Sharafi S, Fouladvand S, Simpson I & Alvarez JAB (2016) Application of pattern recognition in detection of buried archaeological sites based on analysing environmental variables, Khorramabad Plain, West Iran, *Journal of Archaeological Science: Reports*, 8, pp. 206-215.

DOI: 10.1016/j.jasrep.2016.06.024

Application of Pattern Recognition in Detection of Buried Archaeological Sites based on Analysing Environmental Variables, Khorramabad Plain, West Iran

Siyamack Sharafi, academic centre for education, culture and research, Lorestan, Iran.
Sajjad Fouladvand*, academic centre for education, culture and research, Lorestan, Iran.
Tel: +989163696538, Fax: 06633242620, E-mail: sjjd.fouladvand@gmail.com
Ian Simpson, school of natural science, university of stirling, Stirling, Scotland, United Kingdom.
Juan Antonio Barcelo Alvarez, department of prehistory, autonomous university of barcelona, Bellaterra, Spain.

ABSTRACT

Archaeologists continue to search for techniques that enable them to analyze archaeological data efficiently with Artificial Intelligence approaches increasingly employed to create new knowledge from archaeological data. The purpose of this paper is to investigate the application of Pattern Recognition methods in detection of buried archaeological sites of the semi-arid Khorramabad Plain located in west Iran. This environment has provided suitable conditions for human habitation for over 40,000 years. However, environmental changes in the late Pleistocene and Holocene have caused erosion and sedimentation resulting in burial of some archaeological sites making archaeological landscape reconstructions more challenging. In this paper, the environmental variables that have influenced formation of archaeological sites of the Khorramabad Plain are identified through the appliation of Arc GIS. These variables are utilized to create an accurate predictive model based on the application of One-Class classification Pattern Recognition techniques. These techniques can be built using data from one class only, when the data from other classes are difficult to obtain, and are highly suitable in this context. The experimental results of this paper confirm one-class classifiers, including Auto-encoder Neural Network, K-means, Principal Component Analysis Data Descriptor, Minimum Spanning Tree Data Descriptor, K-Nearest neighbor and Gaussian distribution as promising applications in creating an effective model for detecting buried archaeological sites. Among the investigated classifiers, Minimum Spanning Tree Data Descriptor achieved the best performance on the Khorramabad Plain data set.
KEY WORDS: Artificial Intelligence; Pattern Recognition; One-Class classification; Predictive Modeling; Khorramabad Plain; Environmental Variables.

1 INTRODUCTION

The detection and spatial characterisation of archaeological sites based on geomorphological parameters is now an essential aspect of landscape archaeology research (Ayala and French, 2005; Barton et al., 2002, 2010; Butzer, 1982; Schiffer, 1983; Tartaron et al., 2006; Wells, 2001). Increasingly this work is being integrated through the application of GIS based analyses that allows efficient spatial and locational analyses of site – environment relationships. (Gouma, 2011, Kuiper and Wescott, 1999; Bala et al., 2014). Within the suite of quantitative GIS based techniques applied to landscape archaeology, predictive models are enabling researchers to estimate the possibility of presence or absence of archaeological evidence across extensive areas of search (Ebert, 2004; Kamermans and Rensink, 1999). Inductive based approaches used in both Archaeological Heritage Management (AHM) and scientific research, creates a model based on correlations between previously identified archaeological sites and variables that are obtained from the current physical landscape. Deductive approach, which are relatively rare, constructs the predictive model based on prior anthropological and archaeological knowledge, and uses previously identified sites to evaluate the model (Kamermans, 2006). Numerous predictive models have been developed using different methods including Bayesian statistics and Dempster-Shafer modelling to detect archaeological sites (Verhagen et al., 2010; Kvamme, 1990; Lang, 2000; Gibbon, 2000; Konnie et al. 2000; Fernandes et al. 2011) and in developing these approaches Kamermans (2010) has identified a range of problems concerned with quality and quantity of archaeological input data including relevance of the environmental input data, lack of temporal and/or spatial resolution, use of spatial statistics, testing of predictive models, and need to incorporate social and cultural input data. A number of recommendations to address these problems have been developed as archaeological experience with quantitative GIS has emerged (Verhagen et al, 2009).

Artificial intelligence (AI) is the intelligence exhibited by machines or software. In recent years there has been growing interest in applying AI in many fields including data mining (Perumal et. al., 2015). In archaeology its application has been through Artificial Neural Networks (ANNs) and expert systems (Vitali, 1991; Voorrips, 1990; Richards, 1998). Deravignone and Jánica (2006) studied the basic concepts required to bring artificial intelligence, in particular ANNs into archaeological research investigating the application of ANNs in a raster GIS environment with the aim of creating archaeological predictive models. Barceló (2010) reviewed the implication of using Computational Intelligence in archaeology. He explained that artificial intelligence models are feasible in archaeological recognition systems just like other sciences. Puyol-Gruart (1999) has considered the possibility of using more recent subfields including Knowledge Discovery in Databases (KDD), Visual Information Management (VIM) and Multi-agent Systems (MAS) in archaeological research.

This paper is a first comparative analyses of the spatially predictive capabilities of different AI methods in a semi-arid regional context. Environmental variables that have influenced formation of archaeological sites located in the Khorramabad Plain, Iran, are derived through application of Arc GIS. These variables are then utilized to create a predictive model based on Pattern Recognition, one of the most important subfields of AI. The term pattern recognition has evolved substantially from its roots in artificial intelligence, engineering and statistics. Patten recognition is the study of how machines perceive the environment, learn to recognize pattern of desired class from their background, and from these machine based observations make reasonable decisions about the categories of the different patterns (Jain, 2000). One-class classification as a pattern recognition method was developed by Moya and Hush (1996; Pimentel, et. al., 2014). One-class classification endeavours to identify objects of a specific class amongst all samples, by learning from a training set containing only the samples of that class. In one-class classification, it is assumed that only information of one of the classes, the target class, is available (Tax, 2001). So, the most valuable feature of one-class classifiers that makes it important to the objectives of this paper is that these types of classifiers can be built using only data from archaeological sites when the data from non-archaeological site class is difficult to obtain (which they usually are).

In this paper applications of GIS spatial analysis and one-class classification methods are employed to detect buried archaeological sites of the Khorramabad Plain, a geomorphic unit located in the southern part of the Khorramabad Valley with antiquity more than 40,000 years of human settlement. Section 2 of the paper examines the details of defined variables and generated data set together with a brief overview of pattern recognition models considered. Experimental results and discussion are drawn in section 3; in this section the efficacy of using one-class classification in detecting buried archaeological sites is clearly shown and discussed. Section 4 gives a summary of this work and propose some ideas for future research in the field of Pattern Recognition applied to archaeology.

1.1 Geographical and archaeological features of Khorramabad Plain

Khorramabad Plain is located in Lorestan province, west Iran; it lies within E 48 11” to E 48 28” and N 33 19” to N 33 30”. The Khorramabad River passes across the plain which is surrounded by high mountains. Northern and central parts of the plain include urban areas where the possibility of archaeological sites surviving is unlikely and so the southern part of the plain is investigated in this paper. This area of the plain is characterised by alluvial deposits and the Zagros folded zone (Figure 1). The plain has low sloping topography with the minimum altitude 1135 meters and a maximum height of 1436 meters. The annual average of temperature in the area is 17.2 ° C and average precipitation is 502 mm per year.

Khorramabad Plain is one of the oldest residential plains, occupied from the Palaeolithic period through to the Islamic era. The Kunji and Ghamari caves in mountain areas around the plain (Palaeolithic), Masour mound in southern parts of the plain (Neolithic to the Islamic) and Falak-al-aflak castle in central parts of the plain (Sasanian), indicate the rich history as one of the first and longest lasting human habitations in the region. Water resources, fertile soil and a flat topography in the southern parts of this plain serves to indicate the existence of numerous cultures and archaeological sites in this area (Hole, 1970) and this together with the compactness of the plain has led to the plain becoming an important focus for archaeological investigation. Archaeological excavations including Hole and Flannery during the years 1959 to 1960 and 1963 to 1965, Demorgan (1891), Herzfeld (1928), Cl.Coff (1961), Berman (1978), Wright, Nelly and Johnson (1975) and Wenen (1972) and Javadi et al (2000) in caves, rock shelters and archaeological sites indicate the importance of this area of Iran to the understanding the dynamics of social, cultural and environmental change (Javadi and Borazjani , 2000).

Archaeologically, Khorramabad Plain has a sequence of Islamic, historic and prehistoric eras and by the year 2000 some 43 known historical sites have been identified (Table 1). However, environmental changes in the late Pleistocene and Holocene such as formation of the Kar-Gah Lake and morphological changes in the Khorramabad River’s path, has caused erosion and sedimentation processes over time such that some archaeological sites have disappeared and others have been buried under soil and sediment. This paper outlines application of one-class classification methods, in the detection of these buried archaeological sites.

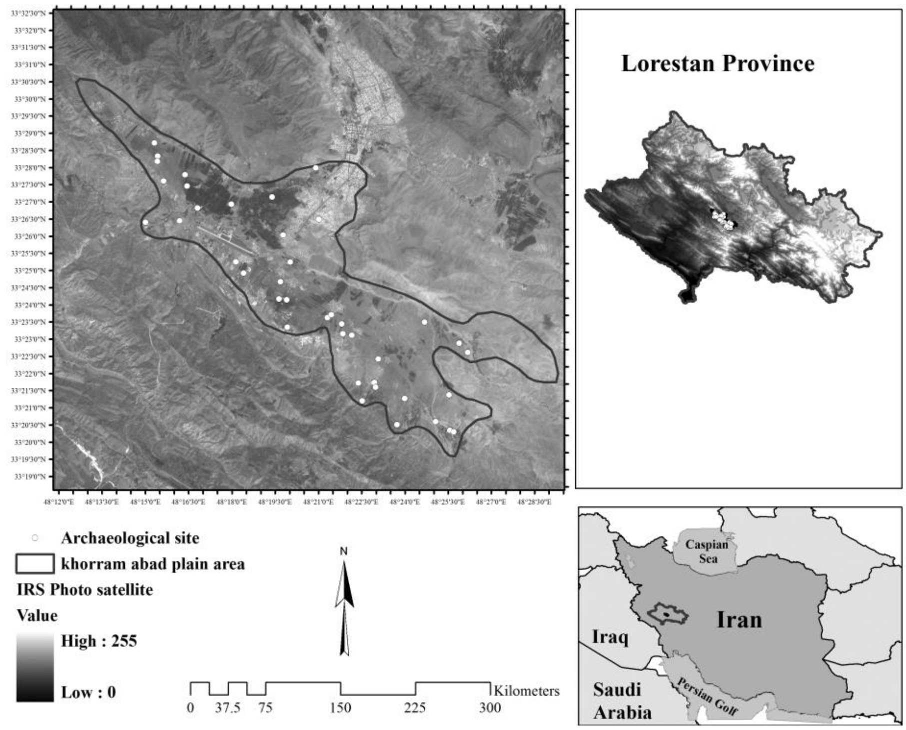


Figure 1) Location of Khorramabad Plain in Iran together with the location of 43 previously known sites in the study area.

Table 1: Archaeological sites of Khorramabad Plain (Javadi and Borazjani, 2000)

Code	Altitude of the site with respect to pre-site ground level (meter)	Name	Period
Kp1	12	Masour mound	**Sasanian to Chalcolithic**
Kp2	10	Tagh and togh mound	**Historical**
Kp3	1	Bagheno mound	**Islamic-Historical-Prehistoric**
Kp4	5.8	Margaymeh tombs	**Islamic**
Kp5	2.5	Dinarvand mound	**Historical**
Kp6	19	Sofreh mound	**Early and Middle Islamic-Historical**
Kp7	3	Khaki mound	**Chalcolithic**
Kp8	1	Dinarvand 1 mound	**Historical-Calcholithic**
Kp9	17	Sorkh deh mound	**Islamic-Historical**
Kp10	3.5	Bazgir mound	**Middle Islamic-Historical**
Kp11	2	Gorbacheh cemetery	**Historical**
Kp12	10	Sarkalak site	**Middle Islamic-Historical-Bronze Age**
Kp13	5.5	Rava hell mound	**Chalcolithic**
Kp14	11	Armani mound	**Early Islamic-Historical**
Kp15	2	Daraei cemetery	**Late Islamic-Historical**
Kp16	1.5	Daraei site	**Historical**
Kp17	5	Naservand 2 mound	**Chalcolithic-Neolithic**
Kp18	3	Sorkh deh 1 site	**Historical**
Kp19	2.3	Dehbagher mound	**Chalcolithic**
Kp20	8	Angoz site	**Middle Islamic-Historical-Bronze Age**
Kp21	5	Naservand 1 mound	**Chalcolithic-Historical**
Kp22	5.5	Chesmeh sorkheh site	**Middle Islamic-Historical**

Kp23	3	Pol baba hossein mound	**Parthian-Bronze**
Kp24	2.5	Pirjed site	**Chalcolithi-Historical**
Kp25	2.5	Sorkh deh 2 site	**Historical**
Kp26	8	Poll babahossein site	**Middle Islamic-Parthian**
Kp27	16	Ali sabz site	**Parthian-Iron age**
Kp28	7.5	Chegahoroshi 1 mound	**Early Islamic-Sasanian-Bronze-Chalcolithic**
Kp29	1.25	Pakoreh mound	**Parthian**
Kp30	1	Asgarabad 1 mound	**Early and Middle Chalcolithic-Parthian**
Kp31	1.5	Chi kham la mound	**Middle Islamic-Parthian**
Kp32	1.5	Deh mohsen mound	**Middle Islamic-Historical**
Kp33	2	Asgharabad mound	**Historical-Chalcolithic-Neolithic**
Kp34	5	Sohel baigi 2 mound	**Historical-Bronze-Calcholithic**
Kp35	5	Sohel baigi 1 mound	**Historical- Chalcolithic-Neolithic**
Kp36	1.25	Chegahoroshi 3 mound	**Historical**
Kp37	1.25	Daymeh araban mound	**Sasanian-Parthian**
Kp38	2.5	Rusi mound	**Early and Middel Islamic-Historical**
Kp39	2.5	Roghani mound	**New Neolithic-Chalcolithic**
Kp40	2	Fathollah mound	**New Chalcolithic**
Kp41	20	Hellat rashno mound	**Early and Middle Islamic-Sasanian-Parthian**
Kp42	2.5	Chegahoroshi 2 mound	**Bronze Age**
Kp43	1	Cham khoregh mound	**Middle Islamic-Sasanian**

2 METHODOLOGY

The methodology used here can be divided into two main stages. Firstly, the environmental factors of 43 archaeological sites of Khorramabad Plain are collected using Arc GIS. These 43 archaeological sites were detected in 2000 by the Cultural Heritage of Lorestan province (Javadi and Borazjani , 2000). Secondly, the results of the first part are applied as input to create predictive models based on one-class classification methods. Environmental factors of the 43 archaeological sites including elevation (1:25000), slope (1:25000), precipitation (1:50000), distance to river (1:25000), distance to accessible roads (the roads which are used in this research highly overlapped with the ancient roads) (1:50000) and water resources (1:50000) were prepared, and then raster layers of these factors were generated utilizing Arc GIS. Slope, elevation, distance to roads and distance to river are generated using digital topography maps which are prepared by National Cartographic Center, Iran. Precipitation is generated using data from synoptic and climatology stations and interpolation methods incorporated in ArcGIS. Water ground level is produced using data from piezometric wells in the area of study and interpolation methods. Figure 2 represents these raster layers in which the sizes of each cell is 20 * 20 m. Digital values for each factor were extracted using the Sample tool in Arc GIS and then they were exported into Microsoft Excel 2010; Min-Max normalization (Han and Kamber, 2006) was performed on the data set to reduce the effect of measurement unit on the learning process of models. As outcomes, a digital database of the environmental characteristics of 43 historical sites is used to build Artificial Intelligence (AI) based predictive models.

Our analyses have access only to data from the target class, the 43 detected archaeological sites data. In order to evaluate the models, samples from the outlier class (parts of the plain where there are no archaeological sites) are required. To do so a new dataset containing 43 target samples (archaeological samples) and another 43 artificially generated non-targets (non-archaeological samples) are created. The non-targets are drawn from a block-shaped uniform distribution that covers the target data (Tax, 2001; Tax, 2014). It is worth noting that the block-shaped distribution works efficiently for this work which models a low dimensional data set. However, an alternative for high dimensional data sets are Gaussian distribution or Gaussian Mixture Models which can be used to cover high dimensional data sets (Bishop, 2006; Tax, 2001). Finally, a dataset which contains 86 archaeological and non-archaeological samples is created and used to train, validate and test the predictive models using a nested 10-fold cross validation (Alpaydin, E., 2004) method. Training a pattern recognition model and testing it on the same data is not reliable because a model that simply repeats the labels of samples that it has just seen would have a perfect performance but would have a high error on yet-unseen samples. To avoid this, it is common practice when performing pattern recognition experiments to hold part of the available data as a set for testing the trained model. Additionally, by using only one test set the results may be biased towards the specific situation of this test set, especially when the data set is small. Consequently, the tuning experiments are repeated many times with different validation and test sets to gather enough statistical validity. More precisely, parameters of the models are trained and optimized using a nested K-fold cross validation (Alpaydin, E., 2004) method in our work.

As a result of the above experiment, a model is identified which can efficiently separate archaeological sites from non-archaeological sites. Then, this best model is assessed using a new test set generated using field studies of the Khorramabad Plain as a ground validation of the model. In this policy of non-archaeological test sample generation, 45 locations on the plain such as areas excavated for building different facilities including roads, tunnels and transects are considered non-archaeological samples. Figure 3 shows some examples of the archaeological and non-archaeological sites. The 45 non-archaeological test sites are combined with the previously known 43 archaeological sites to create a real world data set including both archaeological and non-archaeological samples which are used to better evaluate the best method identified in our first experiment using artificially generated non-archaeological test samples.

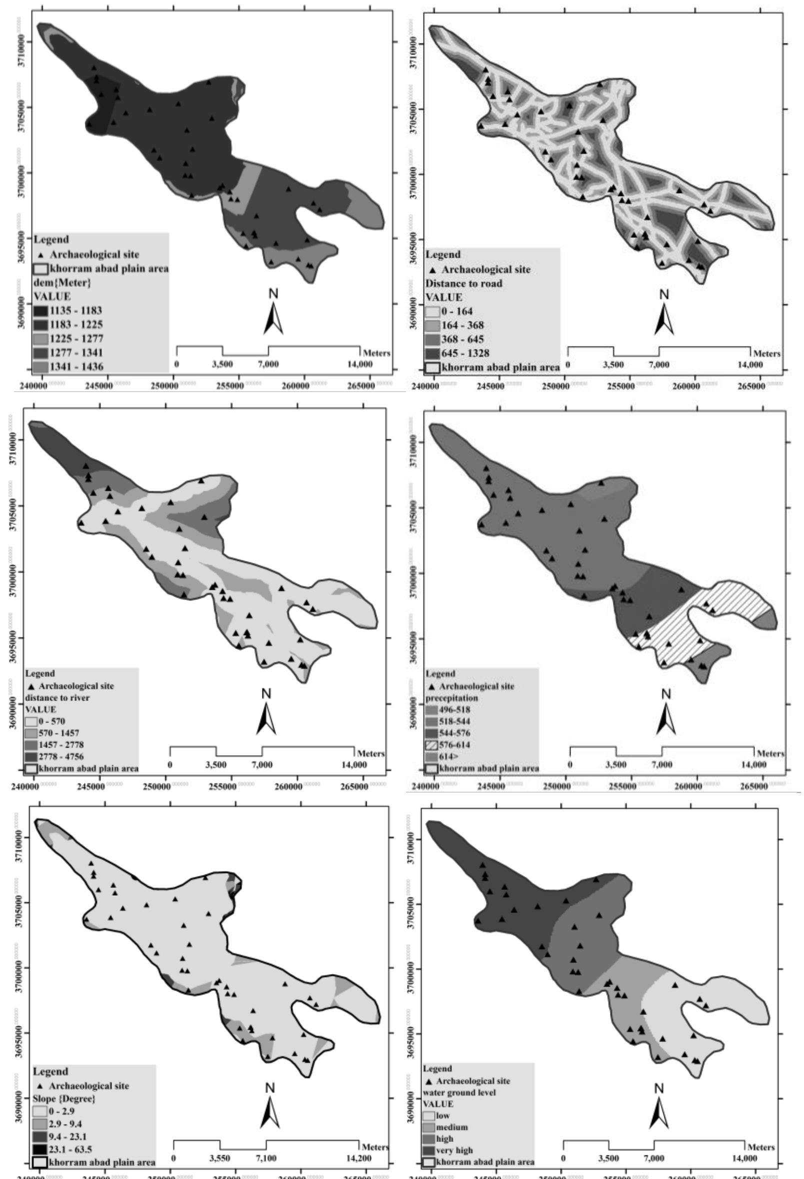


Figure 2) Environmental-based raster layers used in the models together with the location of 43 previously known sites in the study area.

(a) A known archaeological site example (Bagheno mound) **(b) A known archaeological site example (Daraei site)**

(c) A non-archaeological test site (a trench) **(d) A non-archaeological test site (a seasonal river)**

Figure 3) Examples of archaeological and non-archaeological test sites obtained from field studies on the Khorramabad Plain.

There are a number of one-class classifiers in the literature (TAX, 2001; Khan and Madden, 2010). In this paper, several sophisticated one-class classifiers that have been widely used in the literature are investigated. The implemented classifiers are: auto-encoder neural network, k-means data descriptor, Principal Component Analysis data descriptor (PCA_DD), Minimum Spanning Tree data descriptor (MST_DD), k-nearest Neighbour and Gaussian distribution; a short outline for each classifier is given below.

2.1 Auto-encoder neural network

Artificial neural networks provide a general and practical method for learning functions from examples and which are inspired by biological neural networks. ANNs basically consist of inputs (like synapses in the biological neural network), which are multiplied by weights (strength of the respective signals), and then computed by a mathematical function which determines the activation of the neuron. Another function (which is usually the identity function) calculates the output of the artificial neuron. ANNs combine some artificial neurons in order to process data and perform various tasks including classification. A full explanation of artificial neural networks is outside the scope of this paper with details and explanation given by Bishop (1995). The auto-encoder is a one-class classifier algorithm with architecture like a feed-forward neural network. It is very similar to the multilayer perceptron (MLP), with an input layer, an output layer with equally as many nodes as the input layer, and one or more hidden layers connecting them. The functions endeavour to learn an approximation to the identity function; the difference between the input and output pattern is used as a characterization of the target class. This results in:

$$f(x) = (x - NeurN(x))^2 \quad (1)$$

In which, x is the input pattern and NeurN(x) is the output of the network. The classifier then is defined as (TAX, 2001):

$$\mathrm{h(x)} = \begin{cases} \text{target} & \text{if } \mathrm{f(x)} \leq \theta \\ \text{outlier} & \text{if } \mathrm{f(x)} > \theta \end{cases} \quad (2)$$

The threshold θ is a tuning parameter set according to the target error.

2.2 *Gaussian Distribution*

Many random phenomena obey the normal distribution, at least approximately. Therefore, a Gaussian distribution can be used to model the target class in one-class classification. In other words, the Gaussian model can be used to characterise a group of samples of any number of dimensions with two values: a mean vector and a covariance matrix. This model can then be used to find the label of any unknown sample and to find out if the unknown sample belongs to the Gaussian model of training samples or not. This classifier models the training data as a Gaussian distribution using the Mahalanobis distance to model the archaeological sites as a Gaussian distribution (Equation 3):

$$f(x) = (x - \mu)^T \Sigma^{-1} (x - \mu) \qquad (3)$$

Here, x indicates the input pattern, and the mean μ and covariance matrix Σ are sample estimates. The classifier then becomes as in Equation 2. The density function for a Gaussian distribution is defined as Equation 4.

$$\mathrm{N}(\mathrm{x}; \mu, \Sigma) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\mathrm{x} - \mu)^{\mathrm{T}} \Sigma^{-1} (\mathrm{x} - \mu)\right\} \qquad (4)$$

Where, μ specifies the mean of the distribution and Σ is a d × d matrix specifying the covariance of the distribution.

2.3 *K-nearest neighbours*

The k-nearest neighbours algorithm is a method for classifying new samples based on closest training samples in the feature space. The k-nearest neighbours algorithm can be summarized as follow:

- Suppose each sample in the data set has n features which are combine to form an n-dimensional vector (Equation 5):

$$x = (x_1, x_2, .., x_n) \qquad (5)$$

- Given an unknown sample, find the k closest neighbours of this input sample using Euclidian distance function. The Euclidean distance between points x and u is defined as (Equation 6).

$$d(x, u) = \sqrt{\sum_{i=1}^{n} (x_i - u_i)^2} \qquad (6)$$

- The average of these distances is calculated and considered as f(x).
- The classifier then becomes as in Equation 2.

2.4 *K-means*

The k-means algorithm (MacQueen, 1967) is a simple unsupervised learning algorithms that has been utilized in many problem domains. In k-means clustering algorithm, n input patterns are divided into k clusters in which each pattern belongs to the cluster with the nearest mean (cluster centre). The location of cluster centres has an important effect on the final results. So, the k-means algorithm try's to place the cluster centres as distant as possible from each other. The k-means algorithm can be summarized in the following steps:

1. Initialize K points into the feature space of the training samples randomly, as initial cluster's centres.
2. Assign each training sample to the cluster that has the closest centre with regard to the Euclidian distance (Equation 6).
3. When all training samples have been assigned, recalculate the new means of each cluster and consider these means as new centres for new clusters.
4. Repeat Steps 2 and 3 until the centres no longer move. This algorithm minimizes the following error function (Equation 7).

$$j = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - c_j \right\|^2 \qquad (7)$$

In Equation 7, k and n are the number of clusters and the number of training samples, respectively. Also, $\left\| x_i^{(j)} - c_j \right\|^2$ indicates the Euclidian distance from the sample $x_i^{(j)}$ to the cluster centre c_j. Then, a new sample is characterized by:

$$f(x) = min_j (x - c_j)^2 \qquad (8)$$

In which, x is the new input sample and c_j indicate the cluster centres. The classifier then becomes as in Equation 2.

2.5 *Minimum Spanning Tree Data Descriptor*

The MST_DD (Juszczak et. al., 2009) is a non-parametric method based on graphical representation of the target training data. This method assumes that if two examples represent two similar objects in reality, not only these two mentioned points but also the other proper neighbours of these two points should be neighbours in the feature space RN. MST_DD firstly constructs a fully connected and undirected graph on training target samples.

MST_DD algorithm assigns a weight for all edges related to their lengths. A minimum spanning tree of this graph is then extracted. In the recognition phase, the shortest distance of the input pattern x to the minimum spanning tree is used as the similarity to the target class. In other words, first the distance of the input pattern to all edges is calculated and then the shortest distance is considered as the distance of the input pattern to the tree. To calculate distance of an input pattern to an edge, the projection of the new pattern x onto each edges is calculated using Equation 9:

$$P_{e_{ij}}(x) = x_i + \frac{(x_j - x_i)^T (x - x_i)}{\|x_j - x_i\|^2} (x_j - x_i) \qquad (9)$$

If $P_{e_{ij}}(x)$ lies on the edge e_{ij}, then the distance of the pattern x to the edge is computed as the Euclidian distance between x and its projection on e_{ij}. Should $P_{e_{ij}}(x)$ not lie on the edge, the distance of the input pattern x to the edge e_{ij} is calculated as the shortest Euclidian distance to one of the vertices $\{x_i, x_j\}$. Finally, the distances of the pattern x to all edges is calculated and the shortest distance to all edges selected as the distance of input pattern x to the tree.

2.6 Principal Component Analysis Data Descriptor

The missions of principal component analysis are to (1) extract the most important information from the data table; (2) reduce the dimensionality of the data by keeping only the important information; (3) simplify the description of the data set; and (4) visualize and analyze the structure of the data and the variables. However, here it has been utilized to describe the archaeological site data by a linear space. Then, the difference between an original new object and the projection of that new object onto the linear space (in the original data) is calculated and used for classification.

In PCA, the criterion is maximizing variance. The principal component is W_1 such that the sample, after projection onto W_1, has maximum spread so that difference between the sample points becomes most apparent (Alpaydin, 2004). In other words, this method describes the target data by a linear subspace with this subspace defined by eigenvectors of the data covariance matrix Σ. The projection is shown in Equation 10.

$$f(x) = W(W^T W)^{-1} W^T x \qquad (10)$$

Where, W indicates a d×k matrix that includes k eigenvectors of the data covariance matrix. Then, the f(x) function is defined as squared distance from the original sample and its mapped version (TAX, 2001):

$$f(x) = ||x - x_{proj}||^2 \qquad (11)$$

In which, x is the new input pattern and x_{proj} is projection of this object onto the subspace (in the original data). The classifier then becomes as in Equation 2.

3 EXPERIMENTAL RESULTS AND DISCUSSION

As highlighted in Section 2, Methodology, two procedures were used to obtained samples from non-archaeological sites: 1) Generating non-archaeological test samples using a block-shaped uniform distribution that covers the target data (known archaeological samples), 2) Generating non-archaeological samples using field studies of the Khorramabad Plain. The first procedure gives a data set containing 43 previously known archaeological samples and 43 artificially generated non-archaeological samples is obtained and used to train, validate and test the models using a nested 10-fold cross validation. Nested 10-fold cross validation includes two loops; in the inner loop the training data is partitioned into 10 parts in equal sizes, then 9 of the parts are used to train (optimize the parameters of) the model and evaluated on the remaining part. This procedure is repeated for all 10 possible choices for the held-out part and the performance scores from the 10 runs are averaged. The outer loop is executed three times and each time it chooses a different 30 per cent of the whole data and allows the other 70 per cent to be used in the inner loop. The average performance over these three test sets for different classifiers is represented in Table 2. It is worth noting that to build the one class classifiers only target data are used with the outliers used for evaluating and testing the models. All algorithms were implemented in Matlab using ddtools package (Tax, 2014). Each experiment was repeated 10 times and the results in Table 2 are averaged.

In Table 2, the experimental evaluation of the proposed models are represented, based on different measures including False Positive Rate (FPR), False Negative Rate (FNR), Area Under Curve (AUC), Precision and Recall (Alpaydin, 2004). Here, the archaeological site samples are considered as positive samples (target samples) and other samples are considered (non-target) as negative samples. False Positive Rate, False Negative Rate, Precision and Recall are defined in equations 12, 13, 14 and 15, respectively. To fine-tune and to evaluate a classifier, another approach is to calculate the area under the curve (AUC), the receiver operating characteristics (ROC) curve. ROC curve shows recall versus false positive rate for different values of related parameter.

True Positive (TP) is the number of times that the predictive model classifies an input sample as an archaeological site correctly. True Negative (TN) stands for the number of times that the system classifies a sample as a non-archaeological site sample correctly. Similarly, False Positive (FP) refers to the number of times that the predictive model classifies a sample as an archaeological site wrongly. False Negatives (FNs) is the number of times that the system classifies a sample as a non- archaeological site sample wrongly.

$$False\ Positive\ Rate = \frac{FP}{FP+TN} \tag{12}$$

$$False\ Negative\ Rate = \frac{FN}{TP+FN} \tag{13}$$

$$Precision = \frac{TP}{TP+FP} \tag{14}$$

$$Recall = \frac{TP}{TP+FN} \tag{15}$$

Table 2: Experimental results based on different measures including False Positive Rate (FPR), False Negative Rate (FNR), Area Under Curve (AUC), Precision and Recall.

Classifier	FNR	FPR	Precision	Recall	AUC
Minimum Spanning Tree data Descriptor	0.04	0.09	0.91	0.96	0.98
k-nearest neighbour	0.07	0.12	0.90	0.93	0.98
k-means	0.12	0.07	0.93	0.88	0.97
auto-encoder	0.14	0.09	0.92	0.86	0.97
PCA	0.16	0.07	0.93	0.84	0.95
Gaussian Target Distribution	0.20	0.11	0.90	0.79	0.84

Both precision and recall measures are sometimes used together in the F1-measure to provide a single measurement for a system. The F1-measure, represented in Equation 16, can be interpreted as a weighted average of the precision and recall, where an F1-measure reaches its best value at 1 and worst score at 0. To highlight the efficacy of pattern recognition methods in detection of buried archaeological sites, Figure 4 shows a comparison of F1-Measure values for implemented methods.

$$\mathrm{F1-measure} = \frac{\mathrm{2.precision.recall}}{\mathrm{precision+recall}} \tag{16}$$

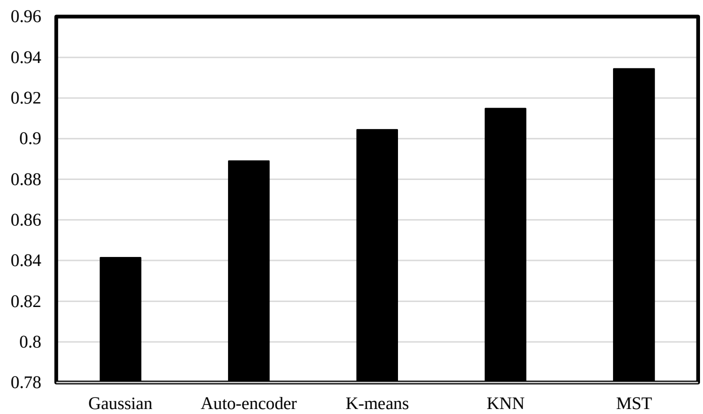
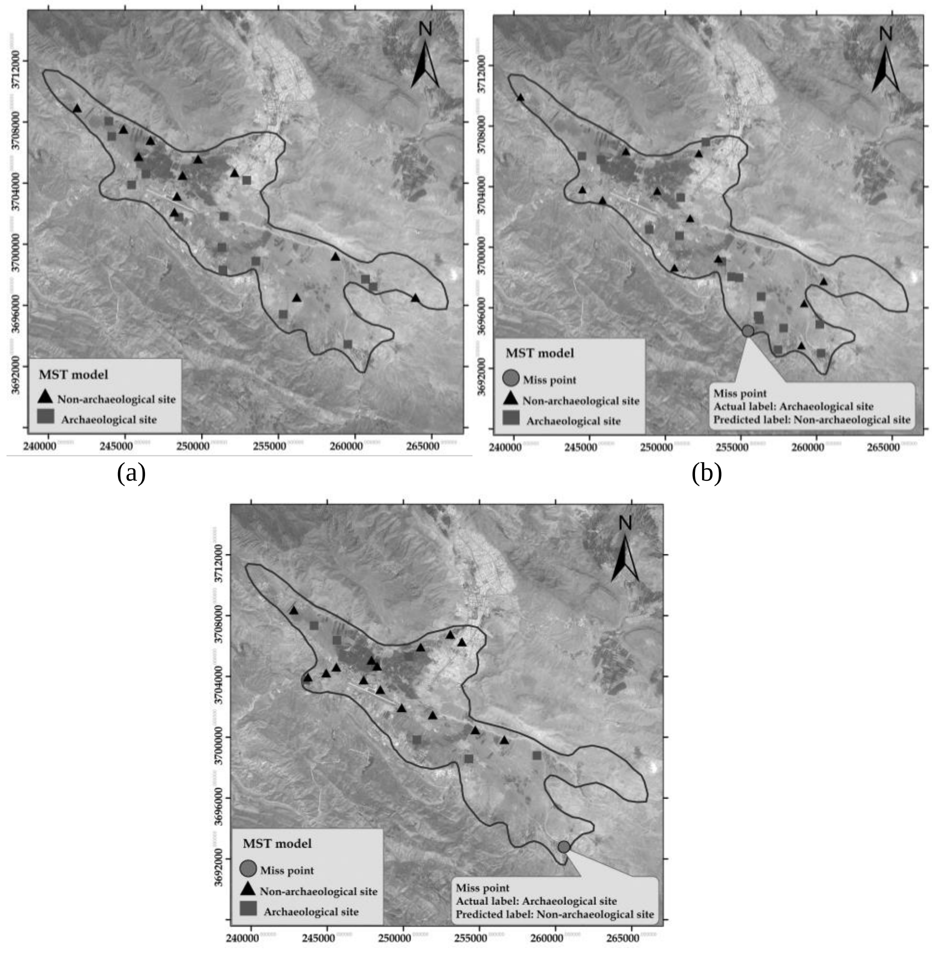


Figure 4) Comparison of F1-Measure values for implemented methods.

It is apparent from Table 2 and Figure 4 that the applied one-class classifiers have a promising performance in building predictive models for detecting buried archaeological sites. With regard to the experiments represented in Table 2, the minimum spanning tree data descriptor has the best performs in comparison to other classifiers. In this classifier, targets and edges are effectively classified, and even neighbourhoods of the (graph) edges can be considered as target classes and are an additional set of virtual target objects. These additional objects, in turn, can help model a target distribution in multi-dimensional spaces and where small sample sizes can otherwise be problematic. Positively, the experiments of this paper corroborate with previous research on MST_DD features with the MST_DD classifier performing well in multi-dimensional spaces and in small sample size problems in comparison to other existing one-class classifiers (Juszczak et. al., 2009).

As mentioned above, we use a second experiment to generate real world non-archaeological test samples in addition to artificially generated non-archaeological test samples. To further evaluate the MST_DD classifier and to visualize its performance, the trained MST_DD model is assessed using a new test set generated using field study of the Khorramabad Plain. 45 points of the plain including the bed of current rivers exposed by water and other areas excavated for building different facilities including roads, tunnels and transects are considered as outliers (non-archaeological samples). These 45 samples are randomly divided into 3 parts of 15 samples. Each of these parts is combined with 14 randomly extracted samples from archaeological site samples. Consequently, three different test sets are generated, each with 29 samples. The test sets are entered into the trained MST_DD model and predicted model results together with the real labels of the test set samples are represented in Figure 5. As is evident from Figure 5, the MST_DD algorithm effectively represents results from the test sets. The MST_DD algorithm is able to recognize all archaeological and non-archaeological sites correctly and did not miss any samples in Figure 5.a. In both Figures 5.b and 5.c the classifier missed only one sample from target samples and predict all non-archaeological sites correctly.



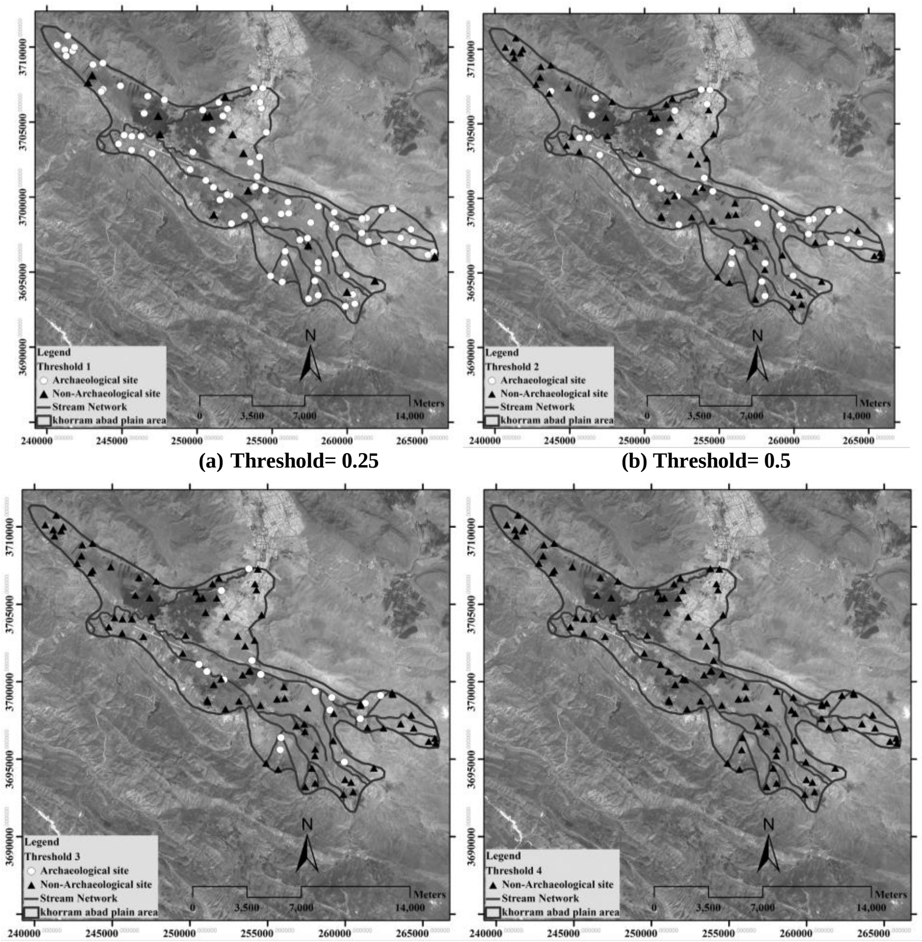
(a) (b)

(c)

Figure 5) Performance of the MST_DD model on three different test sets.

One of the main goals of this experiment is the attempt to study the properties and possible advantages of using pattern recognition methods in archaeology and consequently find new ways to identify new archaeological sites locations by using one-class classification techniques. Experiments have also been carried out to generate a map showing where previously undiscovered sites might be and with varying degrees of certainty. To do so, 100 co-ordinates were generated randomly from the study area using Arc GIS; MST_DD (see Table 2 and Figure 4) was then utilized to investigate the randomly generated coordinates. Figure 6 shows four different maps that represent some of the randomly generated co-ordinates as potential locations for previously undiscovered archaeological sites, with 4 different degree of certainty. Different degrees of certainty were obtained by varying MST_DD's threshold and which is a tuning parameter for the MST_DD model. The MST_DD's threshold determines the training target samples that are allowed to be rejected and classified as outliers during the training process. The MST_DD's threshold was varied from 0.01 to1 by steps equal to 0.01 and the results observed. In practice, the MST_DD model is trained and created using previously known archaeological sites (43 samples) with 100 different thresholds in the interval between [0, 1], and then the trained MST_DD models are used to label each of the randomly generated coordinates as archaeological site or non-archaeological site. As a result, 100 different maps with 100 degree of certainty are generated; each of them introduces some of the randomly generated coordinates as previously undiscovered sites. The interval between [0,1] is divided into 4 equal sub-intervals and Figure 6 shows 4 of the generated maps for Thresholds equal to 0.25, 0.5, 0.75 and 1.

Based on the outcomes presented in Figure 6, as the threshold becomes larger, less randomly generated samples are considered as archaeological sites. The effect of the threshold values on the results in Figure 6 is evident. The smaller threshold would result in higher numbers of predicted sites (and higher true positives) but lower degree of certainty (and higher false positives) are attained and is an appropriate approach when a lower, conservative, prediction is sought. Conversely, the larger threshold would result in lower number of proposed sites (and lower false positives) but higher degree of certainty (and lower true positives) and is appropriate when more speculative approaches are required. The control parameter, threshold, can also be used to achieve a balance between high number of proposed site and low degree of certainty and is perhaps its most useful application.



(a) Threshold= 0.25 (b) Threshold= 0.5

(c) Threshold= 0.75 (d) Threshold= 1

Figure 6) Possible location of previously undiscovered sites with various degree of certainty.

Our analyses of the Khorramabad Plain establishes that there are common environmental constraints on settlement patterns from prehistory through to the Islamic - historical period, making it possible to efficiently model archaeological site distribution by pattern recognition algorithms. This includes the first village based settlements from ca. 7500 BP with an economy based on domestic livestock (goats and ewes) and later arable based settlements that introduced wheat, barley, lentils and flax to the region (Table 1; Javadi and Borazjani, 2000; Hole, 1970).

Artificial Intelligence is attracting widespread interest in many sciences because of its emerging robust predictive capabilities. AI enables archaeologist to more fully exploit knowledge from extensive amount of archaeological data and assists archaeologists in reasoning and making decisions that range from appropriate conservation and protection strategies to where best to excavate in a complex cultural landscape. The experimental results of this paper provides clear evidence that the application of Pattern Recognition has real potential as an effective AI application for the detection of buried archaeological sites. It can ensure that archaeologist avoid expensive and time consuming efforts to survey and excavate more archaeologically limited landscape areas. Although in this study MST_DD represents an indisputably better performance compared with other one-class classifiers, we also highlight that there is no single method that for any data set represents the most accurate method (the No Free Lunch Theorem; Alpaydin, 2004). Although the MST_DD model may be successfully applied as a predictive model in other semi-arid area, especially when there is small number of previously identified sites and where there has been substantial accumulation of eroded soils, we would suggest the approach developed by this paper. That is, an investigation and testing of a range of algorithms with selection based on best performance against a test data set.

4 CONCLUSIONS

Artificial intelligence, in particular one-class classifiers hold promise when applied to the detection of buried archaeological sites. One-class classifiers were trained using extracted data (using Arc GIS) from previously identified archaeological sites of Khorramabad Plain and then able to automatically classify unseen input patterns as potentially archaeological sites (for further investigation) or non-archaeological areas. The results indicated that application of one-class classification methods, and in particular the minimum spanning tree data descriptor, construct efficient predictive models for semi-arid areas. We now anticipate that our findings can be reliably applied in other study areas without a significant degradation in performance. Our future work will now explore the application of one-class classifiers using data sets from other study areas and environments. To aid these endeavours we recommend that a plug-in software of the proposed one-class classifiers for Arc GIS software should be built.

ACKNOWLEDGEMENTS

This work is supported by the Academic Centre for Education, Culture and Research (ACECR) of Lorestan. Moreover, parts of this research have been funded by The Catalan Government-Generalitat de Catalunya (2014 SGR 1169), The Spanish Ministry of Science and Innovation, through Grant No.HAR2012-31036 awarded to J.A. Barceló and Project CSD2010-00034 «Social and environmental transitions: Simulating the past to understand human behavior (CONSOLIDER-INGENIO 2010 program by Spanish Ministry of Science and Innovation, see: http://www.simulpast.es), and Fundació LaCaixa (RecerCaixa project 2015ACUP 00191).

REFERENCES

Alpaydin E. 2004. *Introduction to Machine Learning*. MIT Press, Cambridge; 584 pp.

Ayala, G, French C. 2005. Erosion modelling of past land-use practices in the Fiumedi Sotto di Troina River Valley, North-Central Sicily. *Geoarchaeology* **20**: 149-167.

Balla A, Pavlogeorgatos G, Tsiafakis D, Pavlidis G. 2014. Efficient predictive modelling for archaeological research. *Mediterranean Archaeology and Archaeometry* **14**: 119-129.

Barceló J A. 2010. Computational Intelligence in Archaeology. State of the Art. In *Making History Interactive. Computer Applications and Quantitative Methods in Archaeology (CAA).* Frischer B. Webb C J, Koller D. (eds.) Archaeopress: Williamsburg, pp. 11-21.

Barton C. M, Bernabeu J, Aura J E, Garcia O, Roca N L. 2002. Dynamic landscapes, artefact taphonomy, and land use modelling in the western Mediterranean. *Geoarchaeology* **17**: 155-190.

Barton C M, Ullah I, Mitasova H. 2010. Computational modelling and Neolithic socio ecological dynamics: a case study from southwest Asia. *American Antiquity* **75**: 364-386.

Bishop C M. 2006. *Pattern Recognition and Machine Learning*. Springer, New York; 738 pp.

Bishop C M. 1995. *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford; 504 pp.

Butzer K W. 1982. *Archaeology as Human Ecology: Method and Theory for a Contextual Approach*. Cambridge University Press, Cambridge; 380 pp.

Deravignone L, Jánica G M. 2006. Artificial Neural networks in Archaeology. *Archaeologia e Calcolatori* **17**: 121-136.

Ebert D. 2004. *Predictive modelling and the ecology of hunter- gatherers of the boreal forest of Manitoba*. Archaeopress, Oxford; 81 pp.

Fernandes R, Geeven G, Soetens S, Klontza-Jaklova V. 2011. Deletion/Substitution/Addition (DSA) model selection algorithm applied to the study of archaeological settlement patterning. *Journal of Archaeological Science* **38**: 2293-2300.

Gibbon G. 2000. Appendix A: Archaeological Predictive Modelling: An Overview. In *A Predictive Model of Precontact Archaeologial Site Location for the State of Minnesota*, Hudak G J, Hobbs E, Brooks A, Sersland C A, Phillips C, Paul S T (Eds.). Minnesota Department of Transportation, Minesota.

Han J, Kamber M, Pei J. 2006. *Data Mining: Concepts and Techniques*, Morgan Kaufmann; 744 pp.

Hole F. 1970. The Palaeolithic culture sequence in Western Iran. Actes du VII Congrès International des Sciences Préhistoriqueset Protohistoriques, Pruge, pp. 286-292.

Jain A K, Duin R P W, Mao J. 2000. Statistical Pattern Recognition: A Review. *IEEE Transaction on Pattern Analysis and Machine Intelligence* **22**: 4-37.

Javadi M R, Borazjani N V. 2000. Preliminary identification of hills and caves in Khorram abad valleys, Tehran, Technical Assistance Office to preserve and revive the cultural heritage of Iran.

Juszczak P, Tax D M J, Pekalska E, Duin R P W. 2009. Minimum spanning tree based one-class classifier. *Neurocomputing* **72:** 1859-1869.

Kamermans H. 2006. Searching for Tools to Predict the Past; the Application of Predictive Modelling in Archaeology. In *Reading Historical Spatial Information from around the World: Studies of Culture and Civilization Based on Geographic Information Systems Data*. Uno T (Ed.) International Research Centre for Japanese Studies, Kyoto, pp. 35-46.

Kamermans H, Rensink E. 1999. GIS in palaeolithic archaeology. A case study from the southern Netherlands. In *Archaeology in the age of internet - CAA97. Computer Applications and Quantitative Methods in Archaeology BAR International Series.* : Dingwall L, Exon S, Gaffney V, Laflin S, Leusen M van (Eds.) Archaeopress, Oxford, **81**: 20-33.

Khan S S, Madden M G. 2010. A Survey of Recent Trends in One Class Classification. *Lecture Notes in Computer Science* **6206**:188-197.

Konnie, L. Wescott, R. Joe, B., (2000). *Practical Applications of GIS for Archaeologists a Predictive Modelling Toolkit*, , , CRC Press, Boca Raton; 176 pp.

Kuiper J A, Wescott K L. 1999). A GIS Approach for Predicting Prehistoric Site Locations, Presented at the Nineteenth Annual Esri User Conference, San Diego, California, pp. 1-12.

Kvamme K L. 1990. The Fundamental Principles and Practice of Predictive Archaeological. *Mathematics and information science in archaeology: A flexible framework* **3**: 257-295.

Lang N. 2000. Harmonising research and Cultural resource Management. In *Beyond the map: Archaeological and Spatial Technologies*, Lock, G (Ed.) IOS Press, pp. 214-227.

MacQueen J B. 1967. Some Methods for classification and Analysis of Multivariate Observations, Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, **1**: 281-297.

Moya M M, Hush D. 1996. Network Constraints and Multi-objective Optimization for One-class Classification, *Neural networks* **9**: 463-474.

Myrsini G, Gert J V W, Steven S. 2011. Assessing the effects of geomorphological processes on archaeological densities: a GIS case study on Zakynthos Island, Greece. *Journal of Archaeological Science* **38**:. 2714-2725.

Nabney I T. 2004. *Matlab Algorithms for Pattern Recognition*. Springer, London,; 420 pp.

Osareh A. 2004. Automated identification of diabetic retinal exudates and the optic disc, PhD Thesis, University of Bristol, Bristol.

Pimentel M A F, Clifton D A, Clifton L, Tarassenko L. 2014. A review of novelty detection, *Signal Processing*. **99**: 215-249.

Perumal M, Velumani B, Sadhasivam A, Ramaswamy K. 2015. Spatial Data Mining Approaches for GIS – A Brief Review. *Advances in Intelligent Systems and Computing*. **338**: 579-592.

Puyol-Gruiart J. 1999. Computer Science, Artificial Intelligence and Archaeology, In *New Techniques for Old Times. CAA98. Computer Applications and Quantitative Methods in Archaeology.* : Barceló J A, Briz I, Vila A. (Eds.) Proceedings of the 26th Conference, Barcelona, March 1998 (BAR International Series 757). Archaeopress, Oxford, pp. 19-28.

Richards J D. 1998. Recent Trends in Computer Applications in Archaeology. *Journal of Archaeological Research* **6**: 331-382.

Schiffer M B. 1983. Toward the identification of formation processes. American settlement patterns based on surficial artifact distribution: Replacing humans on the landscape. In *Earth Sciencesand Archaeology*. Goldberg P, Holliday V T, Ferring R. (Eds.) Kluwer Academic/Plenum Publishers, New York, pp. 107-141.

Tartaron T F, Gregory T E, Pullen D J, Noller J S, Rothaus R M, Rife J L, Tzortzopoulou-Gregory L, Schon R, Caraher W R, Pettegrew D K, Nakassis D. 2006. The eastern Korinthia archaeological survey e integrated methods fora dynamic landscape (Corinthian plain). *Hesperia*, **75**: 453-523.

Tax D M J. 2001. One-class classification. M.Sc. Thesis, Delft University of Technology; 202 pp.

Tax, D.M.J., (2014). DDTools, the Data Description Toolbox for Matlab. available at: [http://prlab.tudelft.nl/david-tax/dd_tools.html], accessed : September 08 2015.

Verhagen Ph., Kamermans H, Leusen M V. 2009. The future of archaeological predictive modelling. In *Archaeological Prediction and Risk Management. Alternatives to current practice.* Kamermans H, Leusen M V, Verhagen Ph (Eds.) Leiden University Press, Leiden, pp. 19-25.

Verhagen Ph, Kamermans H, Leusen M V, Ducke B. 2010. New developments in archaeological predictive modelling. In *The Cultural Landscape & Heritage Paradox. Protection and Development of the Dutch Archaeological-Historical Landscape and its European Dimension.* Kars H, Valk A V, Wijnen M., Bloemers T. (Eds.) Amsterdam University Press, Amsterdam, pp. 429-442.

Vitali V. 1991. Formal methods for the analysis of archaeological data: Data analysis vs expert systems. In *Computer Applications and Quantitative Methods in Archaeology*, Lockyear K, Rahtz S. P. Q. (eds.) BAR International Series, Oxford, pp. 207-209.

Voorrips A. 1990. New Tools from Mathematical Archaeology, Presented at the 5th International Symposium on Data Management and Mathematical Methods in Archaeology, Scientific Information Centre of the Polish Academy of Sciences, Warsaw, Poland.

Wells L. 2001. A geomorphological approach to reconstructing archaeological settlement patterns based on surficial artifact distribution: Replacing humans on the landscape. In *Earth Sciences and Archaeology*. Goldberg P, Holliday V T, Ferring R. (Eds.) Kluwer Academic/Plenum Publishers, New York, pp. 107-141.