A neural network based on time series for spatiotemporal relationships prediction

Hana Alouaoui*, Sami Yassine Turki and Sami Faïz

LTSIRS Laboratory of Remote Sensing and Information Systems with Spatial Reference,
ENIT Notianal Engineering School of Tunis,
Tunis El Manar University, Tunisia
Email: aloioui_hana@yahoo.fr
Email: yassine.turki@isetub.rnu.tn
Email: sami.faiz@insat.rnu.tn
*Corresponding author

Abstract: The study of spatial objects and their evolution is a complex process. Most of the techniques used in this field do not consider the evolution of spatiotemporal relationships. The present paper proposes a new approach for the prediction of the future behaviour of spatiotemporal relationships based on spatiotemporal association rules. Such rules demonstrate the evolution of spatial objects and the influence of the spatial distribution of adjacent-areas relationships. We use a predictive neural network based on a nonlinear time series technique to generate spatiotemporal predictive rules. The learning examples correspond to spatiotemporal association rules and the results are in the form of spatiotemporal predictive rules assessing the future spatiotemporal relationships. These relationships can be used to inform about upcoming risks. We conduct an experimentation using a time series of satellite images, describing Megrine zone in the southern coast of Tunis (Tunisia). As a final result, we obtain spatiotemporal predictive rules describing the spatiotemporal relationships evolution between anterior and future dates. A comparison between the predicted values and the ground truth shows good correspondence rates varying between 78% and 90%.

Keywords: spatiotemporal data mining; spatiotemporal association rules; STAR; spatiotemporal relationships prediction; neural network based on time series; natural risk prediction.

Reference to this paper should be made as follows: Alouaoui, H., Turki, S.Y. and Faïz, S. (2016) 'A neural network based on time series for spatiotemporal relationships prediction', *Int. J. Spatial, Temporal and Multimedia Information Systems*, Vol. 1, No. 1, pp.63–86.

Biographical notes: Hana Alouaoui is an Assistant at the University of Jendouba Tunisia, a PhD student in Computer Science and a Researcher at the LTSIRS: Laboratory of Remote Sensing and Information Systems with Spatial Reference-National School of Engineers, University of Tunis El Manar. Her research interests include spatiotemporal knowledge discovery especially mining spatiotemporal association rules in order to predict future risks.

Sami Yassine Turki is an Assistant Professor at the University of Carthage in Tunis and Researcher at the LTSIRS: Laboratory of Remote Sensing and Information Systems with Spatial Reference-National School of Engineers, University of Tunis El Manar. His researches focus on the use of GIS in urban studies and the application of spatial data mining techniques in order to produce urban land use maps.

Sami Faïz is a Professor in Computer Science at the Higher Institute of Multimedia Arts of Manouba and member of the LTSIRS: Laboratory of Remote Sensing and Information Systems with Spatial Reference. His teaching and research interests are in the areas of geographical databases, geographical information systems, spatial data mining and spatiotemporal data mining. He is also the founder of many national and international projects in geomatics.

1 Introduction

The study of spatial objects evolution is a complex process based on a variety of techniques. This impetus is reflected in several areas including the acquisition of spatial data, storage, and analysis for decision support. Subsequently, various studies have come to the fore. They aim to cope with factors leading to the risk occurrence (Isik et al., 2013) and to fulfil tasks related to the prevention of natural risks.

However, the risks related to the evolution of geographical objects remain of great concern and need a big interest from the scientific community. Additionally, most of the developed techniques used for the study of spatial objects evolution do not consider the evolution of spatiotemporal relationships. Such relationships are not to be neglected, they can form a major issue in detecting and predicting natural risks.

In this paper, we study the improvements that data mining techniques can bear in this field, especially the risks related to land cover changes. In fact, we aim to generate spatiotemporal predictive rules (STPR) rich in information about future occurrence of spatiotemporal prohibited relationships. These relationships are considered as risks in our study. The learning examples on which we base the whole predictive process are spatiotemporal association rules (STAR). One of the numerous benefits deriving from the STAR mining is the possibility of finding some prohibited spatiotemporal relationships existing in the consequent of the rule. Hence, such a particular knowledge may be important in predicting upcoming risks.

Our predictive system is based on a neural network (NN) and time series analysis. STARs are processed by the NN as historical data. Then, future prediction is handled with a nonlinear autoregressive exogenous (NARX) time series due to the heterogeneity of inputs and the dynamics of the geographical entities.

This paper is organised as follows: Section 2 gives an overview on the methods proposed in the literature to resolve the natural risk prediction problem. In Section 3, we explain the time series principal and the time series prediction technique based on NN. In Section 4, we describe our method for STPR calculation. In Section 5, we explain our experimental study, discuss the obtained results and provide a comparison to other approaches. Finally, in Section 6, we sum up the main conclusions of this paper as well as point out directions for future works.

2 Modelling of land cover changes and prediction of natural risks

Numerous works are devoted to study future prediction related to land cover changes. In this section, we detail both approaches interested in modelling land cover changes and in predicting natural risks related to these changes. We also give the main limits of the mentioned works.

Various models have been developed to produce simulations of the land cover evolution. They are distinguished by their analysis approaches. Many of these models allocate aggregated area wide projections of population and jobs to subareas. These models are based on a variety of theories and paradigms. Among the numerous approaches in the literature, several ones are interested in land cover changes over time.

Statistical models are used in order to explicitly identify the causes of changes in land cover using multivariate analysis (Serneels and Lambin, 2001). These techniques include multiple logistic regressions and are applied in several case studies like deforestation, and forest fire prediction (Huang and Friedl, 2014).

Cellular automata are also applied to achieve this finality. They are based on a representation of the spatial environment by a grid of cells in which the modeller defines the rules of evolution on the basis of the spatial and temporal autocorrelation (Alcamo et al., 1996). The state of a cell (e.g., an agricultural parcel) at a time $t + 1$ depends on its state and its neighbourhood at time t. System dynamics are thus based on local interactions between neighbouring spatial entities.

Likewise, multi-agent systems (MAS) are used to model the land covers evolution. In Manson (2000), MAS are combined with other types of models in order to include spatial aspects in the modelling process and consequently simulate changes in land cover. Similarly, the proposed approach in Farah et al. (2006), seeks to analyse spatiotemporal geographical information. They use MAS in order to detect changes on multi-temporal satellite images.

Hidden Markov chains are also devoted to handle with land cover change modelling. In Derrode and Carincotte (2005), fuzzy hidden Markov chains detect changes between two synthetic aperture radar (SAR) images. In Essid et al. (2012), a methodology for integrating a new parameter measuring spatial relationships in the hidden Markov models (HMM) is proposed in order to detect, interpret and predict changes in urban areas. The proposed methodology is founded on generating change probabilities between classes referring to the rate changes in the past.

Several other applications are based on time series analysis. The approach presented in Nabel et al. (2014) shows that simulation driven by environmental time series have to take into consideration spatial correlations. The method is applied to spatial correlation in climatic fluctuations. The main idea in Pouliot et al. (2014) is the development of temporally consistent land cover time series (LCTS) from satellite-based earth observation. The detected changes are used to update maps. In Neeti and Eastman (2014), the concept of temporal concatenation of multiple image time series is introduced. Three sets of precipitation time series are used as a case study. The approach presented in Lyons et al. (2013), demonstrates the potential use of time-series analysis to investigate seagrass growth. Compared to traditional seagrass mapping and monitoring approaches, this method proves clear benefits of the use of time-series analysis of remotely sensed seagrass evolution. The approach described in Huang and Friedl (2014), uses distance metrics to measure the similarity between a pixel's annual time series of the same land cover class. In Dinga et al. (2014), a concept of mean length variability is proposed to compare the difference in spatial heterogeneity. Temporal changes in spatial heterogeneity are observed and they are a result of changes in the fraction of vegetation cover.

Most of the methods based on the detection of land cover changes are intended to natural risk prediction applications such as the predictability of catastrophic earthquakes (Molodenskii and Molodenskii, 2011), rainfall erositivity (Xin et al., 2011) and tornadoes.

An increasing number of geographical information systems (GIS) tools have also been established to integrate and analyse data in order to produce developable land units and to project their growth (Karatunga, 2005; Toledano et al., 2008; Singh et al., 2012). These models project future land cover patterns that are directly related to the input policies in order to estimate future land cover changes.

The learning capacities of NN are also explored. The developed model in Pijanowski et al. (2014) uses a GIS to manage spatial data and a NN which learns the spatial data patterns. The obtained model called land transformation model (LTM) simulates local scale patterns and is considered as land use land cover change (LULCC). In Tatem et al. (2002), the authors present an approach to predict spatial patterns of subpixel scale features from remotely sensed imagery. The method uses a NN that converges to a minimum of an energy function defined as a goal. This function incorporates prior information on land cover types. The proposed approach in Gazzaz et al. (2012), aims to predict water quality index in the Kinta River Malaysia. The used method is based on three layer perceptron. The best prediction values are obtained by the quick propagation training algorithm. In Isik et al. (2013), a NN is trained, validated and tested using streamflow data belonging to ten watersheds in Western Georgia, USA. Ten LULCC scenarios are developed by changing the LULCC percentages in the forested and urban test watersheds. The developed model predicts daily streamflows.

Most of the above-cited works deal with statistical perceptions and methods. No works have considered spatiotemporal relationships evolution as indicators of upcoming risks.

Several works consider simple features like texture, gray levels and perimeter to carry out the evolution of land covers. They do not allow to find out how objects evolve, recovering the history of their spatial relationships, and how they will be in a future time. In addition, there are no works describing the use of NN in a prediction finality based on time series method and in a spatiotemporal context. Particularly, in the case of spatiotemporal relationships between discreetly moving regions which is the addressed issue in this paper.

3 NN based on time series

In this section, we explain the time series prediction principle and the use of NN in this field.

3.1 Time series

A time series is a sequence of data depicted at subsequent points in time and spaced at uniform time intervals (Wei, 1990). This sequence of data has a natural temporal ordering.

Time series are used in many applications such as; statistics, signal processing, pattern recognition, econometrics, mathematical finance, weather forecasting, etc.

In Durante et al. (2014), financial time series are used in order to generate specific correlations. Two applications are presented about the use of a new measure for determining linkages in financial systems, and creating clusters of financial time series. The approach studied in Leasage and Krivelyova (1999), is dedicated to macroeconomic-modelling field and it aims to forecast the regional employment. The method is based on the analysis of the monthly employment time series from 1982 to 1995. In Loglisci et al. (2013), authors investigated the task of discovering evolution chains in dynamic networks. The proposed solution is based on the extraction of time-period stamped patterns. Their experiments are applied on social data.

These time series applications are purely static and cannot be used for a prediction finality. However, nonlinear time series modelling improves forecasts (Bowerman and O'Connel, 1993) and produces a richer notion of dynamics than linear time series models allow. Time series predicting makes use of the natural one-way ordering of time. Consequently, values for a given period will be derived from past values. Having time series {x[t – 2], x[t–1], x[t]}, the problem is how to estimate x at some future time x [t + s] (i.e., s is called the horizon of prediction).

In the case we want to predict just one time sample in the future (example: the next day, the next month, etc.), the horizon of prediction is simply equal to 1 (s = 1).

Table 1 Nonlinear time series methods

Method	*Equation*	*Explanation*
Nonlinear autoregressive with external (exogenous) input (NARX)	Y(t) = f(x(t-1),…, x(t-d), y(t-1),…, y(t-d))	The prediction of series y (t) is computed on the basis of d past values of y (t) and other series x (t).
The function f is some nonlinear function, such as a polynomial. It can be a neural network, a wavelet network or a sigmoid network.		
Nonlinear autoregressive (NAR)	Y(t) = f (y(t-1),…, y(t-d))	The prediction of series y (t) is computed on the basis of d past values of y (t).
Nonlinear input-output	Y(t) = f(x(t-1),…, x(t-d))	The prediction of series y (t) is computed on the basis of d past values of series x (t).

Whereas, in the case of predicting the next sample in a time series for a longer forecast, i.e., not x[t + 1] but x[t + s], the horizon of prediction will be s > 1. Here, we are faced with three options:

- Predict x[t + s] by training on {x[t – 2], x[t – 1], x[t]}
- Predict all x[t + i], $1 \le i \le s$ (if s is small).
- Predict x[t + 1] and iterate to get x[t + s] for any s.

Additionally, the nature of the inputs has a strong influence on the chosen method. Table 1 gives an explanation of the most known methods.

The next subsection is devoted to study the use of NN to handle with nonlinear behaviour of dynamical time series, especially, in the prediction task.

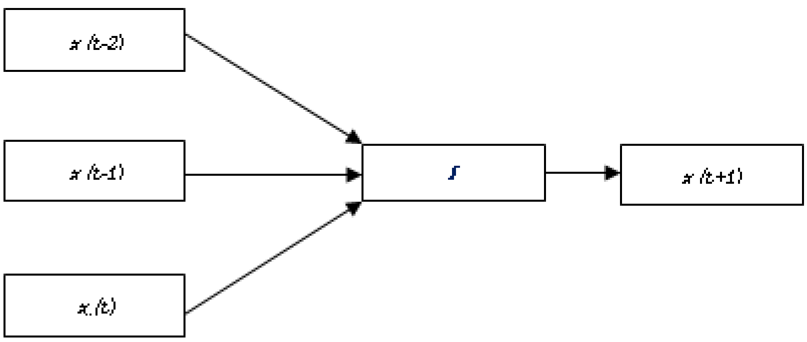
3.2 NNs for time series prediction

NN uses the analogy with the architecture of the human brain. It is composed of several interconnected neurons and a weight is associated with each arc. The input neurons are connected to the attributes of the concerned object (e.g., spatial object, parcel, region, etc.). The output neurons indicate the final value of the decision (e.g., the class of the object, its future state, etc.). The intermediate neurons are organised in layers and the whole constitutes a network. During the learning phase, the objects are presented to the NN and if the answer is not consistent with the decision-supervised, a back-propagation algorithm changes the attributes and the weight of intermediate neurons.

The standard NN method of performing time series prediction (Crone et al., 2011; Hon-lun et al., 2014) is based on the induction of the function f [see equation (1)] using any feed-forward function (Figure 1). NN general architecture uses a set of N-tuples as inputs and a single output as the target value of the network. This method is often called the sliding window technique as the N-tuple input slides over the full training set. Figure 1 gives the basic architecture.

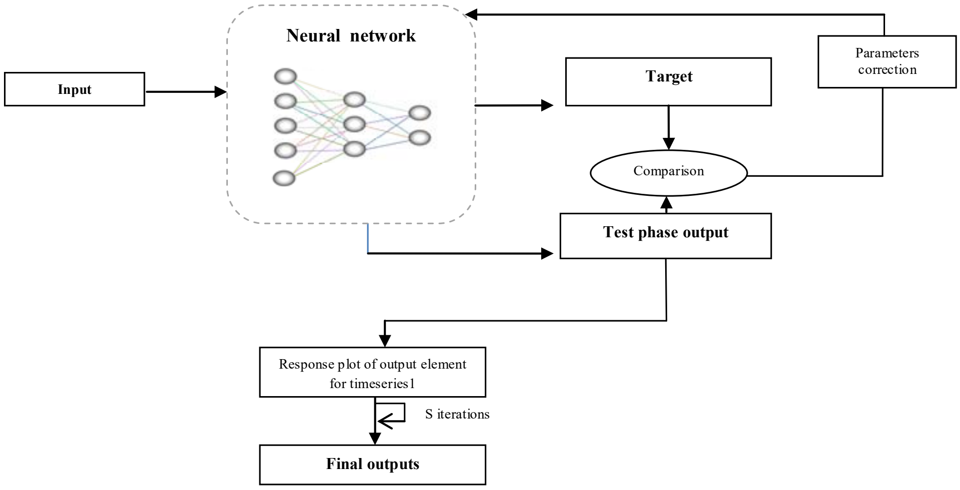
$$x[t+s] = f\left(x[t-2], x[t-1], x[t], \ldots\right) \quad (1)$$

Figure 1 The standard method of performing time series prediction using a sliding window



That architecture cannot handle with more complex prediction problems. In particular cases, we need more than one output. Here after, an architecture of a predictive NN (Figure 2) which aims to predict values on the basis of ‘the response plot of output element for Time series1’. The final outputs are obtained after s iterations.

Figure 2 Architecture of a predictive NN (see online version for colours)



4 The proposed approach

In this section, we present the problem formulation and the architecture of our proposed approach.

4.1 Problem formulation

At each time point ti, the spatiotemporal object is described by the set (ai, gi, pi) where a, g, p denote respectively the non-spatial attributes, geometrical attributes and spatial relations of the reference object X observed in ti.

A time period (or interval) {t1, …, tn} is a sequence of consecutive time points {ti, tj, …, tk} where (t1≤ ti, tj, tk ≤ tn).

{tn…tn+ ε …tm} is another sequence of consecutive time points where tn < tn + ε < tm.

The problem we intend to solve can be formalised as follows:

- *Given* the set of non-spatial attributes (ai) and geometrical attributes (gi) describing the spatial object X at consecutive time periods t1, …, tn.
- *Find* the set of spatiotemporal relationships (pi) relating the reference object X to other relevant spatial objects (Y, Z) in a future time periods tn + ε, …, tm.

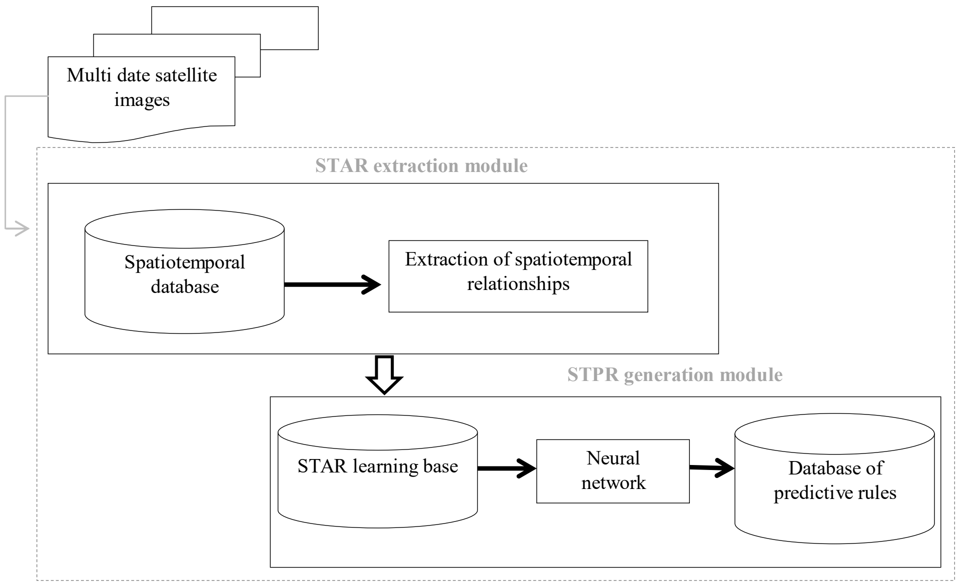
In order to find out the evolution of spatiotemporal relationships, we chose to adopt the association rules mining method (Koperski and Han, 1995; Appice et al., 2005). In our case, we intend to generate STARs that focus on the spatial relationships as object characteristic and show their evolution over the time. These relationships are considered as a structured knowledge, as opposed to other characteristics of a geographical object such as the colour, the texture, the shape, etc.

Our goal, therefore, is to present a methodology to monitor, interpret, and predict the land cover changes taking into account uniquely the spatiotemporal relationships.

4.2 Proposed architecture

The architecture of the proposed approach is based on two modules namely; STAR extraction module and STPR generation module as shown in Figure 3. Furthermore, we explain every module separately.

Figure 3 The proposed architecture



4.2.1 STAR extraction module

This module aims to compute spatiotemporal relationships which describe mutual influences between two objects (Malerba, 2008). This is held through spatial queries with time indication.

We start by specifying the reference object (X) on which the discovery will be achieved and the relevant task objects (having a spatial relationship with the reference object). The relationship can be a proximity relationship, an adjacency relationship or a containment relationship (Table 2).

Table 2 Spatial relationship types

Relationship	*Proximity relationship*	*Adjacency relationship*	*Containment relationship*
Examples	Close_to	Touches	Contains
	Very_close_to	Adjacent_to	Inside
		Intersects	covers

Here after, a query example used to compute the spatiotemporal relationship ‘Water_Very_Close_to_Urban_1987’ is given:

```
Select "name_Urban", "name_Water" as "Water_Very_Close_to_Urban_1987"
from "Urban" U, "Water" W
Where contains (buffer (U.the_geom, 500), W.the_geom)
And "W_time" between '1987-01-01' AND '1987-12-31'
And "U_time" between '1987-01-01' AND '1987-12-31'
```

Contains (buffer (X.the_geom, distance), Y.the_geom) is a predefined function used to compute the relationship on the basis of distance measurements. For example, two objects X and Y have a closeTo relationship if distance (X, Y) ≤ d. Where (d) is a user-specified distance.

The queries execution generates lots of relationships. Algorithmically, frequent itemsets describing frequent spatiotemporal relationships are discovered in each period by exploiting the well-known Apriori algorithm (Agrawal et al., 1993). This same algorithm generates the set of STARs on the basis of the frequent spatiotemporal relationships and referring to the support and confidence measures. The proposed formula of the STAR is given in equation (2):

$$\begin{aligned}&(X, ai, gi, pi, ti) \wedge (X, aj, gj, pj, tj) \wedge \ldots \wedge (X, an, gn, pn, tn)\\&\rightarrow (X, an+\varepsilon, gn+\varepsilon, pn+\varepsilon, tn+\varepsilon)\end{aligned} \tag{2}$$

where

X the reference object

(a) non-spatiotemporal attributes (e.g., name …)

(g) geometrical characteristics of X (e.g., geometry …)

(p) spatial relationships of X with other neighbouring objects

ti, tj, …, tn, tn + ε is a sequence of consecutive time points where (ti < tj, tn < tn+ ε).

4.2.2 STPR generation module

The object of this module is to predict future prohibited spatiotemporal relationships (Table 3). For example, the 'inside' relationship is prohibited and it reveals a metric distance equal to 0 between the urban and the water zones and can be considered as a risk threatening the urban zone.

The inputs of the STPR module are the STARs generated by the first module.

Table 3 Relationships classification

Relationship type	*Example*
Normal	Island_inside _Sea
Possible	Road_Cross_River
Prohibited	Road_inside_River

The proposed formula of the STPR is similar to that of the STAR except for the time of the conclusion (tm) which is unknown (we have not data describing X at that future time). In fact, we have no temporal constraint about the time (tm) and it can be projected in the future.

We propose to use a NN as predictive system because NN is extremely flexible to changes. It can be retrained and adapted to nonlinear problems related to dynamical systems.

When displayed to data, the network gains experience, learns from regularities in the past and sets its own rules.

In our case, it is more appropriate to opt for the nonlinear autoregressive with external (exogenous) input (NARX) (Table 1) since we have exogenous inputs, i.e., the spatiotemporal relationships at a date t2 are not similar to those of t1 and those of t3 do not match with the spatiotemporal relationships of t2.

In addition, the option 'Predict x[t + 1] and iterate to get x[t + s] for any s' (Section 3.1) seems to be more convenient and more adaptable to the NN architecture.

5 Experimentation

This section presents an experimental study of our approach on real data. It first describes the used datasets and the implementation details. The obtained results are then presented.

5.1 Experimental setup

Zones situated on the outskirts of Tunis Lake and Sebkhas of Ariana and Sejoumi are characterised by an important urban growth during the last decades. We use a set of SPOT 3 satellite images of these areas.

The first one (Figure 4) was taken in 1987 and the second one (Figure 5) in 2001, both have a resolution of 20 m and cover an area of 314.86 km^2.

These images are processed by a land cover classification (Figures 6 and 7) into five principal land cover classes, namely: water, agriculture, woodland, urban and bare_soil (Figure 8). We create a spatiotemporal database (STDB) containing five tables (urban, water, woodland, agriculture and Bare_Soil) corresponding to the five-land cover classes. The STDB is implemented with POSTGRESQL and its spatial extension POSTGIS used to handle with geometrical types. To extract the geometrical shapes and then define the geometrical coordinates of each shape we use OpenJUMP GIS. The execution of the first module generates a STAR set indicating spatial land changes affecting the geographical objects. A refinement phase is executed in order to obtain the final STAR set containing only the most interesting and meaningful rules. These rules correspond to the STAR learning base in Figure 3. Here after, an example of an interesting STAR:

urban_close_to_bare_soil_1987 = true	
urban_adjacent_to_water_1987 = true	
→ urban_inside_water_2001 = true	
urban_very_close_to_bare_soil_2001 = true	*conf= (0.9)*

This rule indicates that if an urban zone is close to a bare soil in 1987 and adjacent_to a water source in 1987 then it will be very_close_to the bare soil and inside the water source in 2001. This evolving proximity of the urban zone to the water source is described by the urban expansion and the increasing number of buildings in the zone near to the water source. The 'inside' relationship is prohibited (Table 3) and can be considered as a risk menacing the urban zone.

In the second phase of our experimentation, we build a STAR matrix containing the learning samples going to be processed by the NN. Each relationship of the rule antecedent or the rule conclusion is codified and it takes one among two possible values 0 or 1 (Figure 9), (The value 0 means that the relationship is not verified in the antecedent or the conclusion of the rule. The value 1 means that the relationship is verified in the antecedent or the conclusion of the rule).

We conduct an experiment using a NN under the MATLAB platform. We simulate the NN using a set of 160 different values. For each particular type of relationship we take ten different instances (Figure 9).

Figure 4 Spot image of the study zone in 1987 (see online version for colours)

Figure 5 Spot image of the study zone in 2001 (see online version for colours)

Figure 6 Classified image of the study zone in 1987 (see online version for colours)

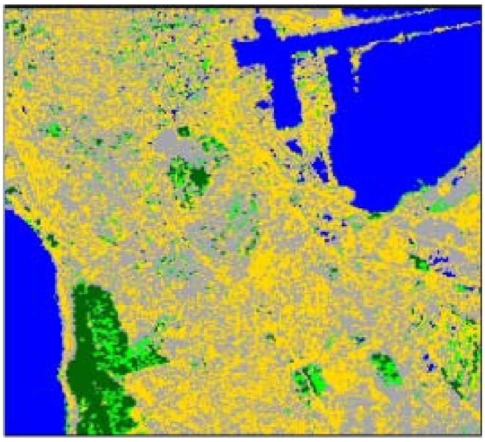


Figure 7 Classified image of the study zone in 2001 (see online version for colours)

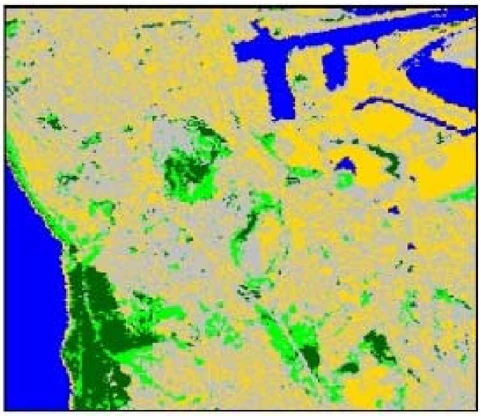


Figure 8 Land cover types (see online version for colours)

The basic architecture given in Figure 2 (Section 3.2) is enriched by our experimental data. The new architecture (Figure 10) summarises the NN processing phases.

The training phase is shown in Figure 11 and Figure 12. The iteration at which the validation performance reaches a minimum (equal to 0.18751) is epoch 11. The training continues for 6 more iterations before it stops.

No difficulties with the training are depicted. The validation and test curves (Figure 11) are very similar.

Figure 9 Learning sample (see online version for colours)

Urban _Close_to_Water_87→ A_1987

Urban _Close_to_Woodland_87→B_1987

Y/P	A-1987	B-1987	....	O-2001	P-2001
R1	0	0		0	1
R2	0	0	0	0	0
R3	0	1	1	1	1
....					
Rn	0	0	1	0	0

Y/P	A1	...	A10	B1	....	....	....	....	Pn
1987	0	0		0	0	....	0	....	1
2001	0	0		1	0	0	1	....	0

10 values

160 values

Figure 10 Processing phases of the predictive NN

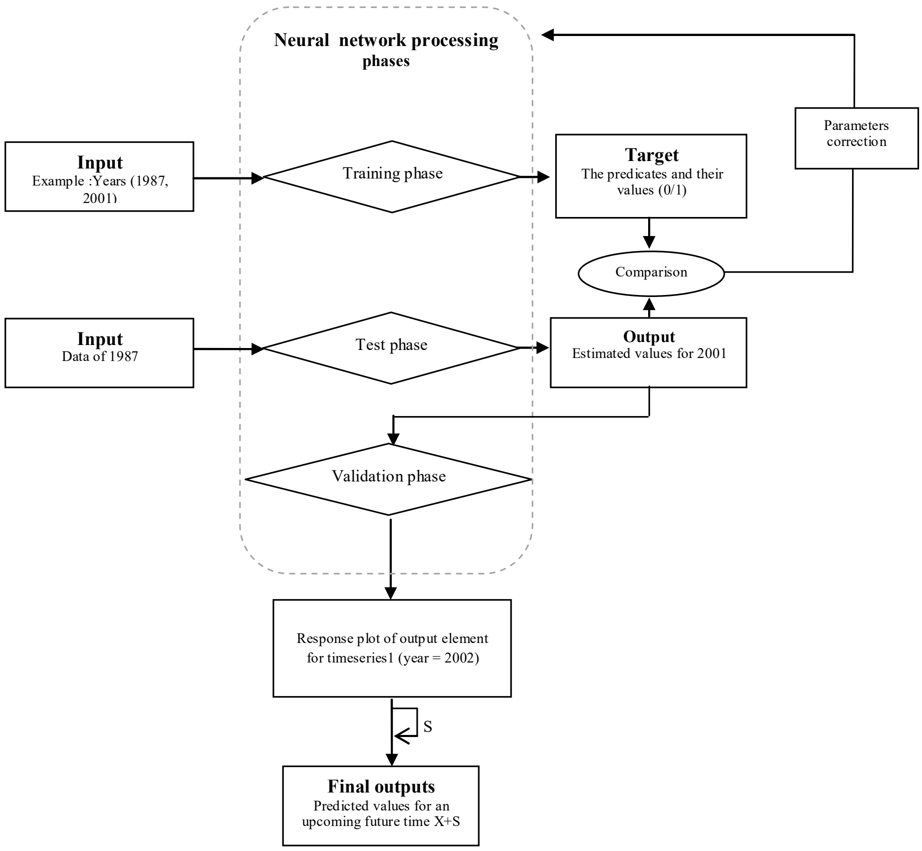


Figure 11 NN training performance (see online version for colours)

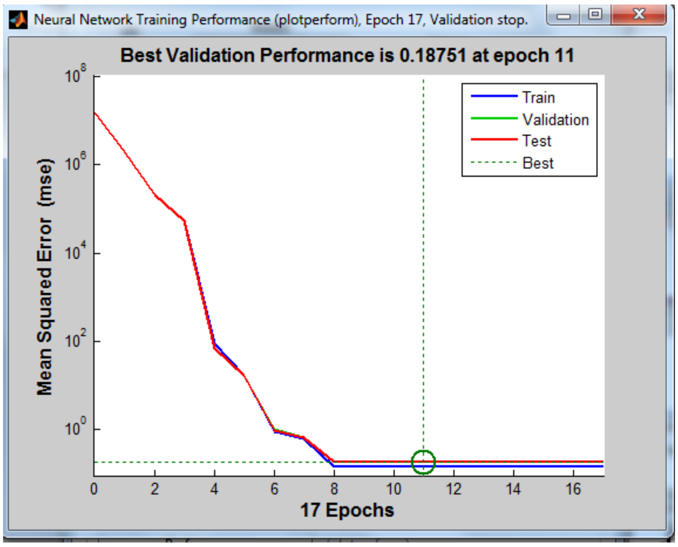
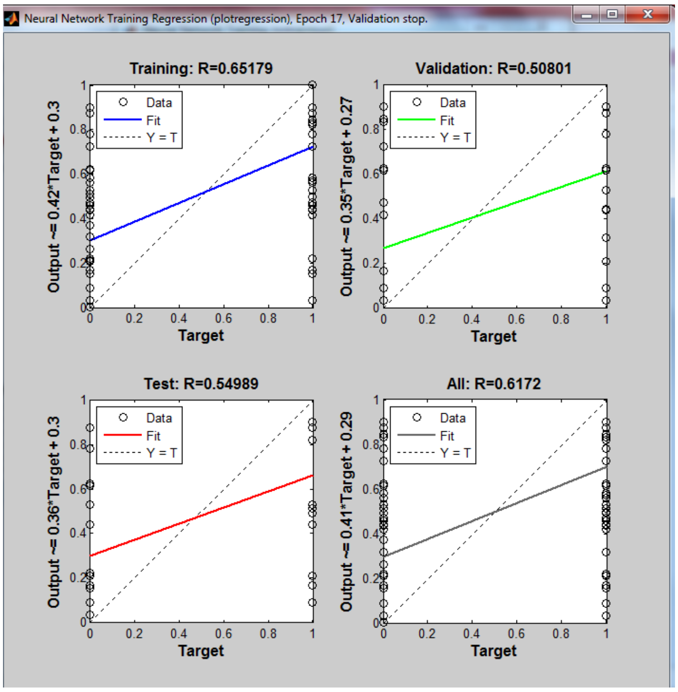


Figure 12 NN regression (see online version for colours)



The following step in the network validation is to create a regression plot, which shows the relationship between the outputs of the network and the targets. The result is shown in Figure 12. The axes represent the training, validation and testing data. The R value is related to the relationship between the outputs and targets. If R = 1, this implies an exact linear correlation between outputs and targets. If R is very close to zero, then there is no correlation between outputs and targets.

Figure 13 Predicted values for 2002 (see online version for colours)

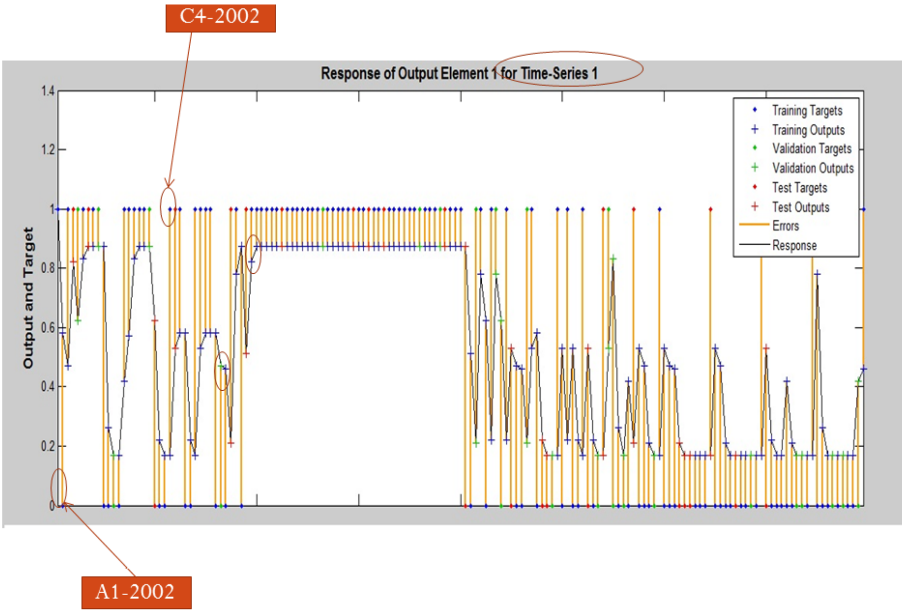


Figure 14 Predicted values for 2015 (see online version for colours)

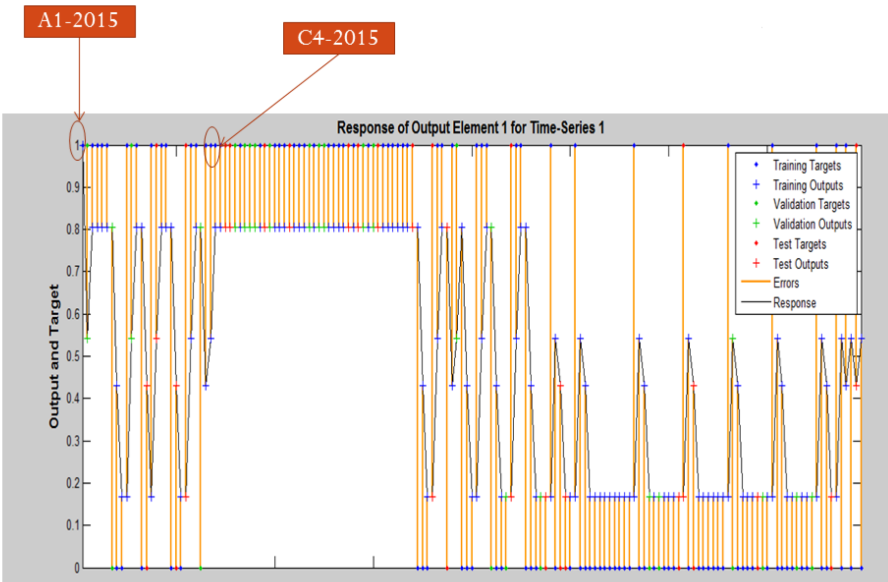


Table 4 Extract of spatiotemporal relationships set varying from 1987 to 2015

P/Y	*A1*	*A2*	*A3*	*A4*	*A5*	*A6*	*A7*	*A8*	*A9*	*A10*	*B1*	*B2*	*B3*	*B4*	*B5*	*B6*	*B7*	*B8*	*B9*	*B10*	*C1*	*C2*	*C3*	*C4*	*C5*	*C6*	*C7*	*C8*	*C9*	*C10*
1987	1	1	0	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	1	1	1	0	0	1	1
2001	0	0	1	1	0	1	1	1	1	1	0	0	0	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
2002	0	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	1	1	1	0	0	1	1	1	1
2003	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	1	1	1	0	0	1	1	1	1	0
2004	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	0	0	0	1	1	1	0	0	1	1	1	1
2005	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	0	0	0	1	1	1	0	0	1	1	1	1	0
2006	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	0	0	0	1	1	1	0	0	1	1	1	1	0	0
2007	1	1	1	1	0	0	0	1	1	1	1	1	1	0	0	0	1	1	1	0	0	1	1	1	1	0	0	0	1	1
2008	1	1	1	0	0	0	1	1	1	1	1	1	0	0	0	1	1	1	0	0	1	1	1	1	0	0	0	1	1	1
2009	1	1	0	0	0	1	1	1	1	1	1	0	0	0	1	1	1	0	0	1	1	1	1	0	0	0	1	1	1	0
2010	1	0	0	0	1	1	1	1	1	1	0	0	0	1	1	1	0	0	1	1	1	1	0	0	0	1	1	1	0	1
2011	0	0	0	1	1	1	1	1	1	0	0	0	1	1	1	0	0	1	1	1	1	0	0	0	1	1	1	0	1	1
2012	0	1	1	1	1	1	1	0	0	0	1	1	1	0	0	1	1	1	1	0	0	0	1	1	1	0	1	1	1	1
2013	0	1	1	1	1	1	1	0	0	0	1	1	1	0	0	1	1	1	1	0	0	0	1	1	1	0	1	1	1	1
2014	1	1	1	1	1	1	0	0	0	1	1	1	0	0	1	1	1	1	0	0	0	1	1	1	0	1	1	1	1	1
2015	1	1	1	1	1	1	0	0	0	1	1	0	0	0	0	1	1	0	0	0	1	0	1	1	0	1	1	1	1	1

During the validation phase, the NN generates the response plot of output element for Time-series 1 corresponding to the curve in Figure 13. This curve gives the values of the relationships in 2002. The NN sets the values > 0.5 to 1 and those < 0.5 to 0. For example, the attribute A1-2002 related to (Urban _Close_to_Water) is not verified in 2002 and takes 0 as value. The attribute C4-2002 related to (Urban _Close_to_Agriculture) is verified and takes 1 as value.

After S iterations corresponding to the number of years separating our reference date of 2002 and the new date to be predicted, we obtain the values we are looking for [e.g., year = 2015 (Figure 14)]. Table 4 contains an extract of the prediction results. Indeed, the results table contains 16 rows corresponding to 16 years ranging from 1987 till 2015 and 160 columns relevant to different spatiotemporal relationships.

5.2 Results and discussion

Applying the Apriori algorithm on this dataset generates predictive rules reassembling the relationships calculated at different time periods (from t1 = 1987 till tm = 2015). An example of an interesting rule is presented as follows:

$$\begin{aligned}&A\_2001 = \text{true} \wedge D\_2001 = \text{false} \wedge M\_2001 = \text{false} \wedge A\_2010 = \text{true}\\&\wedge D\_2010 = \text{true} \wedge M\_2010 = \text{false} \wedge A\_2013 = \text{false} \wedge D\_2013 = \text{true}\\&\wedge I\_2013 = \text{true} \rightarrow I\_2015 = \text{false} \wedge D\_2015 = \text{true} \wedge M\_2015 = \text{true}(0.98)\end{aligned}$$

This rule can be interpreted as shown in Table 5.

Table 5 Rule's interpretation

2001	*2010*	*2013*	*2015*
Urban zone is close to a water zone	Urban zone is close to a water zone	Urban zone is not close to a water zone	Urban zone is not very close to a water zone
Urban zone is not close to a bare-soil zone	Urban zone is close to a bare-soil zone	Urban zone is close to a bare-soil zone	Urban zone is close to a bare-soil zone
Urban zone is not inside a water zone	Urban zone is not inside a water zone	Urban zone is very close to a water zone	Urban zone is inside a water zone

In order to prove the efficiency of our results, we compare our predictions with the reality statements. We use two images of our study zone (taken from Google Earth) in 2010 and 2013 (Figure 15 and Figure 16). We choose randomly an urban zone and a water zone as shown in Figure 17.

We add these records to our STDB. Then, we calculate the distance between the urban and water zones through a spatial query. Results are shown in Figure 18.

The distance between the urban zone and water zone in 2010 is equal to 11.14 km. This distance implies a 'close-to' relationship. The distance between the same urban zone and water zone in 2013 is equal to 0.45 km which indicates a 'very close-to' relationship. Indeed, the reality results totally match with the rule's prediction. Applying the spatial queries on all urban and water zones of the STDB gives the following rates: 78% of urban zones (25 zones among 32 zones) are 'close_to' water zones in 2010 and 90% of urban zones (29 zones among 32 zones) are 'very_close_to' water zones in 2013. An extract of the results is shown in Table 6.

Figure 15 The study zone in 2010 (see online version for colours)

Figure 16 The study zone in 2013 (see online version for colours)

On the basis of these statistics, the predictions of 2015 have a remarkable probability to be exactly near to the reality. Referring to our prediction results, a possible SHP file describing the study zone in 2020 is given in Figure 19. It shows that many urban zones are menaced by the inundation risk.

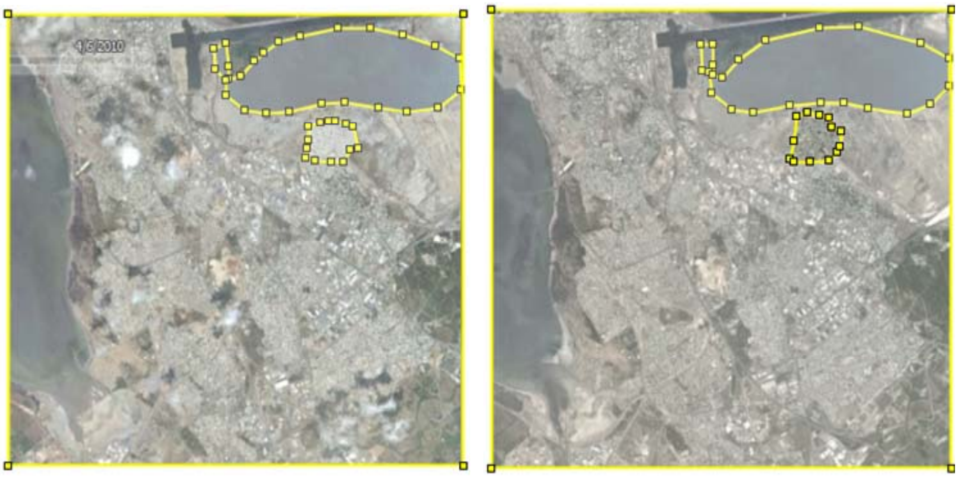


Figure 17 The randomly chosen objects on the image of 2010 and 2013 (see online version for colours)

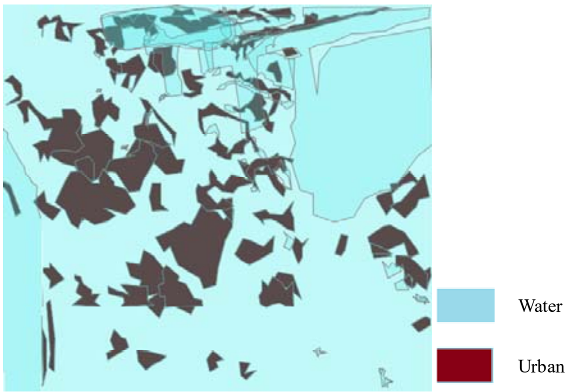
Figure 18 Query's results

3666	urban-2010	water-2010	11.14 km
3792	urban-2013	water-2013	0.45 km

Table 6 Extract of the distance calculation (year 2013)

name_urban	*name_water*	*distance_between*
urban1	water1	0.00
urban1	water2	32.77
urban1	water3	86.04
urban1	water4	375.45
urban1	water5	155.16
urban1	water6	209.53
urban1	water7	177.02
urban1	water8	253.20
urban1	water9	2.43
urban2	water1	0.00
urban2	water2	60.07
urban2	water3	85.16
urban2	water4	298.21
urban2	water5	146.86
urban2	water6	209.31
urban2	water7	180.91
urban2	water8	253.59

Figure 19 Cartographic representation of the prediction results in 2020 (see online version for colours)



5.3 Comparative study

Here after, we compare our approach to some standard approaches among those cited in the state of the art (Section 2). We select two research works based on HMM and MAS. Another approach based on fuzzy sequence extraction is also considered in the comparative study as shown in Table 7. Our choice is driven by the resemblance of the addressed topic in these works to our studied issue.

All these approaches try to model the evolution of spatial objects. Our approach has the particularity to predict the future evolution of spatiotemporal relationships between spatial objects.

Another difference concerns the nature of the results. In the case of HMM or MAS, the results are expressed by probabilities between classes referring to the rate changes in the past. The extracted rules interpretation needs a big knowledge of the method concept and an expert's intervention. Similarly, in the case of fuzzy sequence extraction method, the evolution rules are based on a 'model checking' which qualifies an evolution hypothesis by possible or impossible. Such evaluation is vague and makes the comparison to the real data difficult. In our case, the rules are semantically rich.

6 Conclusions

In this paper, we addressed the issue of spatiotemporal relationships prediction. Our predictive system was a NN based on STAR as learning examples. These rules described the full history of spatiotemporal relationships evolution between known dates.

The prediction method was based on time series containing the relationships between spatial objects at a given time. The results of the experimentation were STPR assessing prohibited upcoming risks implicitly expressed by spatiotemporal relationships. The comparison between predicted values and ground truth gave good correspondence rates ranging between 78% and 90%.

In future works, we will extend our approach of spatiotemporal analysis by adding other indicators of change especially demographic data and climatic features. We also aim to implement our approach in gene regulatory analysis.

Table 7 Comparative study

	Data	*Objective*	*Method basis*	*Technical tools*	*Change nature*	*Obtained results*	*Evaluation*	*Adaptability to other data types*
Hidden Markov models (Essid et al., 2012)	Satellite images	Detection of spatiotemporal changes in an urban area.	Generation of change probabilities between classes referring to the rate changes in the past.	HMM Viterbi algorithm	Change of texture Change of gray levels	Rates indicating low, middle, or strong expansion	Evaluation function included in the Interpretation system.	Not indicated
Multi-agent systems (Saheb Ettabaâ et al., 2004)	Satellite images	Interpretation of a sequence of satellite images taken at different dates in a purpose of spatiotemporal modeling.	Each spatial object is considered as a dynamic system whose evolution can be modeled by temporal automata. Hierarchical blackboard control driven by goals (queries) and data (events)	Temporal automata	Change of texture Change of the land cover type	Graphs Modeling land parcels dynamics	Ground truth comparison	Yes
Fuzzy sequence extraction (Zoghlami, 2013)	Maps	Detection of urban evolution	A degree of possibility is calculated for each pattern of spatial evolution.	Fuzzy queries apriori algorithm	Future trajectory of a town	Evolution Rules	No	Not indicated
Neural network/time series (our approach)	Satellite images STDB	Prediction of Spatiotemporal relationships	Generation of spatiotemporal predictive rules indicating spatiotemporal relationships evolution.	Spatial queries apriori algorithm Neural network Time series analysis	Changes of ST relationships	STPR	Ground truth comparison	Yes

A prediction of the evolutionary behaviours of gene segments and other substances in the cell may be a very interesting issue.

References

Agrawal, R., Imielinski, T. and Swami, A. (1993) 'Mining association rules between sets of items in large databases', *Proceedings of the ACM SIGMOD, Conference on Management of Data, ACM Press*, New York, pp.207–216.

Alcamo, J., Kreileman, G.J.J., Bollen, J.C., Van den Born, G.J., Gerlagh, R., Krol, M.S., Toet, A.M.C. and Vries, H.J.M. (1996) 'Baseline scenarios of global environmental change', *Global Environmental Change*, Vol. 6, No. 4, pp.261–304.

Appice, A., Berardi, M., Ceci, M. and Malerba, D. (2005) 'Mining and filtering multi-level spatial association rules with ARES and discovery of spatial association rules in geo-referenced census data: a relational mining approach', in Hacid, M-S. et al. (Eds.): *ISMIS, LNAI*, Vol. 3488, pp.342–353, Springer-Verlag, Berlin, Heidelberg.

Bowerman, B.L. and O'Connel, R.T. (1993) *Forecasting and Time Series: An Applied Approach*, 3rd ed., p.526, Duxbury Press, Belmont, CA.

Crone, S.F., Hibon, M. and Nikolopoulos, K. (2011) 'Advances in forecasting with neural networks? Empirical evidence from the NN3 competition on time series prediction', *International Journal of Forecasting*, Vol. 27, No. 3, pp.635–660.

Derrode, S. and Carincotte, C. (2005) 'Applications des chaînes de Markov floues à la segmentation non supervisée d'images, Ateliers Traitement et Analyse de l'Information: Méthodes et Applications', *TAIMA'05*, Hammamet, Tunisia, 26 September–1 October, pp.193–198.

Dinga, Y., Zhao, Kai., Zhenga, X. and Jianga, T. (2014) 'Temporal dynamics of spatial heterogeneity over cropland quantified by time-series NDVI, near infrared and red reflectance of Landsat 8OLI imagery', *International Journal of Applied Earth Observation and Geoinformation*, Vol. 30, pp.139–145, DOI.10.1016/j.jag.2014.01.009.

Durante, F., Foscolo, E., Jaworski, P. and Wang, H. (2014) 'A spatial contagion measure for financial time series', *Expert Systems with Applications*, Vol. 41, No. 8, pp.4023–4034.

Essid, H., Sellami, A., Barra, V. and Farah, I.R. (2012) 'Monitoring intra-urban changes with hidden Markov models using the spatial relationships', *ICGST-GVIP Journal*, Vol. 12, No. 1, pp.49–55.

Farah, I.R., Saheb-Ettabaâ, K., Solaiman, B. and Ben-Ahmed, M. (2006) 'Analyzing spatial – temporal geographic information based on blackboard architecture and multi-agent system', *International Journal of Computer Science and Network Security (IJCSNS)*, Vol. 6, No. 8A, pp.4–10.

Gazzaz, N., Yusf, M.K., Aris, A.Z., Juahir, H. and Ramli, M.F. (2012) 'Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors', in *Marine Pollution Bulletin Journal*, Vol. 64, No. 11, pp.2409–2420.

Hon-lun, Y., Hongqin, F. and Yat-hung, C. (2014) 'Predicting the maintenance cost of construction equipment: comparison between general regression neural network and Box-Jenkins time series models', *Automation in Construction*, Vol. 38, pp.30–38, DOI: 10.1016/j.autcon.2013.10.024.

Huang, X. and Friedl, M. (2014) 'Distance metric-based forest cover change detection using MODIS time series', *International Journal of Applied Earth Observation and Geoinformation*, Vol. 29, pp.78–92, DOI: 10.1016/j.jag.2014.01.004.

Isik, S., Kalin, L., Schoonover, J.E., Srivasatava, P. and Lockaby, B.G. (2013) 'Modeling effects of changing land use/cover on daily streamflow: an artificial neural network and curve number based hybrid approach', *Journal of Hydrology*, Vol. 485, pp.103–112, DOI: 10.1016/j.jhydrol.2012.08.032.

Karatunga, A. (2005) 'Role of geographic information systems (GIS) in disaster management, the Ungandan case', *International Workshop on University Network for Natural Disaster Reduction in Africa*.

Koperski, K. and Han, J. (1995) 'Discovery of spatial association rules in geographic information databases', in *Proceedings of the 4th Int.' Symp on Large Databases (SSD'95)*, Portland, pp.47–66.

Leasage, J. and Krivelyova, A. (1999) 'A spatial prior for Bayesian vector autoregressive models', *Journal of Regional Science*, Vol. 39, No. 2, pp.297–317.

Loglisci, C., Ceci, M. and Malerba, D. (2013) 'Discovering evolution chains in dynamic networks', in Appice, A., Ceci, M., Loglisci, C., Manco, G., Masciari, E. and Ras, Z.W. (Eds.): *New Frontiers in Mining Complex Patterns – First International Workshop, NFMCP 2012, Held in Conjunction with ECML/PKDD 2012, Springer 2013 Lecture Notes in Computer Science*, Bristol, UK, 24 September 2012, Revised Selected Papers, Vol. 7765, pp.185–199.

Lyons, M., Roelfsema, C. and Phinn, S. (2013) 'Towards understanding temporal and spatial dynamics of seagrass landscapes using time-series remote sensing Estuarine', *Coastal and Shelf Science*, Vol. 120, pp.42–53, DOI: 10.1016/j.ecss.2013.01.015.

Malerba, D. (2008) 'A relational perspective on spatial data mining', *Int. J. Data Mining, Modelling and Management*, Vol. 1, No. 1, pp.103–118.

Manson, S.M. (2000). 'Agent-based dynamic spatial simulation of land-use/cover change in the Yucatán Peninsula, Mexico', Paper presented in the *Fourth International Conference on Integrating GIS and Environmental Modeling (GIS/EM4)*, Banff, Canada.

Molodenskii, D.S. and Molodenskii, M.S. (2011) 'Variations in tidal response of the medium in the spatiotemporal neighborhood of catastrophic earthquake in Japan', *Izvestiya, Atmospheric and Oceanic Physics*, Vol. 46, No. 8, pp.992–994.

Nabel, J., Kirchner, J., Zurbriggen, N., Kienast, F. and Lischke, H. (2014) 'Extrapolation methods for climate time series revisited spatial correlations in climatic fluctuations influence simulated tree species abundance and migration', *Ecological Complexity Journal*, DOI: 10.1016/j.ecocom.2014.02.006.

Neeti, N. and Eastman, J.R. (2014) 'Novel approaches in extended principal component analysis to compare spatio-temporal patterns among multiple image time series', *Remote Sensing of Environment*, Vol. 148, pp.84–96, DOI: 10.1016/j.rse.2014.03.015.

Pijanowski, B.C., Tayyebi, A., Doucette, J., Pekin, B.K., Braun, D. and Plourde, J. (2014) 'A big data urban growth simulation at a national scale: configuring the GISnd neural network based land transformation model to run in a high performance computing (HPC) environment, *Environmental Modelling & Software*, Vol. 51, pp.250–268.

Pouliot, D., Latifovic, R., Zabcic, N., Guindon, L. and Olthof, I. (2014) 'Development and assessment of a 250 m spatial resolution MODIS annual land cover time series (2000–2011) for the forest region of Canada derived from change-based updating', *Remote Sensing of Environment*, Vol. 140, pp.731–743, DOI: 10.1016/j.rse.2013.10.004.

Saheb Ettabaâ, K., Farah, I.R. and Ahmed, M.B. (2004) 'Mulit-agent system for detecting changes on satellite image sequence', *International Conference on Industrial Technologies*, Hammamet, IEEE Catalogue Number: 04TH8771c, December.

Serneels, S. and Lambin, E.F. (2001) 'Proximate causes of land-use change in Narok District, Kenya: a spatial statistical model', *Agriculture, Ecosystems and Environment Elsevier Science*, pp.65–81

Singh, M., Mishra, V.D., Thakur, N.K. and Sharma, J.D. (2012) 'Remote sensing GIS based statistical modelling for the prediction of natural hazards', *International Journal of Engineering Research & Technology (IJERT)*, Vol. 1, No. 6, pp.1–7.

Tatem, A.J., Lewis, H., Atkinson, P.M. and Nikson, M.S. (2002) 'Super-resolution land cover pattern prediction using a Hopfield neural network', *Journal of Remote Sensing of Environment*, Vol. 79, No. 1, pp.1–14.

Toledano, J., Eastman, J.R., Sangermano, F., Labs, C. and Killeen, T. (2008) 'GIS helps gauge land change effects on the nation's endemic biodiversity', *International ESRI*.

Wei, W. (1990) *Time Series Analysis, Unvariate and Multivariate Methods*, pp.162–167, Addison-Wesley, Boston.

Xin, Z., Yu, X. and Li, Q. (2011) 'Spatiotemporal variation in rainfall erosivity on the Chinese Loess Plateau during the period 1956–2008', *X. X. Lu Reg. Environ. Change*, Vol. 11, No. 1, pp.149–159.

Zoghlami, A. (2013) *Modélisation et conception de systèmes d'information géographique gérant l'imprécision*, Thèse de doctorat en informatique, université de Saint Denis, mars.