The use of Kohonen mapping for the elucidation of space-time trajectories of multiple parameters: potential applications in fluvial geomorphology.

Dan Weaver** and Nick Mount*

*School of Geography, University of Nottingham, Nottingham NG7 2RD
Tel. +44(0)115 951 5428 Fax. +44 (0)115 951 5249
nick.mount@nottingham.ac.uk

***danofweaver@googlemail.com*

1. Introduction

The Earth's large, dynamic river systems remain a major geo-hazard both in terms of flooding and erosion. For example, erosion of the Brahmaputra River in Assam, India, has been responsible for more the 400,000 hectares of land destruction since 1955, with more than 100,000 families displaced (Assam Disaster Management Cell, n.d.). Despite the clear need for improved prediction of the dynamics of such channels they remain poorly understood. Within the disciplines of geomorphology and hydrology approaches to predicting complex channel processes continue to focus on the development of deterministic techniques including computational and physical models developed for individual rivers (e.g. Zanichelli *et al*., 2004) or highly specific channel settings. They are commonly three-dimensional, employ highly-parameterised, reductionist approaches, are difficult to apply outside of the specific river settings for which they were developed and are difficult to scale up for regional or continental-scale application. Consequently it is difficult to imagine them as the key to unlocking successful and rapid prediction of large-scale channel dynamism.

Understanding river channel processes at larger scales requires engagement with available spatial data sets. Analysis of planimetric channel parameters from remotely sensed imagery (Mount *et al.*, 2003; Sarma, 2005) or the estimation of hydraulic parameters from digital elevation, digital surface and terrain models (e.g. Pistocchi and Pennington, 2006) remain popular and established techniques. However many additional parameters known to exert major control on channel dynamism can now be estimated at large scales via analysis of remotely-sensed imagery, although exactly how these parameters combine to control dynamism is still poorly understood. These include land use changes and land cover classes (Akbari *et al*., 2006 Boucher *et al*., 2006), inundation histories (Jain *et al*., 2006), riverbed aggradation (Fan *et al*., 2006), bathymetric measurements (Carbonneau *et al*., 2006) and riparian vegetation (Goetz, 2006). The result is a large-scale, multi-parameter analytical environment, in which techniques are required that are capable of integrating and analysing the changing spatial and temporal patterns within parameters deemed to be important in controlling channel dynamism. Should such techniques be applied, an analytical approach more suited to predicting channel dynamism at large-scales may emerge.

In this paper, spatio-temporal self-organising maps, an extension of the standard Kohonen mapping clustering technique (Kohonen, 1990, 1995), are employed to quantify space-time trajectories amongst the values of multiple parameters in some simple river channel datasets. Once computed, the space-time trajectories associated with locations known to have undergone a particular dynamic process can be compared to emerging trajectories across the entire spatial extents of the analysis. Similar trajectories, indicating a high likelihood of the given processes occurring in new locations can then be identified and isolated. The algorithmic processes outlined in this paper are those employed in GISTSOM, a software package developed by the authors for the definition and comparison of space-time trajectories in multi-parameter, spatio-temporal data.

2. Kohonen mapping of multi-dimensional data

A Kohonen map, commonly known as a self-organising map, is an unsupervised neural network capable of representing high-dimensional data in a low-dimensional form through multi-dimensional clustering. Commonly, the output from a Kohonen map is a 2-dimensional array in which there are a predefined number of elements (e.g. 10 x 10) to which input data are assigned. Importantly, it spatially organises input data according to similarities in the values of the multiple dimensions at given locations, so that locations exhibiting similar values in their associated dimensions are represented proximal to one another and those exhibiting dissimilar values are represented distal to one another.

The classic explanation of Kohonen maps exemplifies the organisation of the three dimensions Red, Green and Blue (RGB) associated with image pixels (the sample data) into a two-dimensional output array, such that similar colours are clustered together according to the values in each of the three RGB dimensions. The individual RGB values are viewed as weights which, together, form a weight vector describing the location of the colour in the three-dimensional, RGB space. In the example given in Figure 1, the Kohonen map algorithm will cluster those pixels with the most similar weight vectors (i.e. the most similar colours) together within a two-dimensional array whilst maintaining the maximum possible distance between those pixels with dissimilar weight vectors.

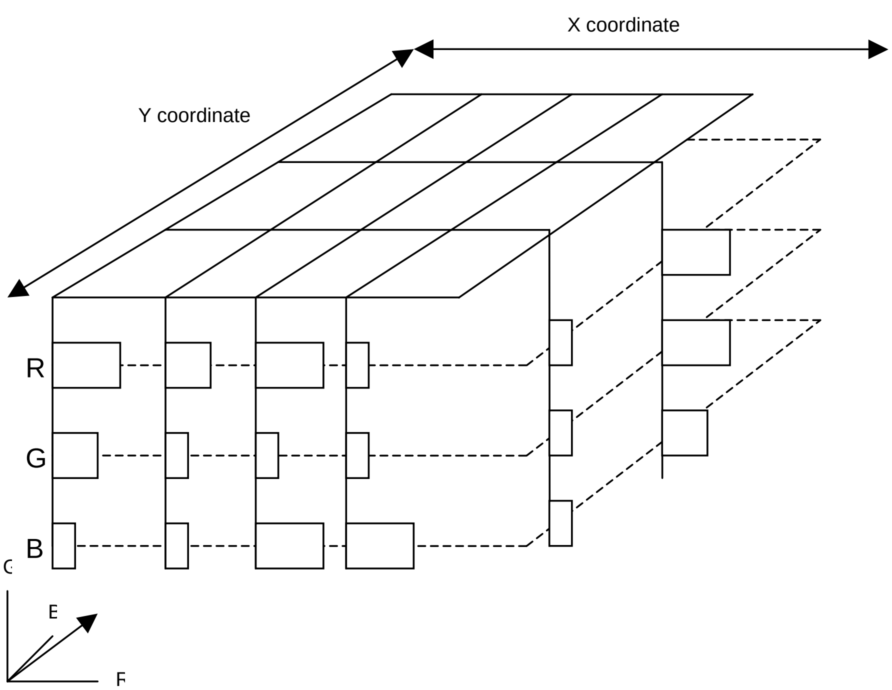

Figure 1. A pixelated array of RGB values (i.e. an image). Each pixel is an input sample to the Kohonen mapping algorithm, with the RGB values forming a weight vector that can be compared to the weight vectors assigned to each element of the Kohonen map. The RGB weight vector for the bottom left pixel is shown schematically.

The Kohonen map clustering algorithm firstly initialises an array of elements (often termed neurons), and assigns weight vectors (most often randomly) to each element of the array according to the dimensionality of the input weight vectors. The distance between each element's weight vector and the weight vectors of the sample datum is then computed according to equation 1.

$$dist = \sqrt{\sum_{i=0}^{n} x_i^{\,2}}$$

(equation 1)

where,

x_i is the value at the *ith* dimension of a sample

n is the number of dimensions to the sample data

The element with the most similar weight vector (i.e. smallest distance) is then termed the best matching unit (BMU) and is made more similar to the sample weight vector according to a learning function which controls the magnitude of modification. Neighbouring elements to the BMU are also made more similar, but to a lesser degree as defined by a neighbourhood function. The key attribute of this process is that the further away an element is from the BMU, the less it learns to be similar to the sample data. The algorithm then iterates, randomly selecting a sample datum each time and identifying the BMU. Crucially, as the number of iterations increases, the size of the neighbourhood around the BMU declines and the learning function decreases. In this way stability in the mapping of the input weight vectors to the weight vectors contained in the output array can be attained. In the case of an RGB image, similar colours will cluster together in the output array and clusters of dissimilar colours will be located distally in the array. The result is a mapping of multidimensional pixel values from the original image to the output array (Figure 2.)

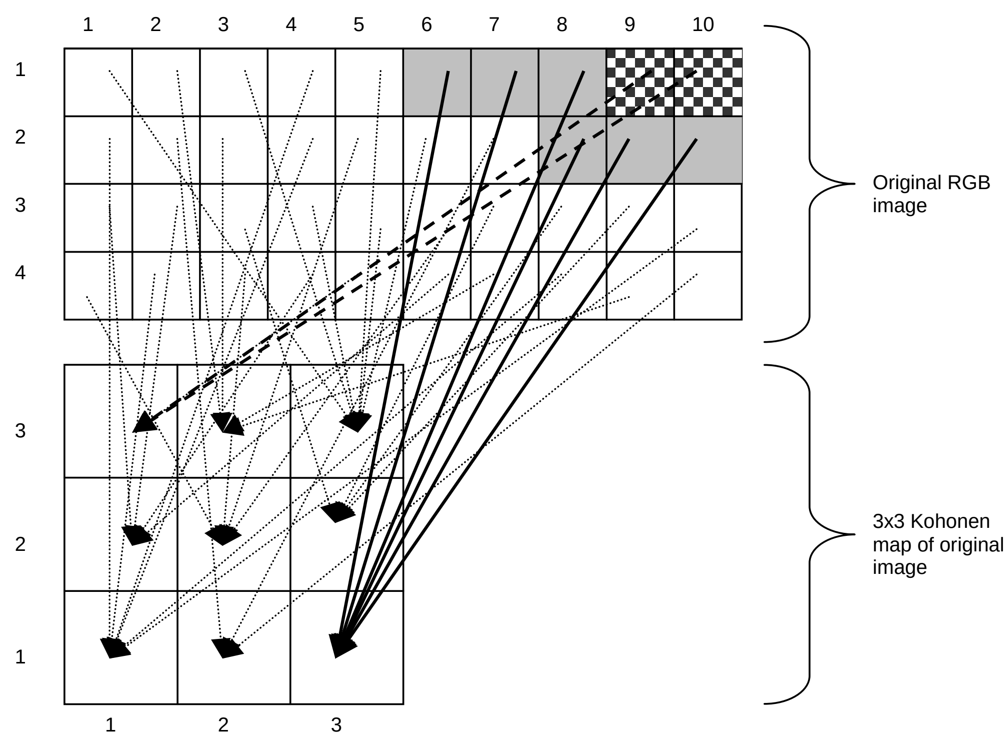

Figure 2. The mapping of a 10 x 4 image onto a 3 x 3 Kohonen map. Pixels (6,1), (7,1), (8,1), (8,2), (9,2) and (10,2) have all mapped to element (3,1) indicating they are all of similar colour. Similarly, pixels (9,1) and (10,1) have both mapped to element (1,3) indicating that

they are of similar colour. These two colour groupings have mapped to elements a long way apart on the Kohonen map indicating that the two groups of pixels contain very different colours.

3. Extending the Kohonen mapping algorithm across time.

The standard Kohonen mapping algorithm described above is applied to sample data collected at one point in time. However, it can be *repeated* on data existing across several time periods, an extension for which GISTSOM has specifically been developed. In GISTSOM, the sample data weight vectors used in the training of the initialized SOM are taken randomly from across *all* of the time periods for which data are available. In this way, the weight vectors against which the Kohonen map learns implicitly include a temporal dimension, and a spatio-temporal Kohonen map evolves.

By comparing the locations in the spatio-temporal Kohonen map to which input data for each time period have been mapped, a set of coordinates describing the trajectory of the sample data across the Kohonen map through time can be extracted (Figure 3). These space-time trajectories represent the changes in the multiple dimensions of each sampled datum (i.e. pixel in an image) through time. By comparing the space-time trajectory coordinates for all of the input sample data, similar trajectories can be identified. In the case of GISTSOM a simple coordinate tolerance measure has been used as a measure of similarlity. Sample data with similar space-time trajectories can then be identified as samples whose multi-dimensional values have changed in a similar manner through time. In the case of multiple RGB images, samples with similar space-time trajectories would have experience similar colour changes through time.

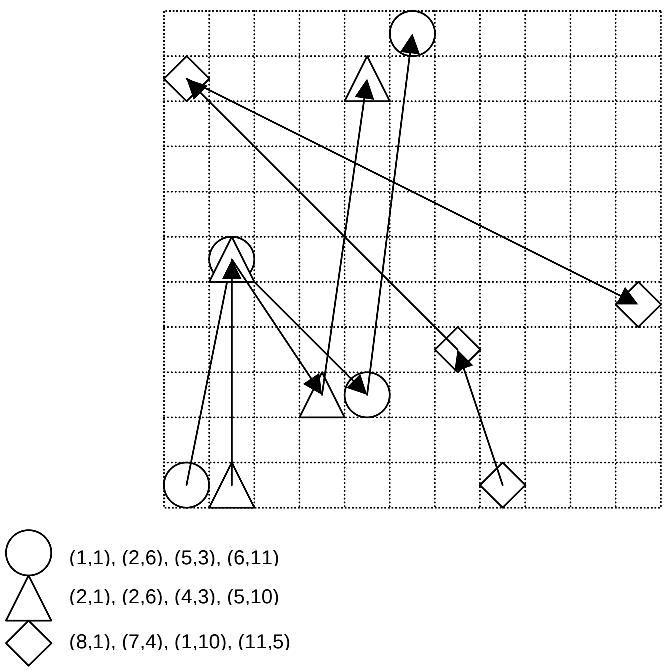

Figure 3. Space time trajectories for three samples in an 11x11 Kohonen map. Clearly the trajectories of the circle and triangle data are very similar indicating that all of their multiple parameter values have responded in very similar manner through time. By contrast, the trajectory of the diamond data is very different indicating a very different response in its multiple parameters through time.

4. Beyond RGB: Kohonen maps in dynamic river channels

Moving beyond the example of RGB values, the dimensions within the sample data of a Kohonen map can be representative of any phenomena. In the case of understanding dynamic river channels, the sample data at a given point in time can be represented by any number of rasters, each defining the values of an individual parameter considered an important control of river dynamism (e.g. roughness, bed load size, vegetation type, land use type, inundation histories etc.) over a given spatial extent. These data can be held in multi-dimensional arrays, sampled, and mapped to the spatio-temporal Kohonen map space. By repeating the mapping for additional time periods, the space-time trajectories associated with each cell in the data's spatial extent can be defined.

According to the analysis outline above, space time trajectories for a cell known to have experienced a given form of channel dynamism can be extracted from available spatio-temporal datasets of relevant parameters recorded *prior* to the occurrence of the dynamism. In this way, the space time trajectory for the multiple parameters which is *predictive* of the given dynamism can be computed. It then follows that by searching for the occurrence (or it's apparent development) of this space-time trajectory throughout contemporary data, cells in which multi-parameter responses are indicative of the given dynamism can be isolated.

5. GISTSOM exemplar

The graphical user interface of GISTSOM during an example analysis of panchromatic river channel data over three temporal periods is given in figure 4. GISTSOM provides both a visual and numerical output of the space-time trajectories associated with any cell in the input raster data set. It allows user selection of a given trajectory for any cell and the selection of all cells displaying similar trajectories according to a tolerance value. Data may be output to any proprietary GIS for further analysis (Figure 5).

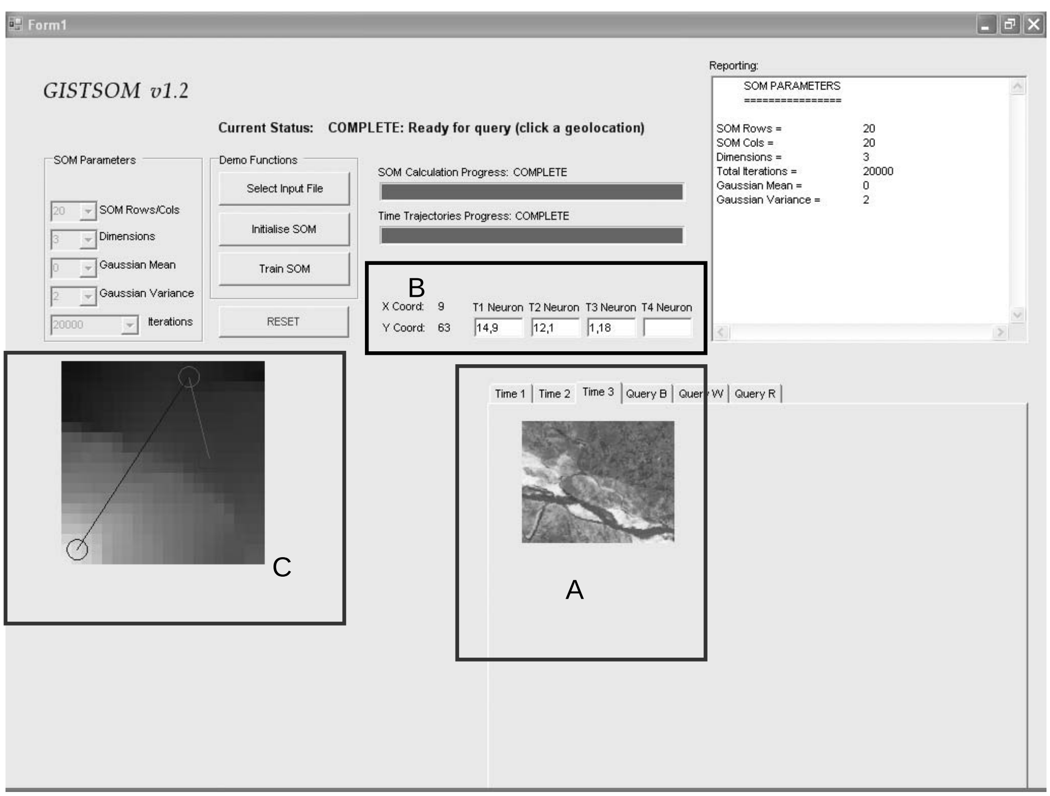

Figure 4. The GISTSOM graphical user interface (GUI). Three parameters of the input data for each time period are displayed using the RGB colour guns (A) allowing input data visualisation. Following the initialisation and training of the Kohonen map, interaction with A allows the space-time trajectory for any cell of the input rasters to be displayed as coordinates (B) or visually in the Kohonen map (C).

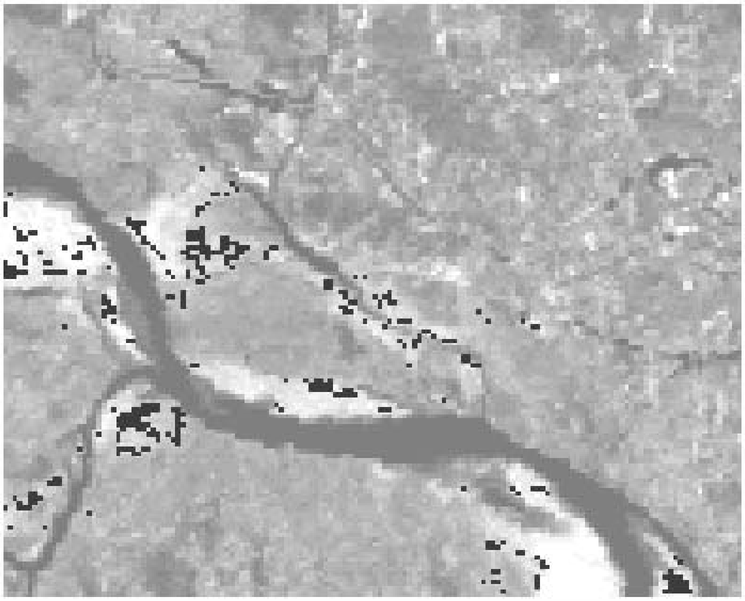

Figure 5. GISTSOM output. Black cells indicate the locations of cells which display a similar space-time trajectory amongst the panchromatic parameters to that selected within the GISTSOM GUI.

6. Discussion

There is little doubt that the ability to analyse the multi-parameter data sets relevant to the prediction of river channel dynamism through space and time represents a significant advance over the present analytical procedures employed by fluvial geomorphologists. These often rely on the visual interpretation of data, contain an over emphasis on either the spatial, or temporal dimensions contained within the data, and are not automated. Consequently they remain restricted in their ability to predict channel dynamism at large spatial and temporal scales. The analysis of Kohonen map space-time trajectories offers a solution to these issues, affording equal emphasis to space and time, removing the need for visual interpretation and analysis of complex data and automating the analytical process.

However, the procedures are dependent on the availability of raster data sets encompassing all of the parameters which drive channel dynamism. For many important parameters good spatio-temporal data are available (e.g. land cover mapping and inundation histories from Landsat TM imagery), but for others, particularly those reliant on 'flown-for-purpose' data, data availability may be a significant restriction to the application of the technique. In addition, during high river stages it is likely that increased inundation area will limit the completeness of the data records for many important parameters, falsely indicating major changes in key parameters. For example, a highly vegetated area at time t_1 which is inundated at time t_2 will appear as 'losing' its vegetation under the inundated area. Therefore, enormous care is needed in selecting data for use in analysing space-time trajectories.

7. Conclusions

The quantification of space-time trajectories of spatio-temporal Kohonen maps offers an exciting development for those analysing dynamic systems driven by changes in multiple parameters in both space and time. It offers automation of frequently manual analyses and removes bias in the analysis of the spatial or temporal dimensions. Importantly, it offers a method for the prediction of locations of future dynamism, by analysing the patterns of spatio-temporal change in the parameters governing the dynamism at locations where it has already been observed. However, the technique is dependent on the availability of raster data sets which fully describe the governing parameters through space and time, and this may limit its application to certain problems.

References

Akbari, M., Mamanpoush, A.R., Gieske, A. (2006). Crop and land cover classification in Iran using Landsat 7 imagery. International Journal of Remote Sensing, 27 (19), 4117-4135.

Assam Disaster Management Cell (n.d.). Assam Disaster Management. Accessed online at: *www.sdmassam.nic.in*. [Accessed 12.12.06].

Boucher, A., Seto, K.C., Journel, A.G. (2006). A novel method for mapping land cover changes: incorporating time and space with geostatistics. IEEE Transactions on Geoscience and Remote Sensing, 44 (11), 3427-3435

Carbonneau, P.E., Lane, S.N., Bergeron, N. (2006). Feature based image processing methods applied to bathymetric measurements from airborne remote sensing in fluvial environments. Earth Surface Processes and Landforms, 31 (11), 1413-1423.

Fan, H., Huang, H.J., Zeng, T.Q. (2006). River mouth bar formation, riverbed aggradation and channel migration in the modern Huanghe (Yellow) River delta, China. Geomorphology, 74 (1-4), 124-136.

Goetz, S.J. (2006). Remote sensing of riparian buffers: Past progress and future prospects Journal of the American Water Resources Association. 42 (1), 133-143.

Jain, S.K., Saraf, A.K., Goswami, A. (2006). Flood inundation mapping using NOAA AVHRR data. Water Resources Management, 20 (6), 949-959.

Kohonen, T. (1990). The self-organizing map. Proceedings of the IEEE, 78 (9), 1464-1480.

Kohonen, T. (1995). Self-organizing maps. Springer, Berlin.

Mount, N.J., Louis, J., Teeuw, R.M., Zukowskyj, P., Stott, T. (2003). Estimation of error in bankfull width comparisons from temporally sequenced raw and corrected aerial photographs. Geomorphology, 56, 65-77.

Pistocchi, A., Pennington, D. (2006). European hydraulic geometries for continental SCALE environmental modelling. Journal of Hydrology, 329 (3-4), 553-567.

Sarma, J. (2005). Fluvial processes and morphology of the Brahmaputra in Assam, India. Geomorphology, 70 (3-4), 226-256.

Zanichelli, G., Caroni, E., Fiorotto, V. (2004). River bifurcation analysis by physical and numerical modelling. Journal of Hydraulic Engineering -ASCE 130 (3), 237-242.