e. time series of GWR parameter estimates for each variable) and examine the behaviour of each parameter estimate separately. Approach II was to look at subspaces consisting of time series of several parameter estimates, grouped according to similarity. Similarity grouping was based on the behaviour of the original parameter surfaces. We looked at two separate groups, the first one containing parameter estimates $\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4, \hat{\beta}_5$ and $\hat{\beta}_6$ where the original parameters are linear surfaces (planes) and the second one consisting of the estimates $\hat{\beta}_7, \hat{\beta}_8, \hat{\beta}_9$ and $\hat{\beta}_{10}$ where the original parameters are non-linear surfaces. Approach III was to explore the entire data set at once.

The exploration process consisted of looking for three types of patterns that would help the interpretation of GWR results:

Type 1 – patterns that identify areas of temporal stability for each separate parameter estimate (approach I). We do not consider temporal stability as mathematically constant because of the noise present in GWR modelling. For our purpose it is enough to identify areas where parameter estimates maintain approximately similar values throughout the entire time series.

Type 2 – patterns that identify areas that behave similarly in time:

- (a) for each separate parameter estimate (approach I)
- (b) for a group of parameters (approaches II and III)

Type 3 – patterns that identify parameters with similar temporal behaviour. (approaches II and III)

These particular types of patterns were chosen based on the application of spatio-temporal GWR on hedonic modelling of house prices.² The behaviour of the original parameter surfaces $\beta_j(x, y, t)$ in the simulated spatio-temporal data set was designed with these patterns in mind. Table 1 explains how the three types of patterns appear in the behaviour of each of the 10 original parameter surfaces.

The next section presents the results of our experiment in an attempt to answer the question posed in this paper: is the combination of GWR and visual exploration able to identify the patterns that we know exist in the original parameter surfaces $\beta_j(x, y, t)$?

Results

We have been able to visually identify all of the expected pattern types (1, 2 and 3) in all relevant exploratory approaches (I, II and III). Table 2 summarises the results for approach I (i.e. exploration of separate parameter estimate-based subspaces) and Table 3 summarises the results for approaches II and III (i.e. for groups of similar parameter estimates and for the entire data set). The tables also describe how we looked for particular patterns and whether we were able to identify them for each particular exploration approach.

In the rest of this section we present a selection of the results.

Approach I – Exploring separate parameter-defined subspaces

Parameter β_1 is an inclined plane that rotates around a vertical axis. Figure 2 shows the visualisation of the time series of its GWR estimate $\hat{\beta}_1$ in the exploratory environment. The SOM in Figure 2B was obtained by clustering the time series of $\hat{\beta}_1$, that is, its eight time steps as the variables of the SOM input space. As in all GeoVISTA Studio-based exploration systems, all visualisations are linked to each other through visual brushing with colour, meaning in this case (and in the rest of the examples presented) that all visualisations inherit their colours from the SOM. Visualisations are also connected through interactive brushing and selection, which is important for a successful exploration process.

This result is an example of how pattern type 2a – areas with similar temporal behaviour for one particular parameter estimate – can be observed. In the case of $\hat{\beta}_1$, because of the variation in the parameter surface over time, we expected to see areas with similar behaviour in terms of

the variation in local estimates through time distributed in concentric circles around the central point of the region. Additionally, diagonally opposed areas would be expected to have delayed (and therefore opposite) rotational behaviour by four time steps. All this can be seen in the visualisations of $\hat{\beta}_1$ in Figure 2. The map (Figure 2A) shows a circular spatial structure. The colour of the grid cells in the map is inherited from the SOM (Figure 2B), where it was defined by the clustering performed in the attribute space. Spatial position was not a part of the SOM clustering and yet these attribute-defined (non-spatial) SOM clusters are located in spatially adjacent areas in the map and form a circular pattern. That areas located diagonally opposite on the map are very different from each other, can be deduced from the SOM: the SOM shows a green–yellow cluster on the top border, a pink cluster on the right border and a turquoise–blue cluster in the bottom left area of the lattice. These clusters are very different from each other as evidenced by their separation with areas of dark-coloured cells and by their locations on opposite sides of the SOM. What we expected to see was that diagonally opposed areas in the map should have a sinusoidal track through both PCPs and that one of these tracks is delayed for four time steps. The SOM clustering picks up this behaviour, which is even easier to see in Figures 2E, 2F and 2G, where we selected the two opposite clusters in the SOM, the green–yellow one on the top and the blue one on the bottom. The Temporal PCP (Figure 2G) shows the expected two sinusoidal tracks, delayed by four time steps. A comparison with the map (Figure 2E) also shows that data elements that belong to these two clusters/tracks are located at the opposite corners of the unit square as expected.

Figure 3 shows visualisations of the time series of the parameter estimate $\hat{\beta}_5$. The SOM in Figure 3B shows two clusters located opposite to each other, a green–turquoise cluster in the upper left area of the SOM and a violet–pink cluster in the lower right area. These clusters are very different in the attribute space because they are separated by large groups of dark-coloured cells as well as being located far from each other in the SOM. When these clusters are checked on the map (Figure 3A), the green–turquoise one defines the band on the right border of the unit square and the pink–violet one the band on the left border. The behaviour of these two clusters can then be further investigated in both PCPs, where the respective green–turquoise and pink–violet tracks increase/decrease steadily through the temporal PCP (Figure 3D) and are nearly horizontal in the SOM PCP (Figure 3C). This example also shows the difference between the nested-means scaling of the SOM PCP and the min-max-over-the-entire-temporal-range linear scaling of the temporal PCP. The identified pattern corresponds very well to the areas of similarity in the original parameter β_5 (Tables 1 and 2).

Additionally, we can see the area of temporal stability (pattern type 1) in this visualisation: according to Table 1,

Table 1 Patterns in original parameter functions, related to GWR interpretation

Parameter function	Pattern 1: areas of temporal stability and the constant value in these areas	Pattern 2: areas with similar temporal behaviour	Pattern 3: parameter functions with similar behaviour
$\beta_1(x, y, t)$	A small circular area around the central point $(x, y) = (0.5, 0.5)$. Value = 0.5.	Concentric circles around the central point, where diagonally opposed areas have delayed rotational behaviour by four time steps.	β_2, β_3 and β_4 have identical temporal behaviour, with a difference of two time steps, respectively.
$\beta_2(x, y, t)$	A small circular area around the central point $(x, y) = (0.5, 0.5)$. Value = 0.5.	Concentric circles around the central point, where diagonally opposed areas have delayed rotational behaviour by four time steps.	β_1, β_3 and β_4 have identical temporal behaviour, with a difference of two time steps, respectively.
$\beta_3(x, y, t)$	A small circular area around the central point $(x, y) = (0.5, 0.5)$. Value = 0.5.	Concentric circles around the central point, where diagonally opposed areas have delayed rotational behaviour by four time steps.	β_1, β_2 and β_4 have identical temporal behaviour, with a difference of two time steps, respectively.
$\beta_4(x, y, t)$	A small circular area around the central point $(x, y) = (0.5, 0.5)$. Value = 0.5.	Concentric circles around the central point, where diagonally opposed areas have delayed rotational behaviour by four time steps.	β_1, β_2 and β_3 have identical temporal behaviour, with a difference of two time steps, respectively.
$\beta_5(x, y, t)$	A narrow band around the axis of rotation, $x = 0.5$. Value = 0.5.	Bands parallel to the axis of rotation, $x=0.5$. Bands on the left side of the axis move upwards through time, bands on the right side move downwards.	β_6 has matching temporal behaviour – in reverse.
$\beta_6(x, y, t)$	A narrow band around the axis of rotation, $x = 0.5$. Value = 0.5.	Bands parallel to the axis of rotation, $x=0.5$. Bands on the left side of the axis move downwards through time, bands on the right side move upwards.	β_5 has matching temporal behaviour – in reverse.
$\beta_7(x, y, t)$	A narrow band around the central axis where the minimum of the parabola is located, $x = 0.5$. Value = 0.	Bands parallel to the central axis $x = 0.5$. Bands at the same distance from the axis on both sides of the axis move upwards with the same speed.	β_8 has similar but upside-down temporal behaviour, β_9 has similar temporal behaviour.
$\beta_8(x, y, t)$	Bands on both sides of the unit square, $x = 0, x = 1$. Value = 0.	Bands parallel to the central axis $x = 0.5$. Bands at the same distance from the axis on both sides of the axis move upwards with the same speed.	β_7 and β_9 have similar but upside-down temporal behaviour.
$\beta_9(x, y, t)$	A narrow band around the central axis where the minimum of the parabola is located, $x = 0.5$. Value = 0.	Hyperbolic bands 'parallel' to the central axis $x = 0.5$. Bands at the same distance from the axis on both sides of the axis move upwards with the same speed.	β_7 has similar temporal behaviour. β_8 has similar but upside-down temporal behaviour.
$\beta_{10}(x, y, t)$	A small circular area around the central point of the Gaussian kernel $(x, y)=(0.5, 0.5)$. Value=1.	Concentric circles around the central point move upwards with the same speed.	Behaviour of no other parameter behaviour is similar to this one.

The table explains how each original parameter function $\beta_j(x, y, t)$ is linked to the three types of patterns we were looking for.

what we expected to see is that the values in a narrow band around the central axis are relatively stable. This band can be identified on the map as the area with light-blue and light-orange/beige colours. The lines of these colours follow a relatively flat track through both PCPs and they cross all of the axes more or less around the centre. The

effect can be seen even more clearly in an appropriate interactive selection of only these areas.

An interesting question is what would we expect to see if there were no spatial or temporal patterns in the time series of a GWR parameter estimate? This does not occur for any of the original parameters that

Table 2 Exploration framework for Approach I – in separate parameter-defined subspaces of the GWR results

Pattern type	Parameter estimates	Expected pattern	Identified?	How?
1 – areas of temporal stability	$\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4$ (each separately)	A small circular area in the centre of the square. Value = 0.5.	Yes, to some extent.	In all cases the exploration was done in three separate steps: 1. A combination of SOM, SOM PCP and map was first used to select clusters in relevant spatial location. Selection checked in temporal PCP to see if the values in this selection were relatively constant and close to expected value. 2. A new selection was then made in the first axis of the temporal PCP around the expected value. The track in the temp PCP checked for relatively constant values. Map checked for correct spatial position of the selection. 3. Other values (apart from expected value) were iteratively selected in the first axis of the temporal PCP and tracks checked if they were relatively constant or not. None of them were.
	$\hat{\beta}_5, \hat{\beta}_6$ (each separately)	A narrow band around the axis of rotation. Value = 0.5.	Yes, clearly.	
	$\hat{\beta}_7, \hat{\beta}_9$ (each separately)	A narrow band around the central axis. Value = 0.	Yes, clearly.	
	$\hat{\beta}_8$	Bands on both sides of the square. Value = 0.	Yes, clearly.	
	$\hat{\beta}_{10}$	A small circular area in the centre of the square. Value = 1.	Yes, clearly.	
	$\hat{\beta}_0$	No such pattern exists.	Yes, clearly.	
2a – areas of similar temporal behaviour for one parameter estimate	$\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4$ (each separately)	Concentric circles around the centre. Diagonally opposed areas have delayed behaviour.	Yes, clearly.	In all cases, the exploration procedure was as follows: The most similar/most different clusters were identified in the SOM and selected in the SOM PCP. The spatial distribution of elements belonging to these clusters was checked by looking at the selection in the map. The tracks belonging to this same selection were checked in the temporal PCP to see if the expected temporal behaviour was present or not.
	$\hat{\beta}_5, \hat{\beta}_6$ (each separately)	Bands parallel to the central axis. Upward/downward rotation at the sides.	Yes, clearly.	
	$\hat{\beta}_7, \hat{\beta}_8$ (each separately)	Bands parallel to the central axis move upwards.	Yes, clearly.	
	$\hat{\beta}_9$	Hyperbolic bands parallel to the central axis move upwards.	Yes, clearly.	
	$\hat{\beta}_{10}$	Concentric circular bands around the centre move upwards.	Yes, clearly.	
	$\hat{\beta}_0$	No such pattern exists.	Yes, clearly.	

The table lists the pattern types that we looked for in each of the parameter-defined subspaces, the expected pattern, if it was identified visually ('yes, clearly', 'yes, to some extent', 'no') and how it was identified, that is, which visualisations/functionalities were used for this purpose.

Table 3 Exploration framework for Approaches II and III – for groups of similar parameter estimates and for the entire data set

Approach	Pattern type	Parameter estimates	Expected pattern	Identified?	How?
II – groups of similar parameters	2b – areas of similar temporal behaviour for a group of parameter estimates	$\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4, \hat{\beta}_5, \hat{\beta}_6$	A combination of concentric circles around the centre and parallel bands. Diagonally opposed areas have delayed behaviour.	Circular and diagonally opposed pattern yes, clearly. Left-to-right parallel bands, yes to some extent, but less clearly than the circles.	The exploration procedure was as follows: SOM was used for identification of most similar/most different clusters and SOM PCP for the selection of these. Spatial distribution of elements belonging to these clusters was checked by looking at the selection in the map. Tracks belonging to this same selection were checked in the temporal PCP to see if the expected temporal behaviour as in the original surfaces was present or not.
		$\hat{\beta}_7, \hat{\beta}_8, \hat{\beta}_9, \hat{\beta}_{10}$	A combination of concentric circles around the centre and two types of parallel bands, straight and hyperbolic.	Yes, clearly.	
	3 – parameters with similar temporal behaviour	$\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4, \hat{\beta}_5, \hat{\beta}_6$	1–4 have identical delayed rotational temporal behaviour. Five and six have mirrored temporal behaviour.	Yes, clearly.	
		$\hat{\beta}_7, \hat{\beta}_8, \hat{\beta}_9, \hat{\beta}_{10}$	Seven and nine have similar behaviour, eight mirrors that. Ten is different from all other parameters	Yes, clearly.	
III – entire data set	2b – areas of similar temporal behaviour for the entire data set	$\hat{\beta}_1, \dots, \hat{\beta}_{10}$	A combination of concentric circles around the centre and two types of parallel bands, straight and hyperbolic.	Yes, to some extent.	As before, the exploration procedure was as follows: SOM was used for identification of most similar/most different clusters and SOM PCP for the selection of these. Spatial distribution of the clusters was checked in the map. Tracks belonging to this same selection were checked in the temporal PCP to see if the expected temporal behaviour as in the original surfaces was present or not.
		$\hat{\beta}_1, \dots, \hat{\beta}_{10}$	As above, see the corresponding two fields in Approach II.	Yes, clearly.	

The table lists the approach, the pattern types that we looked for, the expected pattern, if it was identified visually ('yes, clearly', 'yes, to some extent', 'no') and how it was identified, that is, which visualisations/functionality were used and how.

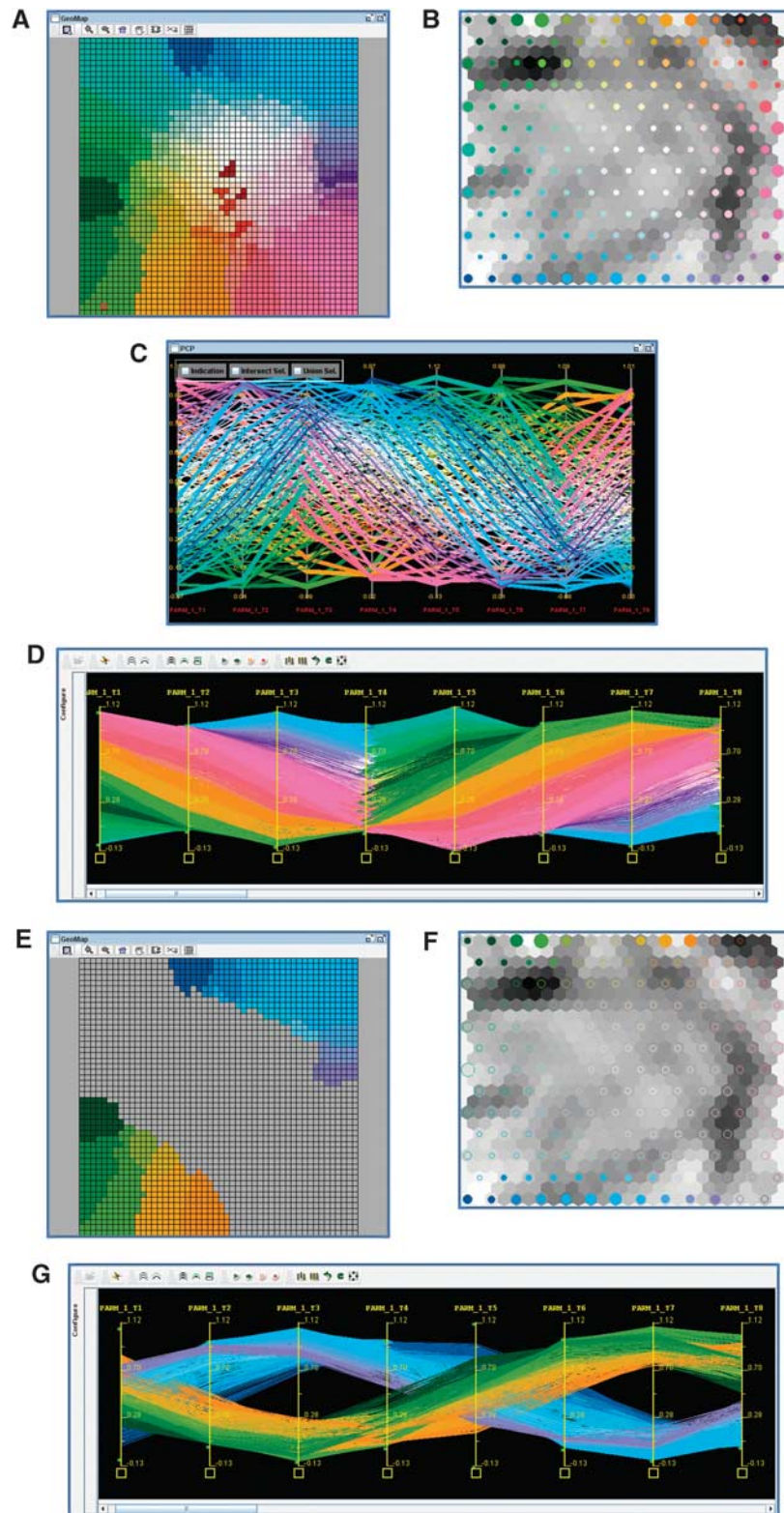


Figure 2 Visualisations of the time series of parameter estimate $\hat{\beta}_1$. Results are presented in (A) the geoMap, (B) the SOM, (C) the SOM PCP and (D) the temporal PCP. A selection of two areas with different temporal behaviour is shown in (E) the geoMap, (F) the SOM and (G) the temporal PCP. The selection shows two different sinusoidal tracks in the temporal SOM with a delay of four time steps and identifies two areas with opposite temporal behaviour. Colours in all visualisations are inherited from the SOM.

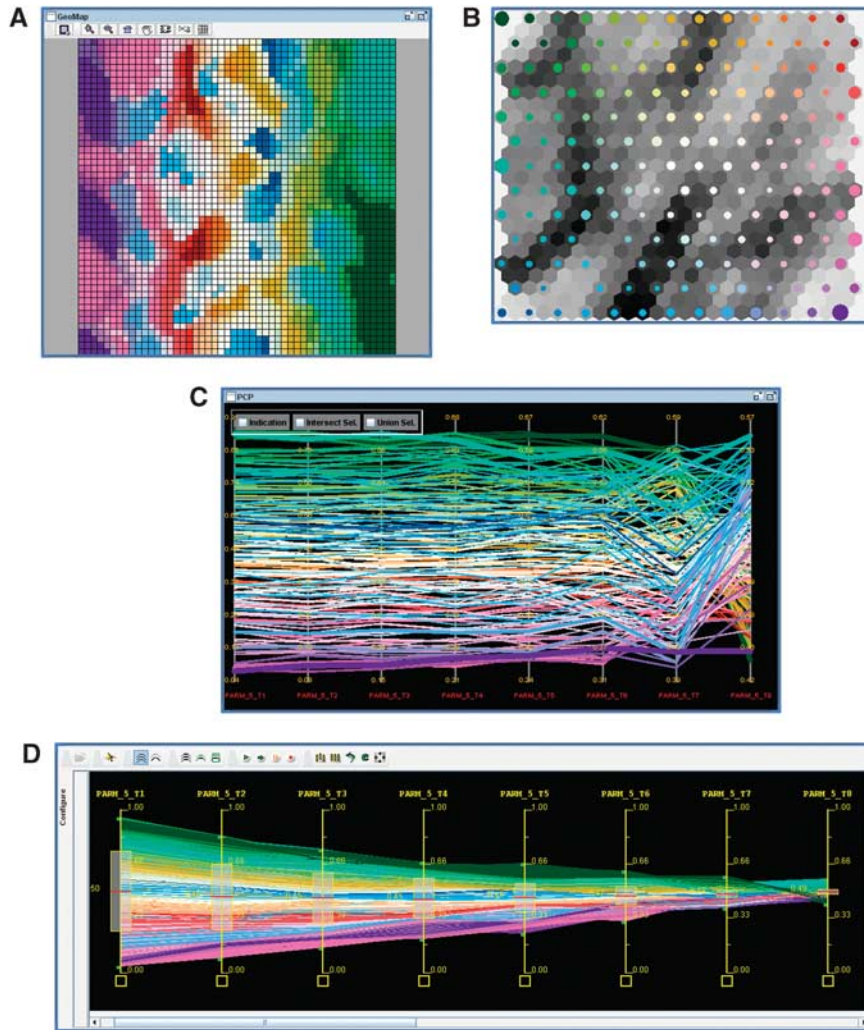


Figure 3 Visualisations of the time series of GWR parameter estimate $\hat{\beta}_5$. Results are presented in (A) the geoMap, (B) the SOM, (C) the SOM PCP and (D) the temporal PCP. Colours in all visualisations are again inherited from the SOM.

belong to the independent regression variables in our model. However, this situation does arise for the estimated intercept. Regression variables X_1, \dots, X_{10} were chosen in such a way that the intercept in the GWR model is constant and equal to 0. Because of this, the intercept estimate $\hat{\beta}_0$ should only show random noise with no particular spatial or temporal pattern. To confirm or reject this expectation, we visualised the time series of the intercept estimate $\hat{\beta}_0$ (Figure 4). While there seems to be a small amount of attribute clustering present in the SOM (Figure 4B), there is clearly no spatial clustering present in the map (Figure 4A). The values of the intercept estimate are very low (the temporal PCP is scaled from -0.06 to 0.08 which is low compared to all other parameter estimates that range from 0 to 1). Neither of the two PCPs contains any clearly recognisable temporal tracks (Figures 4C and D). As expected, we only see random noise.

Approach II – Exploring subspaces defined by groups of parameters

In approach II we explored two separate subspaces, the first defined by the set of the parameter estimates $\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4, \hat{\beta}_5$ and $\hat{\beta}_6$ and the second by the set of the estimates $\hat{\beta}_7, \hat{\beta}_8, \hat{\beta}_9$ and $\hat{\beta}_{10}$. Here we present the exploration results for the second group. Figure 5 shows the relevant visualisations. The SOM in Figure 5B was produced by clustering $\hat{\beta}_7, \hat{\beta}_8, \hat{\beta}_9$ and $\hat{\beta}_{10}$ together. The colours from the SOM are again transferred to the map in Figure 5A and all the PCPs. The SOM PCP (Figure 5C) has 32 axes, which are four time series with eight steps for each of the four parameter estimates in this group. Each parameter estimate also has its own temporal PCP (Figures 5D, E, F, and G for estimates $\hat{\beta}_7, \hat{\beta}_8, \hat{\beta}_9$ and $\hat{\beta}_{10}$, respectively). Each temporal PCP is scaled from min–max across all of its eight axes.

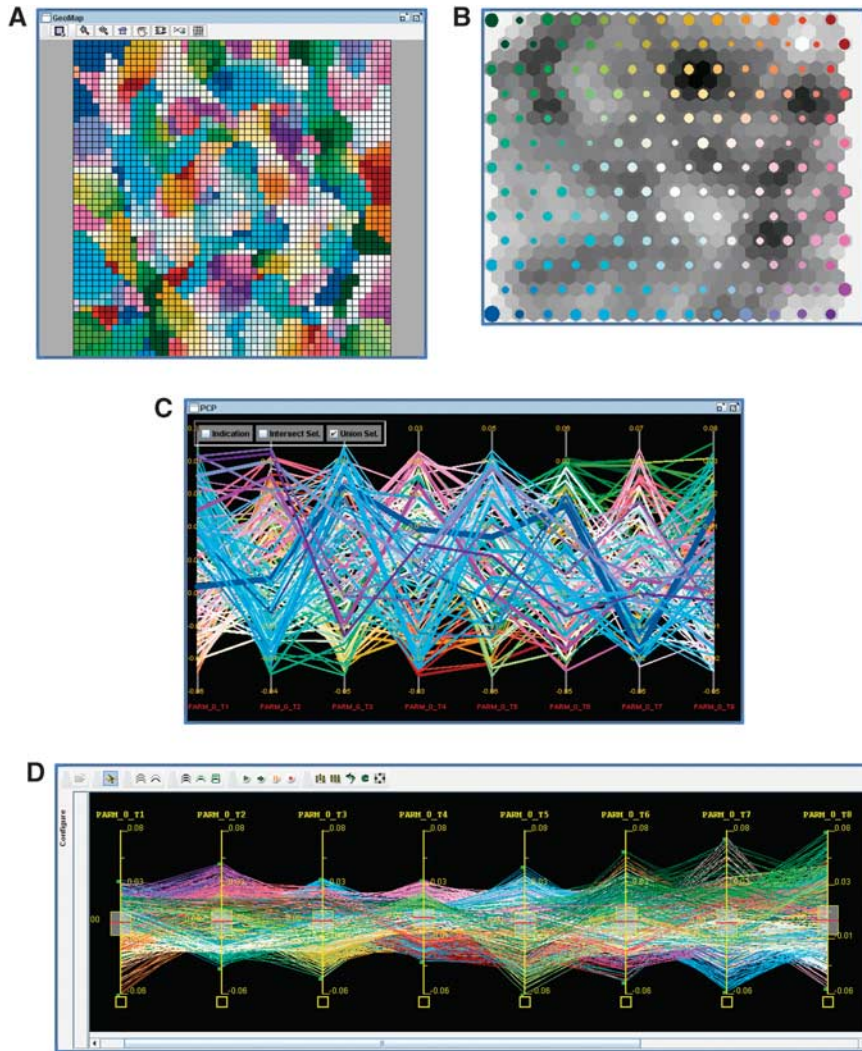


Figure 4 Visualisations of the time series of the intercept estimate $\hat{\beta}_0$. Results are presented in (A) the geoMap, (B) the SOM, (C) the SOM PCP and (D) the temporal PCP. Colours in all visualisations are again inherited from the SOM. There is no spatial or temporal clustering present in this estimate.

To identify the areas that behave similarly in time across all four parameter estimates (pattern type 2b), we need to look at how areas with similar behaviour were defined in the original parameters β_7 , β_8 , β_9 and β_{10} (Table 1). Parameters β_7 , β_8 and β_9 have similar behaviour in bands parallel to both sides of the unit square with the additional property that the β_7 and β_8 bands are straight, while the β_9 bands are hyperbolic. The behaviour of β_{10} is similar in concentric circles around the central point. The geometrical pattern we expected to see in the visualisations of the group of estimates $\hat{\beta}_7$, $\hat{\beta}_8$, $\hat{\beta}_9$ and $\hat{\beta}_{10}$ should therefore show a combination of parallel bands (straight and hyperbolic) and concentric circles (Table 3). This is exactly what the map in Figure 5A shows.

For the parameters β_7 , β_8 and β_9 the side bands and the band around the central axis behave in an opposite

manner. This expected pattern can be clearly seen if a selection of these areas is made in the map (Figure 6A). The opposite behaviour is very clear in the SOM PCP (Figure 6B, notice how the two tracks, red-pink and blue, are very separate in the first three time series for $\hat{\beta}_7$, $\hat{\beta}_8$ and $\hat{\beta}_9$ while there is no separation in the time series for the estimate $\hat{\beta}_{10}$) and in the relevant temporal PCPs (Figure 6C shows the temporal PCP for $\hat{\beta}_9$). A different selection of the central and peripheral circular areas in the map in Figure 6D confirms that the behaviour of the estimate $\hat{\beta}_{10}$ is different from the behaviour of the other three estimates. This is what we expected to see based on the behaviour of the original parameter β_{10} (Tables 1 and 3). The separation between the red and blue tracks is now clear in the last (fourth) time series in the SOM

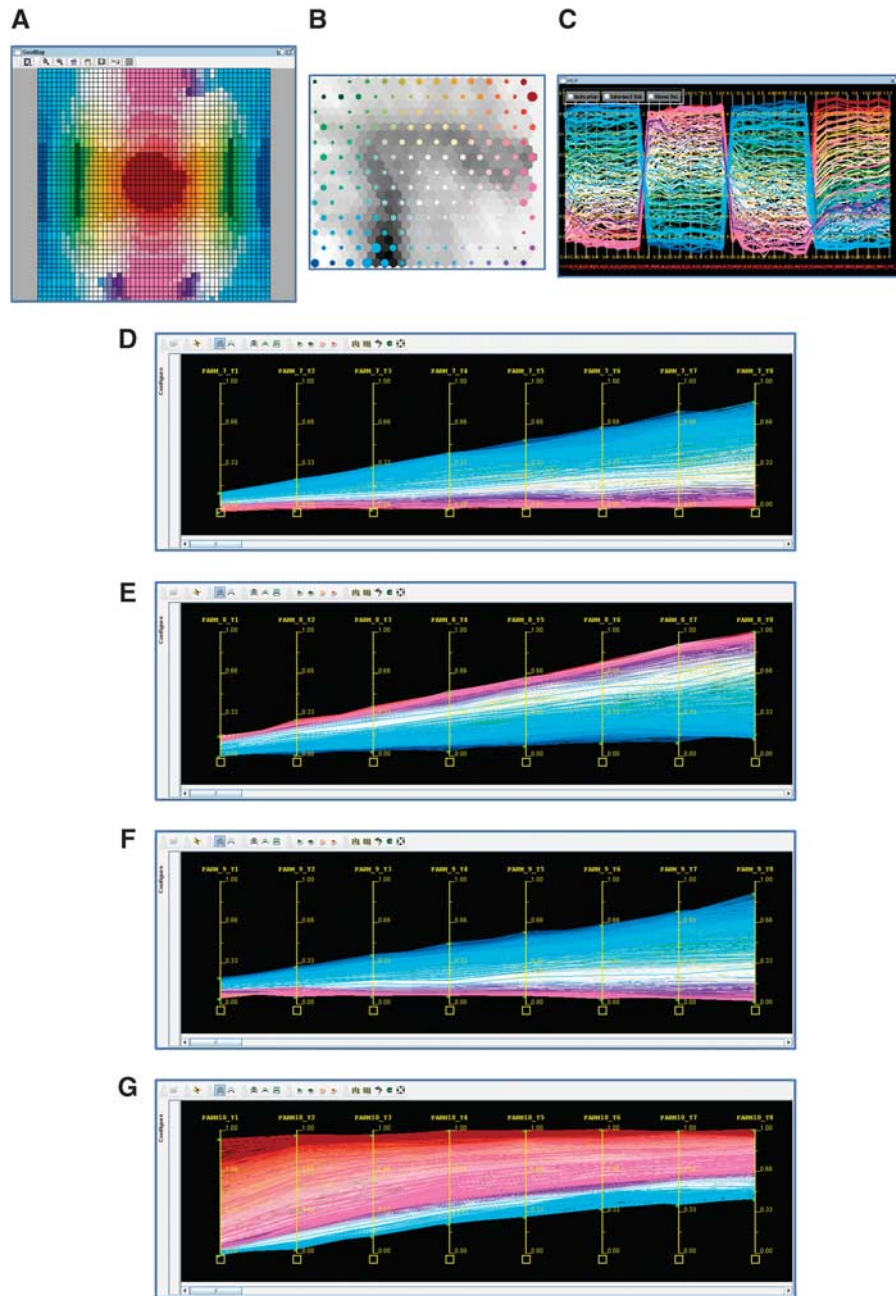


Figure 5 Visualising parameter estimates $\hat{\beta}_7$, $\hat{\beta}_8$, $\hat{\beta}_9$ and $\hat{\beta}_{10}$. (A) shows the map, (B) the SOM and (C) the SOM PCP. The rest of the figure shows temporal PCPs for (D) $\hat{\beta}_7$, (E) $\hat{\beta}_8$, (F) $\hat{\beta}_9$ and (G) $\hat{\beta}_{10}$. Colours in all visualisations are inherited from the SOM.

PCP (Figure 6E) and in the temporal PCP for $\hat{\beta}_{10}$ (Figure 6G). There is no separation in the first three time series in the SOM PCP (Figure 6E), nor in the temporal PCP for $\hat{\beta}_9$ (Figure 6F, and there is also no separation in the temporal PCPs for $\hat{\beta}_7$ and $\hat{\beta}_8$ which are not shown here). These two selections therefore help to uncover which spatial areas behave similarly/differently across the entire set of four parameter estimates.

The last type of pattern we were looking for concerned the identification of attributes with similar behaviour over time (pattern type 3). What we expected to see from the original parameters was the following (compare with Tables 1 and 3 and Figure 1): β_7 , β_8 and β_9 all exhibit a similar increase across time but β_8 is spatially mirrored from the other two (i.e. β_7 , and β_9 have the largest increase on the sides, while β_8 has the largest increase

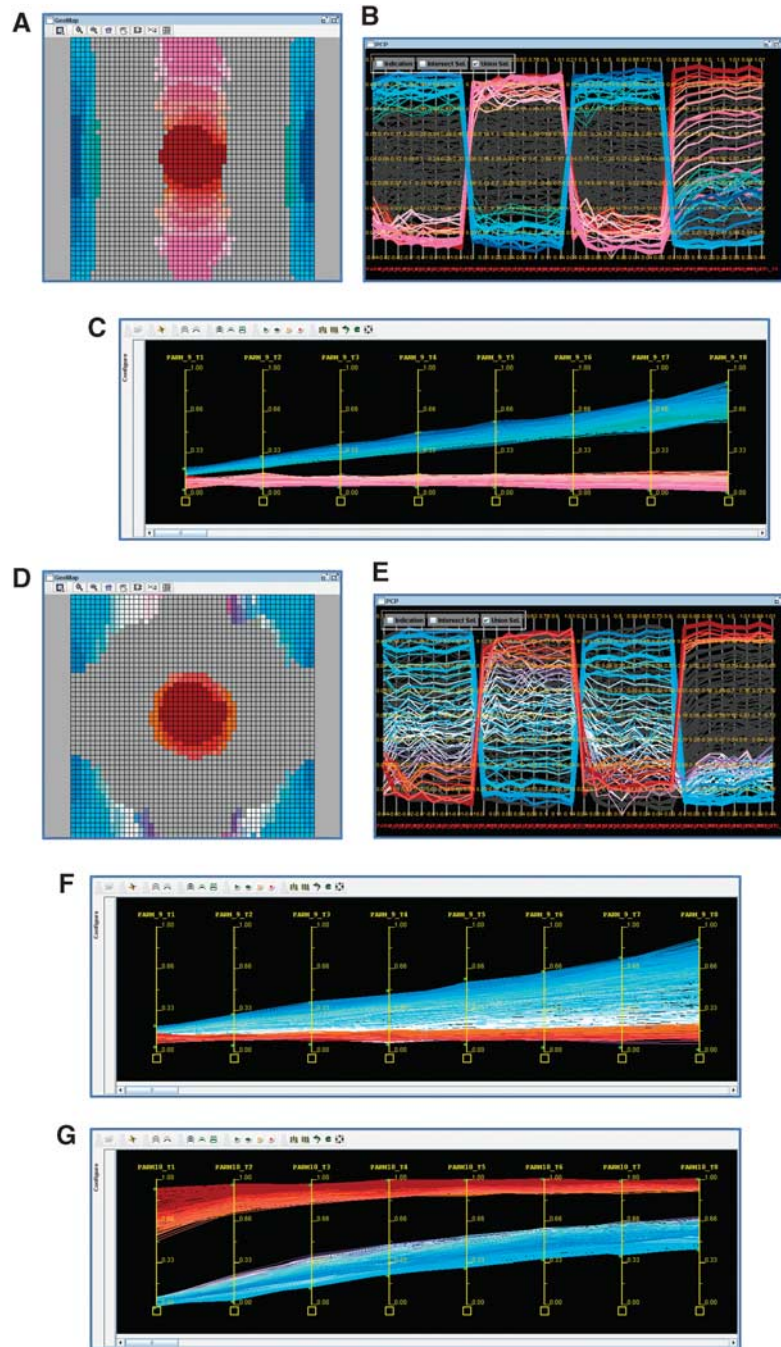


Figure 6 Two different spatial selections in the visualisations of the $\hat{\beta}_7, \hat{\beta}_8, \hat{\beta}_9$ and $\hat{\beta}_{10}$ subset. The first selection shows the central band and the sides of the unit square in (A) the map, (B) the SOM PCP and (C) the temporal PCP for $\hat{\beta}_9$. In the second selection, the central and peripheral circular areas (D) exhibit different temporal behaviour in (E) the SOM PCP, (G) the temporal PCP for $\hat{\beta}_{10}$, but not in (F) the temporal PCP for $\hat{\beta}_9$.

in the central band). The pattern of increase in β_{10} is different from the rest, as it is circular.

The anticipated similarity in the temporal patterns of the parameter estimates $\hat{\beta}_7, \hat{\beta}_8$ and $\hat{\beta}_9$ is clearly noticeable

if we compare their respective temporal PCPs (Figure 5D, 5E and 5F), in spite of a large amount of overprinting. The PCPs of $\hat{\beta}_7$ and $\hat{\beta}_9$ show a similar blue to pink pattern, while the pattern in $\hat{\beta}_8$ is mirrored (pink to blue). The

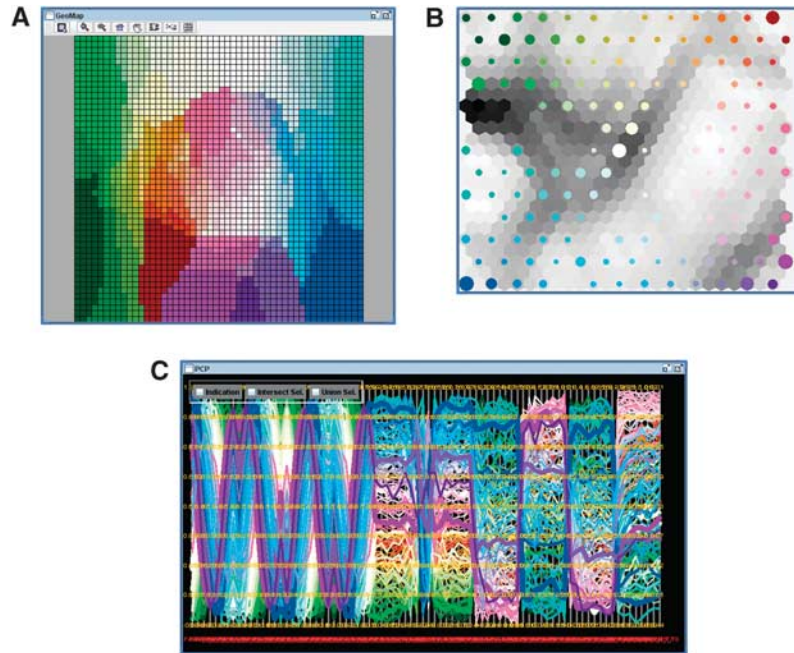


Figure 7 Visualising the entire data set. With (A) the map, (B) the SOM and (C) the SOM PCP with 80 dimensions. Dimensions are shown in the temporal order for each parameter estimate, $\hat{\beta}_1$ – $\hat{\beta}_{10}$.

pattern for the estimate $\hat{\beta}_{10}$ (Figure 5G) is also clearly different from the other three.

Approach III – Exploring the entire data set

In the final step of the experiment, we looked at the entire data set simultaneously, that is, for the time series of all 10 parameter estimates $\hat{\beta}_j$, $j = 1, \dots, 10$, but excluding the intercept estimate as that did not exhibit any spatial or temporal pattern (as described above). Figure 7 shows three of the visualisations produced in this step. The SOM clustering in Figure 7B was based on 80 variables (i.e. eight time steps in 10 time series) and its colour scheme is transferred to the map (Figure 7A) and the SOM PCP (Figure 7C). We were looking for two types of patterns in these visualisations: areas with similar temporal behaviour (pattern type 2b) and parameter estimates with similar behaviour (pattern type 3).

According to the areas of similar temporal behaviour of all 10 original parameters (Table 1), we expected to see a combination of the following geometrical patterns in the map (pattern type 2b): concentric circular areas and bands (straight and hyperbolic) parallel to both sides of the unit square (Table 3). The pattern in the map in Figure 7A does show these characteristics, although the impression is not as clear as in the previous approach. The side bands on the map in Figure 7A are quite clear but the circular pattern is not very obvious. The difference between the behaviour of the side bands and the central areas is large – this can be deduced by looking at the respective colour clusters in the SOM and their distribution in the SOM.

Differences and similarities of the behaviour of various spatial areas can be further confirmed when the areas in question are interactively selected in the map and/or the SOM PCP and the temporal trends in various PCPs compared to each other.

The SOM PCP of the entire data set (Figure 7C) helps with the identification of similar parameter estimates (pattern type 3). Based on the temporal behaviour of the original parameters we expected to find delayed rotation of $\hat{\beta}_1$ to $\hat{\beta}_4$; a mirrored behaviour of $\hat{\beta}_5$ and $\hat{\beta}_6$; similar increases in $\hat{\beta}_7$, $\hat{\beta}_8$, $\hat{\beta}_9$ where $\hat{\beta}_8$ mirrors the other two; and $\hat{\beta}_{10}$ exhibiting different temporal behaviour from the rest (compare with Tables 1 and 3 and Figure 1). Most of this can be seen in the SOM PCP (Figure 7C) and the temporal PCPs for each respective parameter estimate (these are not shown here due to limitations in paper length).

Conclusions and discussion

This paper presents an experiment designed to examine the viability of a visual exploration of the results of a spatial statistical method, GWR, for easier interpretation of spatio-temporal processes. This was done by constructing a simulated spatio-temporal data set where we controlled the behaviour of regression parameters across space and time. The goal was to see if patterns confirming this behaviour could be visually identified when we explored the results of GWR run with this data set. The model employed was a linear regression model although the methodology is readily transferable to other

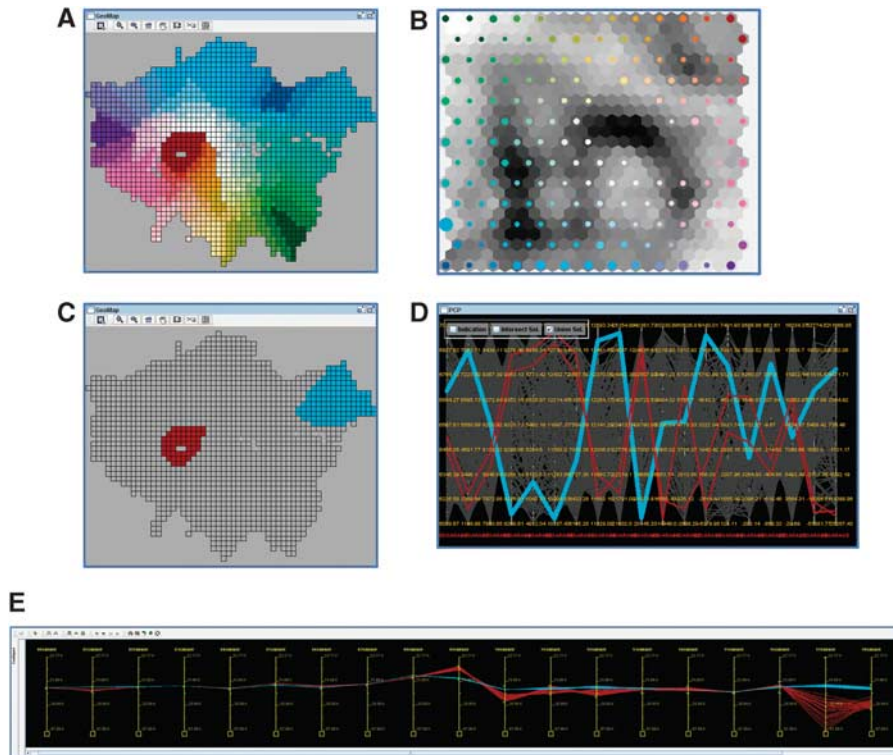


Figure 8 Using the combined statistical–exploratory method on a real data set of hedonic models of house prices in Greater London Area. Visualisations show the spatio-temporal GWR estimates for one regression parameter, the presence of the garage. (A) shows the map of the 1 km² grid covering the Greater London area and (B) the SOM clustering. Colours in the map and other visualisations are inherited from the SOM. A selection of two areas in central and eastern London is shown (C) on the map, (D) in the SOM PCP and (E) in the temporal PCP.

forms of modelling. Indeed, the concept of geographical weighting can be applied to any form of statistical methodology employed to analyse spatial data.

Here we looked for three types of patterns related to the interpretation of the results of spatio-temporal GWR modelling of hedonic pricing of houses in London.² These three types of patterns (areas of temporal stability in each separate regression parameter, areas of similar temporal behaviour in one or more parameters and groups of similar parameters) were successfully identified. The results answer the main question of the paper affirmatively: the combination of GWR and visual exploration is successful in identifying patterns in the spatio-temporal data set of parameter estimate surfaces $\hat{\beta}_j(x, y, t)$ that exist in the original parameter surfaces $\beta_j(x, y, t)$.

Since the experiment was performed on a simulated data set, these interpretations do not have any real meaning. When working with real data however, the patterns of local parameter estimates describe geographical processes that generate the data. Successful identification of these patterns in the GWR results through visual exploration should therefore facilitate interpretation of the spatio-temporal trends and the influence of background geographical processes on the phenomenon reflected in the data.

One of the issues of the combined statistical–exploratory methodology is to find a combination of visualisations and computational data mining methods that is most appropriate for the interpretation of spatio-temporal GWR or other geostatistical methods. Our experiment recycled an existing visual exploration methodology,⁴ but further work is required to determine if a different visualisation framework would be more suited to facilitate the interpretation of the results of spatio-temporal GWR. This is related not only to the chosen geostatistical methodology but also to the required exploration tasks, which in turn depend on the spatio-temporal geographical phenomenon that the combined methodology is intended to investigate and explain. It is therefore necessary to consult a typology of tasks for spatio-temporal exploration (such as for example Andrienko *et al.*²⁵ or Andrienko and Andrienko)²⁶ to select and/or develop the most appropriate visualisation methodology.

The study reported here uses simulated data that has the advantage that we can examine the accuracy of the patterns of the parameter estimates uncovered through the combination of GWR and Geovisual Analytics because we know the real values. Such validation would be impossible in the real world where we can only guess at the real processes. While simulated data are therefore

essential to this experiment, a question mark on their use exists in that how well will the methodology work in the real world where the data and processes are 'noisy'? To examine this issue briefly, we return to the motivation for our experiment which was a spatio-temporal analysis of house prices in London. GWR was recently used to investigate spatial variations in house price determinants across London separately for the years between 1980 and 1998.² We constructed a spatio-temporal data set from these GWR analyses in the same way as the simulated spatio-temporal data set for the experiment presented in this paper and performed a short preliminary visual exploration.²⁷ The spatial base for the analysis of London housing price market was a 1 km grid, covering the area of Greater London. As with the simulated data set presented in this paper, we interpolated the results of 19 separate GWR analyses into this grid – one analysis for each of the years 1980–1998 – and merged the results into one large spatio-temporal data set. Each GWR was run on a sample of property transactions, where the house price was regressed on the following variables:² floor area, time of construction of the property, property type, presence of garage, central heating, two or more bathrooms, and two area-related variables: percentage of professionals and percentage of unemployment in the census area where each property was located. Based on the outcome of the experiment presented in this paper, we now know how to look for visual evidence for phenomena that have influenced the housing market in London.

As an illustration, Figure 8 shows an attempt to identify patterns in the time series of GWR estimates for one particular regression parameter – the presence of a garage. This corresponds to approach I from the experiment on simulated data, presented above. We first produced a SOM clustering (Figure 8B), which shows several clusters (light areas) separated with dark areas. There are two clusters in the SOM that are very different from each other: the blue group in the lower left corner and the red group in the upper right corner. The spatial positions of these two clusters can be investigated in the map (Figure 8A) where we can see that the elements forming these clusters are also spatially adjacent. This pattern is even more noticeable in Figure 8C, which shows the map with a selection of only these two clusters. The observed pattern corresponds to the type 2a from our simulated experiment, that is, each of these two clusters is an area where the parameter estimate under investigation behaves similarly through time. Figures 8D and 8E show the same selection in the SOM PCP and the temporal PCP, respectively. As with the simulated data above, these two PCPs show that the behaviours of these two clusters are also very different from each other, that is, there are two completely different temporal trends at work in these two areas. Combining this observation with the actual locations of these two clusters in London gives an indication of the influence of this particular parameter on house prices in two different socio-economic

environments – the affluent area of central London and a more deprived area in eastern London.

This brief example shows how the combined statistical-exploratory methodology tested here can be used to identify relatively homogeneous housing sub-markets in London where processes are generally similar and change in a similar manner over time. We plan to perform a full analysis of the spatio-temporal GWR hedonic model for London housing market according to the exploration framework presented in this paper. The idea is to look for evidence for specific known events in the housing market such as for example the collapse of the market in 1990. We also expect to be able to identify areas with specific social processes such as gentrification or studentification. However, it is only through a fairly rigorous test of the methodology in a controlled experiment, as demonstrated here, that we can have confidence in the interpretation of our results when we apply the methodology to real-world data. Hence, the paper represents an important step in future geovisual analytical interpretations of GWR results.

Acknowledgements

Research presented in this paper was funded by a Research Professorship grant (03/RPI/1382) and a Strategic Research Cluster grant (07/SRC/I1168) both awarded to Prof. Fotheringham by Science Foundation Ireland under the National Development Plan. The authors gratefully acknowledge this support. The authors would additionally like to thank Ricardo Crespo from the National Centre of Geocomputation for help with the GWR processing. Thanks also to the anonymous reviewers for providing comments and suggestions that helped to improve this paper.

References

- 1 Fotheringham AS, Brunsdon C, Charlton M. *Geographically Weighted Regression – the Analysis of Spatially Varying Relationships*. John Wiley & Sons Inc.: New York, 2002.
- 2 Crespo R, Fotheringham AS, Charlton M. Application of geographically weighted regression to a 19-year set of house price data in London to calibrate local hedonic price models. In: *Proceedings of the 9th International Conference on Geocomputation 2007* (Maynooth, Ireland), National University of Ireland Maynooth: Maynooth, Ireland, 2007.
- 3 Fotheringham AS, Brunsdon C, Charlton M. *Quantitative Geography – Perspectives on Spatial Data Analysis*. Sage Publications: Beverly Hills, CA, 2000.
- 4 Demšar U, Fotheringham AS, Charlton M. Combining geovisual analytics with spatial statistics: the example of geographically weighted regression. *The Cartographic Journal* 2008; **45**: 182–192.
- 5 Unwin A, Unwin D. Exploratory spatial data analysis with local statistics. *The Statistician* 1998; **47**: 415–421.
- 6 Keim DA, Ward M. Visualization. In: Berthold M, Hand DJ (Eds). *Intelligent Data Analysis*. 2nd edn, Springer Verlag: New York, 2003; 403–428.
- 7 MacEachren AM. Invited paper on Geovisual Analytics (in preparation). *Information Visualization* 2008; forthcoming.
- 8 National Visualization and Analytics Center (NVAC), Illuminating the path: Creating the R&D agenda for visual analytics [WWW document], 2005. <http://nvac.pnl.gov/agenda.stm> (accessed 2 April 2008).

- 9 MacEachren AM, Wachowicz M, Edsall R, Haug D, Masters R. Constructing knowledge from multivariate spatio-temporal data: integrating geographical visualization with knowledge discovery in database methods. *International Journal of Geographic Information Science* 1999; **13**: 311–334.
- 10 MacEachren AM, Kraak M-J. Research challenges in geo-visualization. *Cartography and Geographic Information Science* 2001; **28**: 3–12.
- 11 Fotheringham AS, Demšar U, Charlton M. Performance of Geographically Weighted Regression (GWR) on a simulated spatio-temporal data set. Submitted to *Geographical Analysis*. In preparation, May 2008.
- 12 Gahegan M, Takatsuka M, Wheeler M, Hardisty F. Introducing Geo-VISTA Studio: an integrated suite of visualization and computational methods for exploration and knowledge construction in geography. *Computers, Environment and Urban Systems* 2002; **26**: 267–292.
- 13 Takatsuka M, Gahegan M. GeoVISTA Studio: a codeless visual programming environment for geoscientific data analysis and visualization. *Computers & Geosciences* 2002; **28**: 1131–1144.
- 14 Kohonen T. *Self-Organizing Maps*. 2nd edn. Springer Verlag: New York, 1997.
- 15 Vesanto J. SOM-based data visualization methods. *Intelligent Data Analysis* 1999; **3**: 111–126.
- 16 Bação F, Lobo V, Painho M. The self-organizing map, the Geo-SOM and relevant variants for geosciences. *Computers & Geosciences* 2005; **31**: 155–163.
- 17 Koua EL, Kraak M-J. Alternative visualization of large geospatial data sets. *The Cartographic Journal* 2004; **41**: 217–228.
- 18 Guo D, Gahegan M, MacEachren AM, Zhou B. Multivariate analysis and geovisualization with an integrated geographic knowledge discovery approach. *Cartography and Geographic Information Science* 2005; **32**: 113–132.
- 19 Skupin A, Hagelman R. Visualizing demographic trajectories with self-organizing maps. *Geoinformatica* 2005; **9**: 159–179.
- 20 Demšar U. Knowledge discovery in environmental sciences: visual and automatic data mining for radon problems in groundwater. *Transactions in GIS* 2007; **11**: 255–281.
- 21 Špatenková O, Demšar U, Kršp JM. Self-organising maps for exploration of spatio-temporal emergency response data. In: *Proceedings of the 9th International Conference on Geocomputation 2007* (Maynooth, Ireland), National University of Ireland Maynooth: Maynooth, Ireland, 2007.
- 22 Guo D. Coordinating computational and visual approaches for interactive feature selection and multivariate clustering. *Information Visualization* 2003; **2**: 232–246.
- 23 Inselberg A. Visualization and data mining of high-dimensional data. *Chemometrics and Intelligent Laboratory Systems* 2002; **60**: 147–159.
- 24 Edsall RM. The parallel coordinate plot in action: design and use for geographic visualization. *Computational Statistics and Data Analysis* 2003; **43**: 605–619.
- 25 Andrienko N, Andrienko G, Gatalsky P. Exploratory spatio-temporal visualization: an analytical review. *Journal of Visual Languages and Computing* 2003; **14**: 503–541.
- 26 Andrienko N, Andrienko G. *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. Springer Verlag: New York, 2005.
- 27 Demšar U, Fotheringham AS, Charlton M. Combining geographically weighted regression and geovisual analytics to investigate temporal variations in house price determinants across London in the period 1980–1998. In: *ICA Visualisation Workshop at the AGILE 2008 conference, 'GeoVisualization of Dynamics, Movement and Change' 2008* (Girona, Spain), ICA Commission on Geovisualisation, 2008.