Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
ISPRS Journal of Photogrammetry & Remote Sensing 60 (2006) 225–238
PHOTOGRAMMETRY & REMOTE SENSING
www.elsevier.com/locate/isprsjprs

The performance of pixel window algorithms in the classification of habitats using VHSR imagery

Iphigenia Keramitsoglou [a,*], Haralambos Sarimveis [b], Chris T. Kiranoudis [b], Charalambos Kontoes [c], Nicolaos Sifakis [c], Eleni Fitoka [d]

[a] *University of Athens, Department of Applied Physics, Remote Sensing and Image Processing Team, Panepistimioupolis, Build. PHYS-V, GR-15784, Athens, Greece*
[b] *School of Chemical Engineering, National Technical University of Athens, Zografou Campus, Athens, GR-15780, Greece*
[c] *Institute for Space Applications and Remote Sensing, National Observatory of Athens, Metaxa and Vas. Pavlou St, Pendeli, Athens, GR-15236, Greece*
[d] *The Goulandris Natural History Museum, Greek Biotope/Wetland Centre (EKBY), 14th km Thessaloniki, Mihaniona, 57001 Thermi, Thessaloniki, Greece*

Received 2 August 2005; received in revised form 3 January 2006; accepted 3 January 2006
Available online 20 March 2006

Abstract

This study investigates the potential of three advanced pixel window classification methods for habitat mapping, namely Kernel based spatial Re-Classification (KRC), Radial Basis Function (RBF) neural networks (NN) and Support Vector Machines (SVM). KRC classifier takes into account the spatial arrangement and frequency of spectral classes present within a predefined square kernel. On the other hand, RBF-NN and SVM classifiers use a set of spectral parameters (digital numbers of training pixels, mean values and standard deviations within a specified window kernel) as input information. The fuzzy means clustering algorithm is utilized for training the RBF networks. This method is based on a fuzzy partition of the input space and requires only a short amount of time to determine both the structure and the parameters of the RBF-NN classifier. The radial basis function is also adopted as the kernel function in the implementation of the SVM methodology. The test area of the present study is Lake Kerkini, a wetland ecosystem located in Macedonia (Northern Greece). The methods are applied to a very high spatial resolution multispectral satellite image acquired by IKONOS-2. The nomenclature used is EUNIS, a detailed hierarchical habitat classification scheme. Several classification experiments are carried out using the same training samples in order to study the behaviour of the three classifiers and perform meaningful comparisons. Overall, all three classifiers performed satisfactorily; however the SVM and RBF-NN classifiers consistently outperformed KRC, reaching overall accuracies of 72% and 69%, respectively.

Keywords: habitat classification; RBF neural networks; kernel based re-classification; support vector machines; EUNIS

1. Introduction

Intensification of agriculture, tourism and fragmentation are only a few of the serious threats imposed to natural and semi-natural habitats. In most of the EU member states conservation of biotopes is a high priority

* Corresponding author.
E-mail addresses: ikeram@phys.uoa.gr (I. Keramitsoglou), hsarimv@central.ntua.gr (H. Sarimveis), kyr@chemeng.ntua.gr (C.T. Kiranoudis), kontoes@space.noa.gr (C. Kontoes), sifakis@space.noa.gr (N. Sifakis), helenf@ekby.gr (E. Fitoka).

0924-2716/$ - see front matter © 2006 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.isprsjprs.2006.01.002

issue in environmental policy, making the need to be able to detect changes in the natural environment increasingly pressing. The conservation status of protected habitats is analysed using habitat maps and spatial indicators to assess fragmentation, spatial distribution and neighbourhood relations of key habitats. Temporal dynamics and biophysical parameters of habitats and surrounding land cover can be examined to gain a better understanding of the pressures that affect protected habitats. The ecological significance of indicators is assured through empirically sound investigations of the relationship of the indicators with biodiversity data on the species level. Present and potential users addressed are local and regional environmental authorities, forest departments, environmental information centres and the European Environmental Agency as well as environmental authorities on the European level (www.spin-project.org).

However, reduced financial resources have raised the awareness of ecologists and conservation biologists for innovative techniques for habitat mapping. A powerful suite of tools and data can now be found in the field of Earth Observation. The strength of this approach is the spatial and temporal consistency of satellite data as well as cost effectiveness (Kerr and Ostrovsky, 2003; Bock, 2003).

The spatial resolution of satellite imagery suitable for earth resources studies has improved from 80 m (Landsat-MSS, launched in 1972) to 0.6 m (QuickBird, launched in 2001), i.e. more than 10,000 times by area, during a quarter of a century. The need of remote sensing community for very high spatial resolution images (VHSR; ground sampling distance of sensor smaller than 10 m) has now been satisfied, yet there is a limitation in image classification, as conventional pixel based methods developed for medium to low spatial resolution images do not seem to be suitable for VHSR images. In VHSR images the spectral responses of particular habitat classes are much more variable as being composed of the spectral responses of individual class elements (intra-class spectral variability). This results in the 'scene-noise' problem. Thus, it is needed to employ new techniques that take into account not only the spectral signature of an individual pixel but also the spatial features extracted from the vicinity of the pixel, within a specified pixel window (Zhang et al., 2003). Three different pixel window classifiers are examined specifically for habitat mapping in the present paper.

The first method is kernel based re-classification software developed by Barnsely and Barr (1996). Originally the concept of land use classification using a cover frequency method was introduced by Gong and Howarth (1992) who appreciated the weaknesses of pixel based classification of high spatial resolution images and applied it to SPOT-HRV images. The kernel re-classification algorithm used in the present study derives information on wetland classes in two stages. The first step is to transform the original multispectral image into a single channel image. This is usually achieved by pixel based supervised or unsupervised clustering. The number of initial classes is not fixed and may vary from six to twelve. The kernel re-classifier is then applied to the transformed image and the pixel labels are grouped into discrete land cover (or habitat) classes on the basis of their frequency of occurrence and spatial arrangement within a specified pixel window (kernel). The overall classification accuracy reported until now in the literature varies from 74% (Kontoes et al., 2000) to 96% (Barnsely and Barr, 1996) using a two-date multispectral set of IRS-1C LISS-III and Pan, and a single multispectral SPOT-1 HRV image, respectively. It is important to note that both studies address the urban environment. KRC examines labels of adjacent pixels within the square kernel and calculates the so-called adjacency event matrix, accounting for the spatial arrangement and frequency of the labels. Criterion for pixel re-labelling is the degree of matching between the adjacency event matrix and the Template Matrices produced during training. Thus, the algorithm accounts for texture and spectral components of the information classes.

The second approach is based on the radial basis function (RBF) neural network architecture (Moody and Darken, 1989). RBFs constitute a special type of artificial neural networks, which has certain advantages over other network types such as the Feedforward Neural Networks (FNNs), including simpler network configurations and faster training procedures. Neural networks have been utilized extensively in solving image classification problems (Abuelgasim et al., 1996; Chettri et al., 1992; Decatur, 1989; Hepner et al., 1990; Kanellopoulos et al., 1990). Surprisingly, the applications of the RBF architecture in solving this type of problems are very few. In a previous work (Keramitsoglou et al., 2003) we have compared RBF-NN classifier with maximum likelihood classifier (MLH). For the purposes of the comparison, the multispectral IKONOS-2 image was resampled to 11 m spatial resolution. Resampling allowed averaging of the responses of different spectral classes (e.g. vegetation or water) and registration as one pixel value producing a coarser resolution image. This was necessary for a meaningful application of MLH classification, as per-pixel classification methods are prone to generate significant misclassification errors when applied to very high spatial resolution imagery for mapping complex information

classes, as already explained. In all comparison experiments the neural network classifiers performed better overall than MLH, in some cases by as much as 17%. For the training of the neural network classifiers, a method based on the fuzzy means algorithm (Sarimveis et al., 2002) was used.

Support Vector Machines (SVM) is the third classification method that is studied in this work. It is a supervised learning technique rooted in the Statistical Learning Theory developed by Vladimir Vapnik and coworkers at AT and T Bell Laboratories (Vapnik, 1995, 1998), which is gaining popularity due to many attractive features and promising empirical performance (Schölkopf and Smola, 2002). Originally the SVM method was worked out for linear two-class classification with margin, where margin means the minimal distance from the separating hyperplane to the closest data points. SVM learning machine seeks for an optimal separating hyperplane, where the margin is maximal. An important and unique feature of this approach is that the solution is based only on the marginal data points, called support vectors. The linear SVM can be extended to non-linear one using a set of non-linear basis functions. Several successful application of SVMs in image classification have been reported in the literature (Camps-Valls and Bruzzone, 2005; Foody and Mathur, 2004a,b; Kim et al., 2002).

The present work exploits the potential of using the three previously introduced pixel window classifiers, namely KRC, RBF-NN and SVM, *specifically for habitat mapping* using 4-m spatial resolution multispectral satellite imagery. The test area is the international importance wetland of Lake Kerkini in Macedonia (Northern Greece) and the objective is to classify it into five classes as determined by the European Nature Information System (EUNIS) developed by the European Environment Agency (EEA, 2002).

The rest of the paper is structured as follows: In Section 2, the three different classification approaches are presented. Section 3 gives the characteristics of the study area as well as the data used for this study, whilst Section 4 describes the classification experiments and the results. Finally, the conclusions drawn are given in Section 5.

2. Specific methodological aspects

2.1. Kernel re-classification (KRC) algorithm

The basic input datasets for the kernel re-classification algorithm include a very high spatial resolution satellite image and a ground truth map of the same year and season. The user then selects the training sets in order to provide information on the different classes. For each final class, a number of training samples are required in order to include all different phenomenal representations of the class (e.g., dense and sparse riparian forest). The first step is to perform an initial unsupervised clustering of the original image into a few clusters; in our case, eight clusters are sufficient to represent the spectral variation of the scene. For unsupervised clustering, an Iterative Self-Organising Data Analysis Technique Algorithm (ISODATA); (Tou and Gonzalez, 1974) is used iteratively until either a 0.95 convergence threshold is achieved or 10 iterations are complete.

Following the definition of training clusters within the originally classified scene, the signature files are generated. These files include the adjacency event matrices, defined according to (Barnsely and Barr, 1996) as

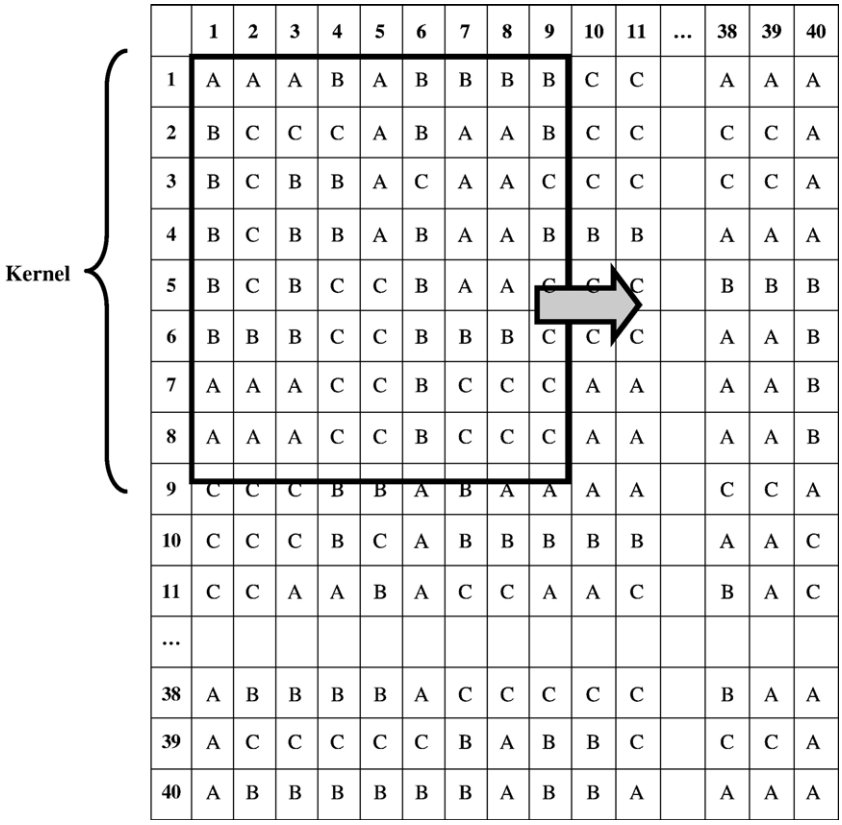
$$M = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \cdots & f_{ij} & \cdots & \cdots \\ f_{n1} & f_{n2} & \cdots & f_{nn} \end{bmatrix}, \quad (1)$$

where f_{ij} is the frequency with which pixels belonging to class i are adjacent to those belonging to class j ($i, j = 1, \ldots, n$) within the specified kernel. From the definition, in Eq. (1) $f_{ij}=f_{ji}$. The size of the square matrix M depends on the number of classes present in the initially classified image (n). The size of the signature file, on the other hand, depends on the kernel size, as it contains all non-overlapping matrices M that can fit in the training cluster. For instance, if the number of classes is three (Fig. 1) – namely A, B and C – the training cluster is 40×40 pixels and the kernel size is 9, then the corresponding signature file contains sixteen 3×3 adjacency event matrices. The total number of signature files equals the number of training clusters, as there is usually more than one cluster per final class.

Following the generation of signature files, the algorithm extracts the mean training adjacency event matrices taking into account the most representative ones defining the so-called template matrix (signature adjacency matrix). The final stage of the process is the reclassification of the initially classified pixels to the final habitat class which is selected as the one with the highest similarity index, defined as:

$$\Delta_k = 1 - \sqrt{0.5N^{-2} \sum_{i=1}^{C} \sum_{i=j}^{C} (M_{ij} - T_{kij})^2}, \quad (2)$$

where M_{ij} is defined from Eq. (1), T_{kij} is the corresponding element for the template matrix of habitat class k, N is the total number of adjacency events in the kernel and C is the number of final habitat classes in the



	1	2	3	4	5	6	7	8	9	10	11	…	38	39	40
1	A	A	A	B	A	B	B	B	B	C	C		A	A	A
2	B	C	C	C	A	B	A	A	B	C	C		C	C	A
3	B	C	B	B	A	C	A	A	C	C	C		C	C	A
4	B	C	B	B	A	B	A	A	B	B	B		A	A	A
5	B	C	B	C	C	B	A	A	C	C	C		B	B	B
6	B	B	B	C	C	B	B	B	C	C	C		A	A	B
7	A	A	A	C	C	B	C	C	C	A	A		A	A	B
8	A	A	A	C	C	B	C	C	C	A	A		A	A	B
9	C	C	C	B	B	A	B	A	A	A	A		C	C	A
10	C	C	C	B	C	A	B	B	B	B	B		A	A	C
11	C	C	A	A	B	A	C	C	A	A	C		B	A	C
…															
38	A	B	B	B	B	A	C	C	C	C	C		B	A	A
39	A	C	C	C	C	C	B	A	B	B	C		C	C	A
40	A	B	B	B	B	B	B	A	B	B	A		A	A	A

Fig. 1. A schematic example of 40×40 pixels image initially classified into 3 classes (A, B and C). A kernel of 9×9 pixels is also illustrated (Keramitsoglou et al., 2003).

image. From the above it is clear that the values of Δ_k fall in the range between 0 (no similarity) and 1 (perfect match), i.e., $0 \leq \Delta_k \leq 1$.

During the final re-classification stage a kernel of defined size passes over the whole image and computes the so-called similarity index between the current and the training adjacency event matrices. The algorithm assigns to each pixel the thematic habitat class for which Δ_k is maximum. It also provides as a standard output product a post-classification probability map, each pixel of which equals the maximum Δ_k. The post-classification probability map has the same number of lines and columns as the classified image, and is necessary to validate the results of re-classification. It can subsequently be used to create a mask of the areas re-classified with confidence larger than a threshold value specified by the user.

2.2. Radial basis function (RBF) neural networks

In this section we present the basic characteristics of the RBF neural network architecture and the training method that was used to develop the image classification neural network models.

2.2.1. RBF network topology and node characteristics

RBF networks consist of three layers: the input layer, the hidden layer and the output layer. The input layer collects the input information and formulates the input vector $\mathbf{x}$. The hidden layer consists of L hidden nodes, which apply non-linear transformations to the input vector. The output layer delivers the neural network responses to the environment. A typical hidden node l in an RBF network is described by a vector $\hat{\mathbf{x}}_l$, equal in dimension to the input vector and a scalar width σ_l. The activity $v_l(\mathbf{x})$ of the node is calculated as the Euclidean norm of the difference between the input vector and the node centre and is given by:

$$v_l(\mathbf{x}) = ||\mathbf{x} - \hat{\mathbf{x}}_l|| \tag{3}$$

The response of the hidden node is determined by passing the activity through the radially symmetric Gaussian function:

$$f_l(\mathbf{x}) = \exp\left(-\frac{v_l(\mathbf{x})^2}{\sigma_l^2}\right) \tag{4}$$

Finally, the output values of the network are computed as linear combinations of the hidden layer responses:

$$\hat{y}_m = g_m(\mathbf{x}) = \sum_{l=1}^{L} f_l(\mathbf{x}) w_{l,m} \quad m = l, \ldots, M \tag{5}$$

where $[w_{1,m}, w_{2,m}, \ldots, w_{L,m}]$ is the vector of weights, which multiply the hidden node responses in order to calculate the mth output of the network.

2.2.2. RBF network training methodology

Training methodologies for the RBF network architecture are based on a set of input–output training pairs ($\mathbf{x}(k)$; $\mathbf{y}(k)$) ($k=1,2,\ldots,K$). The training procedure used in this work consists of three distinct phases:

A) Selection of the network structure and calculation of the hidden node centres using the fuzzy means clustering algorithm (Sarimveis et al., 2002). The algorithm is based on a fuzzy partition of the input space, which is produced by defining a number of triangular fuzzy sets in the domain of each input variable. The centres of these fuzzy sets produce a multidimensional grid on the input space. A rigorous selection algorithm chooses the most appropriate knots of the grid, which are used as hidden node centres in the produced RBF network model. The idea behind the selection algorithm is to place the centres in the multidimensional input space, so that there is a minimum distance between the centre locations. At the same time the algorithm assures that for any input example in the training set there is at least one selected hidden node that is close enough according to a distance criterion. It must be emphasized that opposed to both the *k*-means (Darken and Moody, 1990) and the *c*-means clustering (Dunn, 1973) algorithms, the fuzzy means technique does not need the number of clusters to be fixed before the execution of the method. Moreover, due to the fact that it is a one-pass algorithm, it is extremely fast even if a large database of input–output examples is available.

B) Following the determination of the hidden node centres, the widths of the Gaussian activation function are calculated using the *P*-nearest neighbour heuristic (Leonard and Kramer, 1991):

$$\sigma_l = \left(\frac{l}{p} \sum_{i=1}^{p} ||\hat{\mathbf{x}}_l - \hat{\mathbf{x}}_i||^2 \right)^{1/2} \tag{6}$$

where $\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots, \hat{\mathbf{x}}_p$ are the p nearest node centres to the hidden node l. The parameter p is selected, so that many nodes are activated when an input vector is presented to the neural network model.

C) The connection weights are determined using linear regression between the hidden layer responses and the corresponding output training set, which is formulated as follows: For each input training vector, an output vector is created, where each component represents a different class. Therefore, the dimension of each output vector is equal to the number of different classes. The value of 1 is assigned to the position corresponding to the correct class, while the value of 0 is assigned to all other positions of the output vector.

2.3. Support vector machines (SVMs)

The original support vector machine (SVM) can be characterized as a supervised learning algorithm capable of solving linear and non-linear classification problems. The main building blocks of SVMs are structural risk minimization, originating from statistical learning theory, non-linear optimization and duality and kernel induced features spaces, underlining the technique with an exact mathematical framework. Meanwhile, several extensions to the basic SVM have been introduced, e.g. for multi-class classification as well as regression and clustering problems, making the technique broadly applicable in the data mining area.

The main idea of support vector classification is to separate examples with a decision surface and maximize the margin between the different classes. As complexity is considered directly during the learning stage, the risk of overfitting the training data is less severe for SVM. For constructing general non-linear decision functions, SVMs implement the idea to map the examples from the input space into a high-dimensional feature space function. The construction of a separating hyperplane in the features space leads to a non-linear decision surface in the original space. Expensive calculation of dot products in a high-dimensional space can be avoided by introducing a kernel function. This reduces numerical complexity significantly and allows efficient support vector learning for up to hundreds of thousands examples. Common kernels include polynomials and radial basis function classifiers with smoothing parameters. The design parameters that influence the speed of convergence and the quality of results are the choice of a kernel function and the corresponding kernel parameters.

In order to perform the comparison we used the radial basis function kernel function

$$K(x_i, x_j) = \exp(-\gamma||x_i - x_j||^2) \quad (7)$$

and optimized the results with respect to two parameters: C which is a constant cost parameter, controlling the number of support vectors and enabling the user to control the trade-off between learning error and model complexity, and γ which appears in the above definition of the kernel function.

3. Application area and available data

The application area is a wetland ecosystem of international importance (Ramsar Convention, NATURA 2000 proposed site, IBA etc.) located in Northern Greece (latitude: 42° 12′ N, longitude: 23° 09′ E). Lake Kerkini is a large artificial freshwater lake created on the site of a former natural swamp, after the construction of a dam across River Strymon in 1932, primarily for flood control. Subsequently, the irrigation value also developed. It occupies an alluvial plain between two mountain massives (Kerkini, Mavrovouni) and receives a large quantity of sediments from River Strymon. Following siltation by river sediments, which led to a loss of 61% of the Kerkini storage capacity, and an increase in the surface of irrigated land, it proved necessary to build a new, higher dam and a new dyke to the west (1982). The purposes of the damming were flood protection and the provision of water resources for irrigation. The maximum depth of the lake is 10 m with an annual fluctuation of water level between 4.5 and 5 m. The water level in the lake falls to a minimum each year between September and February and rises to a maximum level between early May and early June. One of the most important ecological features of the wetland is the *riparian forest* occupying the northern central part of the lake and now consists mainly of species of wild willow. It is the most important habitat of the wetland not only for birds but also for reptiles, amphibians and fishes. It is the nesting and feeding habitat for a lot of rare bird species during long periods and it is used by fishes of different species for spawning. Each year it is flooded by water resulting in its continuous shrinkage since 1983, as there is no natural regeneration to replace trees that die out due to protracted inundation. The lake and its riparian areas are important habitats for migratory bird species. In total the area hosts 102 breeding bird species, 171 migrating species and 171 wintering species. Fish fauna includes at least 30 species, there of 2 endemic sub-species.

The nomenclature used in the present study for wetland classes is EUNIS, which is a common European reference set of habitat units. Nomenclatures serve as classification keys that provide the link between ecologically defined habitats and the classes distinguishable on a satellite image (based on spectral values, texture and/or shape). The EUNIS Habitat classification has been developed to facilitate harmonised description and collection of data across Europe through the use of criteria for habitat identification. It is a comprehensive pan-European system, covering all types of habitats from natural to artificial, from terrestrial to freshwater and marine habitat types. In order to have a meaningful comparison, the two different classifiers were trained using the same set of ground truth data.

The satellite image used for the present application was acquired on 15 June 2000 by the IKONOS-2 satellite, which is on a sun-synchronous low earth orbit at a nominal altitude of 681 km and has a revisiting capability of 3 days. The IFOV is such that it collects images of the Earth with a very high spatial resolution of 1 m and 4 m in the panchromatic and multispectral modes, respectively. For the purposes of our study, a sub-scene of about 18 km^2 centred at the Delta of River Strymon is used. The four IKONOS-2 multispectral channels are tuned to detect radiation in the visible spectrum (0.45–0.53 μm centred in the Blue, 0.52–0.61 μm centred in the Green and 0.64–0.72 μm in the Red) as well as in the near infrared spectrum (NIR; 0.77–0.88 μm).

For the extraction of the training sets as well as the validation of classification results, a habitat classification map of the year 2000, conducted using black and white orthophotos (scale: 1:5000; date: 1997; source: Hellenic Army Geographic Service), was updated based on field work. The output habitat map accuracy was assessed with ground truth data and the user's accuracy was found to be 91%. It is worth noting that the ground truth data cover a wider area than the study area of the present work. The accuracy analysis of the classified images using the proposed methods was based on pixel-by-pixel comparison between the re-classified image and the ground truth map. The reference EUNIS habitat ground truth map, used for the extraction of the training sets and the assessment of the performance of the classifiers can be found in Fig. 2.

The samples extracted from the image are 843,290 in total. The selection of training samples for both classification schemes is carried out by extracting representative sample areas of same size (30×30 pixels) from each EUNIS class. The classes present in the study area together with the total number of samples, the number of training clusters and the number of training samples for

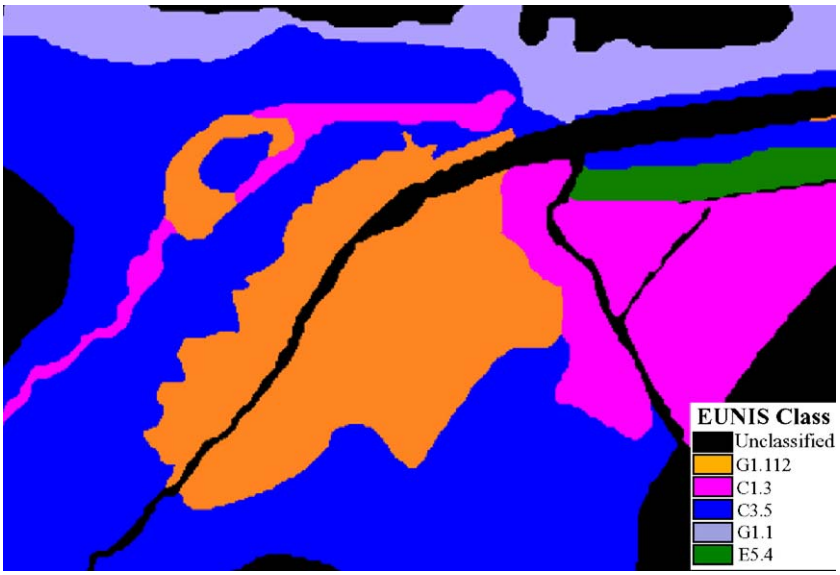


Fig. 2. Ground truth map.

each class are shown in Table 1. It is important to note that the percentage of the training samples is less than 2% of the total number of samples. As it is not possible to distinguish free from rooted floating vegetation of eutrophic waterbodies on a satellite image (classes C1.32 and C1.34), these are considered as one class (C1.3). Another point of interest is the exclusion of the river (EUNIS code C2.3) due to the shape of the class, which does not allow the extraction of 30×30 training clusters. It has to be noted that the current version of KRC does not allow variable training cluster size. Thus, the shape and size of the samples are imposed by the requirements of kernel re-classification algorithm. In total, five different classes will be considered for the re-classification.

Table 1
EUNIS classes encountered in Lake Kerkini test site

Class code	Class definition	No. of training clusters	Total samples	Training samples
G1.112	Mediterranean tall galleries (riparian forest)	4	157,975	3600
C1.32	Free floating vegetation of eutrophic waterbodies	3	97,613	2700
C1.34	Rooted floating vegetation of eutrophic waterbodies			
C3.5	Pioneer and ephemeral vegetation of periodically inundated shores	4	357,344	3600
G1.1	Riparian woodland	1	48,768	900
E5.4	Moist or wet tall herb and fern fringes and meadows	1	8635	900

4. Results

4.1. Kernel re-classification algorithm

Following the selection of the appropriate training samples, the KRC software is executed. The algorithm is tested employing several kernel sizes varying between 3×3, 5×5, 7×7, 9×9, 11×11 pixels. The results are compared in Table 2, where, for each kernel size, the overall accuracy is presented, i.e. the total number of pixels correctly re-classified to the total number of image pixels (including the training samples). We observe that the kernel of 9×9 and 11×11 pixels return the best class discrimination. The kernel of 9×9 pixels is chosen in order not to deteriorate the classified image by eliminating objects of smaller size. The re-classified image, which is the final product of KRC is shown in Fig. 3.

In order to investigate the performance of the kernel classifier in terms of better class discrimination and overall accuracy, the confusion matrix was computed. This was based on pixel-by-pixel comparison between the re-classified image and the ground truth map. The

Table 2
Description of KRC classification experiments

Case	Kernel size	Overall performance (%)
1	3×3	51.02
2	5×5	54.60
3	7×7	56.02
4	9×9	56.71
5	11×11	56.58

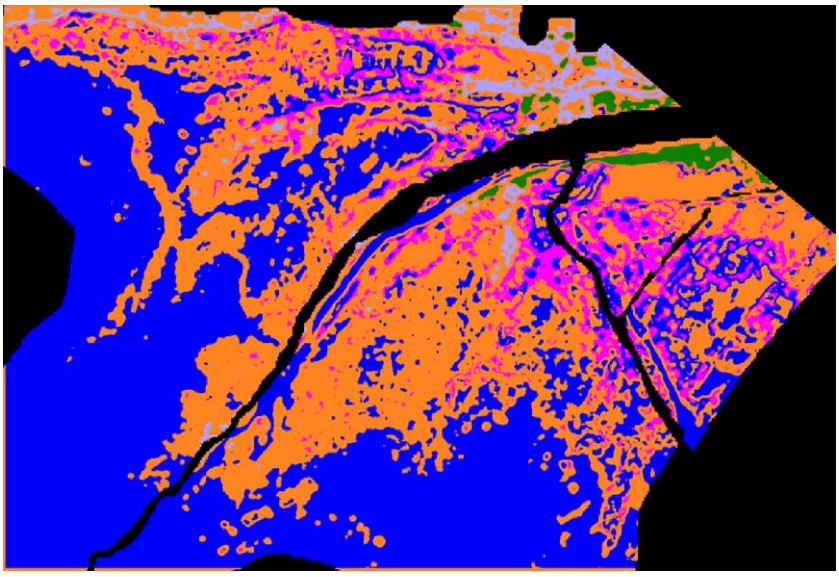


Fig. 3. KRC classification output.

corresponding confusion matrix is presented in Table 3. Each element of the Table represents a percentage, such that each row sums to 100% (user's accuracy representing errors of commission). The overall accuracy, i.e., the total number of pixels correctly re-classified to the total number of image pixels, was 56.71%. In shaded cells along the diagonal, the individual class performances are shown for all five classes, ranging from 3.99% for E5.4 to 71.12% for C3.5.

4.2. RBF-NN classifier

The same samples as in the kernel re-classifier were extracted from the original 4-band image for training the RBF-NN classifiers. This resulted in 11,700 training samples (pixels). For each training pixel, a set of spectral parameters were calculated as follows:

a. original digital number of the pixel (DN value),
b. mean value of the digital numbers within a specified window, and
c. standard deviation of the digital numbers within a specified window.

Textural parameters were not included in the analysis since it was shown (Keramitsoglou et al., 2003) that its incorporation does not affect the performance of the RBF-NN classifiers substantially.

In order to apply the algorithm presented in Section 2.2 to the specific case under study, we prepared the output training set as follows: For each input training vector which consists of the spectral information, an output vector of dimension 5 was assigned, where each element corresponds to a class. Knowing the correct class for the training example we assigned the value of 1 to the respective position in the output vector, while the value of 0 was assigned to all other positions.

When a validation input vector is presented to a RBF network, the classifier should respond with a value close to 1 in the correct position of the output vector and values close to 0 in all other positions. In fact, the way we used the RBF classifiers to assign classes to all the different pixels of the map was the following: For each pixel, the appropriate input vector was first formulated by placing the input parameters in the correct order that was used during the training phase. Then all the output responses were computed and the pixel was assigned to the class corresponding to the element of the output vector with the highest value. This value also gives the confidence level in accepting the classification produced

Table 3
Confusion matrix illustrating the results of the KRC classification

	Habitat map classification					
	EUNIS class	G1.112	C1.3	C3.5	G1.1	E5.4
Image Classification	G1.112	66.26	6.31	26.07	1.18	0.18
	C1.3	46.16	21.53	31.14	0.74	0.42
	C3.5	24.11	2.99	71.12	0.57	1.21
	G1.1	58.77	10.39	4.79	22.44	3.62
	E5.4	78.77	6.51	10.67	0.05	3.99

Numbers denote percentage values.

by the neural network model for the particular pixel. Obviously the confidence level increases as the highest output value is getting closer to 1.

For the purposes of this study, a number of experiments were designed. The experiments did not differ in the physical parameters used for training the neural networks. Only the window size in which statistical and spectral parameters are calculated and the number of fuzzy sets in the initial fuzzy partition of the input space varied. More specifically the experiments were carried out using 3×3, 5×5 and 7×7 windows for the calculation of the spectral features, and 12, 15, 20 and 23 fuzzy sets in each input dimension in the fuzzy means algorithm. In each experiment, the size of the network was determined by the fuzzy-means algorithm based on the initial fuzzy partition of the input space. Table 4 gives all the attributes used for the different experiments and the resulting overall accuracies. The training times were low as expected, since the fuzzy means algorithm needs only one pass of the training data. It must be noted that even in case 4, where the complexity of the network architecture is the highest, the total training time did not exceed 2 min on a Pentium IV 1400 MHz machine. The most accurate neural network model is the one produced by a 7×7 window and 20 fuzzy sets in each input dimension in the fuzzy means algorithm, corresponding to an overall performance of 68.81%. The corresponding confusion matrix and visual interpretation of the classified images are presented in Table 5 and Fig. 4, respectively. The individual class performance ranges from 39.65% for E5.4 to 74.41% for C3.5.

4.3. SVM classifier

The SVM method was tested on exactly the same data that was used for the two other classification schemes. The training set consisted of the information that was used to train the RBF-NN classifiers, i.e. 11,700 training samples extracted from the original 4-band image, containing for each band the digital number of the pixel and mean values and standard deviations within a specified window. The method was tested on 3×3, 5×5 and 7×7 window sizes. To perform the classifications using the SVM method we utilized the LIBSVM software (Chang and Lin, 2001), which provides not only class labels, but also probabilities estimates. The computational effort required by the SVM method is very low, since training time never exceeded 15 s in a Pentium IV 1400 MHz machine. A sensitivity analysis on the values of the two design parameters C and γ was performed. Table 6 gives all the attributes used in various representative experiments and the resulting overall accuracies. The most accurate SVM model was the one produced by experiment 23 (7×7 window size, $C=1$, $\gamma=1$), for which the overall accuracy was 71.68%. The corresponding confusion matrix and visual interpretation of the classified images are presented in Table 7 and Fig. 5, respectively. The individual class performance ranges from 40.54% for G1.1 to 79.65% for C3.5.

Table 4
Description of neural network classification experiments

Case	Window size	Number of fuzzy sets	Number of hidden nodes	Overall performance (%)
1	3×3	12	78	67.01
2		15	164	67.60
3		20	383	62.34
4		23	580	59.09
5	5×5	12	78	67.12
6		15	150	67.78
7		20	327	66.71
8		23	456	62.73
9	7×7	12	81	66.34
10		15	134	68.54
11		20	299	68.81
12		23	422	64.86

Table 5
Confusion matrix illustrating the results of the RBF-NN classifier

	Habitat map classification					
	EUNIS class	G1.112	C1.3	C3.5	G1.1	E5.4
Image Classification	G1.112	72.73	7.88	17.30	0.91	1.18
	C1.3	13.72	62.57	17.33	1.59	4.78
	C3.5	14.46	8.21	74.41	1.32	1.60
	G1.1	40.22	1.44	11.96	40.21	6.18
	E5.4	4.40	48.59	3.65	3.72	39.65

Numbers denote percentage values.

4.4. Comparison between the three classification methods

4.4.1. Overall and individual class performance

As mentioned before, the overall performance of the three classifiers was carried out in an identical manner, i.e., by pixel-to-pixel comparison of the classified images. Experimentation with the design parameters (Tables 2, 4, 6) showed that in general the RBF-NN and SVM methodologies outperform the KRC classifiers. More precisely, the performance of the different RBF-NN classifiers was superior to the best KRC classifier by 2% to 12%, depending on the window and network size. It is interesting to note that the performance of the RBF-NN classifier is deteriorated when a large number of

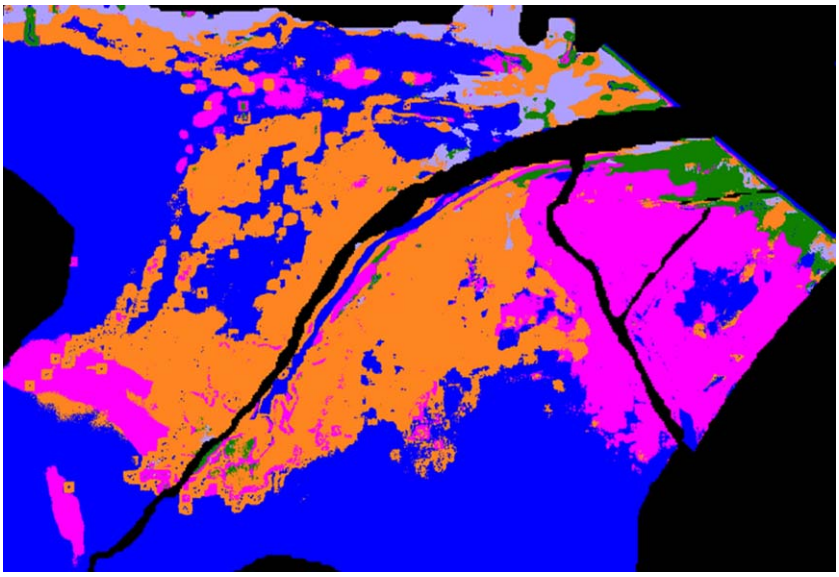


Fig. 4. RBF-NN classification output.

fuzzy sets is utilized in the initial fuzzy partition of the input space. This can easily be explained, by noticing that dense fuzzy partitions result in large network sizes, which suffer from the well-known overfitting phenomenon. The SVM classification method seems less robust compared to the RBF-NN methodology with respect to the design parameters. Although the SVM method produced the best classifier in terms of the overall performance, there are cases (experiments 16, 25) where the resulting SVM classifier has a lower overall accuracy compared to the best KRC model. Robustness in the SVM method is clearly lost when we increase the window size. Window size does not affect considerably the performance of the RBF-NN methodology.

Table 6
Description of SVM classification experiments

Case	Window size	C	Γ	Overall performance (%)
1	3×3	0.1	0.1	68.28
2		1	0.1	67.48
3		1000	0.1	65.97
4		0.1	1	69.00
5		1	1	69.61
6		1000	1	70.12
7		0.1	10	66.03
8		1	10	68.51
9		1000	10	68.43
10	5×5	0.1	0.1	69.06
11		1	0.1	68.22
12		1000	0.1	70.45
13		0.1	1	69.81
14		1	1	70.20
15		1000	1	71.30
16		0.1	10	56.32
17		1	10	60.83
18		1000	10	61.80
19	7×7	0.1	0.1	69.07
20		1	0.1	69.73
21		1000	0.1	66.03
22		0.1	1	70.21
23		1	1	71.68
24		1000	1	71.28
25		0.1	10	54.37
26		1	10	56.90
27		1000	10	57.18

The best classifiers produced by the three methodologies were further compared. These are the classifiers provided by Experiment No. 4 in the KRC methodology (9×9 kernel size, 56.71% performance), Experiment No. 11 in the RBF-NN methodology (7×7 window size, 20 fuzzy sets, 68.81% performance) and Experiment No. 23 in the SVM methodology (7×7 window size, $C=1$, $\gamma=1$, 71.68% performance). In terms of overall accuracy, the best SVM model outperforms the most

Table 7
Confusion matrix illustrating the results of the SVM classifier

	Habitat map classification					
	EUNIS class	G1.112	C1.3	C3.5	G1.1	E5.4
Image classification	G1.112	69.20	3.94	24.22	1.14	1.49
	C1.3	11.48	64.39	19.06	1.52	3.55
	C3.5	11.22	3.92	79.65	1.97	3.25
	G1.1	29.97	1.40	24.48	40.54	3.61
	E5.4	3.08	35.47	5.60	0.00	55.84

Numbers denote percentage values.

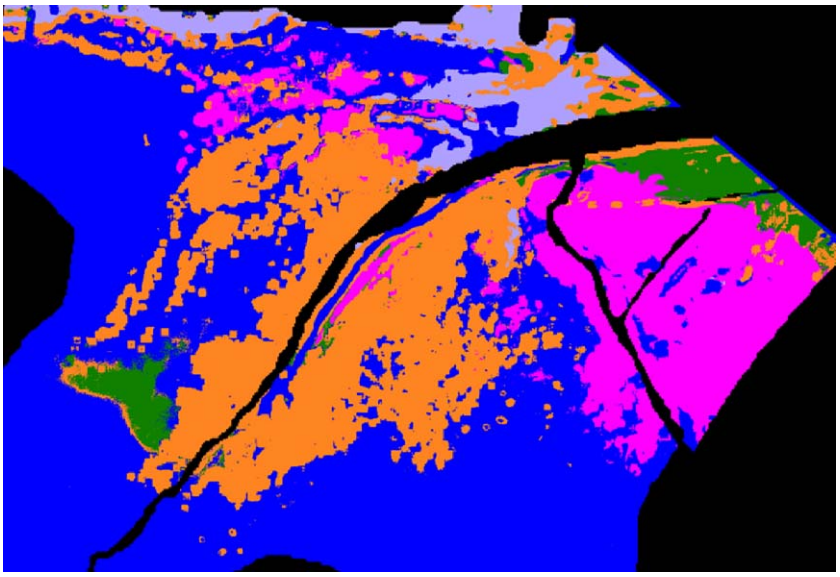


Fig. 5. SVM classification output.

accurate classifiers produced by the rest of the methods. Among the three, the KRC classifier is clearly the less accurate. In order to examine if a significant difference exist in the performances of the two best methods, we applied the McNemar test (Dietterich, 1998). We recorded for each test example how it was classified by the two algorithms and we measured the number n_{10} of examples misclassified by the best SVM classifier but not by the best RBF-NN classifier and the number n_{01} of examples misclassified by the best RBF-NN classifiers but not by the best SVM classifier. We found that n_{10} and n_{01} were equal to 51,719 and 72,129, respectively. Then we calculated the following quantity:

$$\frac{(|n_{01}-n_{10}|-1)^2}{n_{01}+n_{10}} = 3363.2$$

According to the McNemar test, this quantity is large enough to favor the hypothesis that the performance of the best SVM classifier is superior to the best RBF-NN for this particular training set.

Individual class performances are assessed by means of confusion matrices (Tables 3, 5 and 7). The riparian forest, which is a key habitat of the wetland, was classified satisfactorily by all three classifiers. This was the only class where RBF-NN outperformed SVM reaching an individual class performance of 73%. Free and rooted floating vegetation of eutrophic water bodies (C1.3) was poorly classified by KRC as it was confused with G1.112; on the other hand, RBF-NN and SVM reached an accuracy of 62% and 64%, respectively. Pioneer and ephemeral vegetation of periodically inundated shores (C3.5) was the wetland class best classified by all algorithms. More specifically, accuracies range from 71% (KRC) to 80% (SVM). The riparian woodland (G1.1) was poorly classified by KRC and better (but still not satisfactorily) by the other two classifiers. SVM exhibited the highest performance compared to the other two by reaching approximately 40%. Finally, class E5.4 (moist or wet tall herb and fern fringes and meadows) completely failed to be classified by KRC, whereas was classified with an accuracy of 56% by SVM. Overall, SVM had the highest individual class performance, with only one exception (G1.112 where RBF-NN performed better) and KRC the poorest; it has to be noted that RBF-NN was closely behind SVM.

4.4.2. Sensitivity to confidence level

For all three classifiers presented in this work there is an index, which specifies the degree of confidence in the classification result of each pixel. For the kernel based re-classifier (KRC) this is expressed by the similarity index (Eq. (2)), whereas for RBF-NN by the confidence level and for the SVM method by the probability estimate. It is worth examining the sensitivity of the output habitat maps in terms of the threshold value of these indices. In order to get meaningful results, one has to analyse not only the overall performance of the classifier but the percentage of the pixels classified with confidence higher than the threshold value as well. As the threshold value increases, fewer pixels are included in

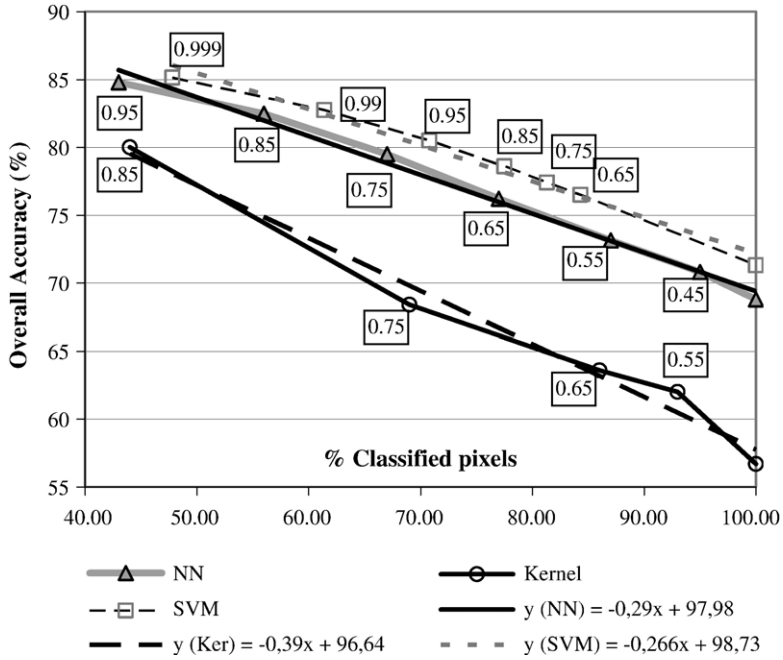


Fig. 6. Sensitivity of classification product to confidence level. Number in box indicates the corresponding confidence level.

the final product, however, the pixels which are included in the map are the ones classified with higher confidence and therefore, the overall accuracy is increased.

To investigate the behaviour of the three classifiers, different confidence threshold values were applied and the pixels classified with confidence below that level were masked out. The results are illustrated in Fig. 6, where the *x*-axis represents the percentage of pixels classified and the *y*-axis the overall accuracy of the classifier. The shape of all three curves is similar. For rather high values of confidence, only a small portion of classified pixels is observed with a high overall accuracy. For the core body of classification a linear decay of confidence is noticed as the population of pixels accepted for classification increases. The results favour the SVM classifier since for the entire *x*-axis (any percentage of classified images), we observe differences ranging from 2% to 3% from the RBF-NN classifier and from 7% to 15% from the KRC classifier.

5. Summary and conclusions

It is a well-known problem in the remote sensing community that, as the latest very high spatial resolution (<10 m) satellite data become available, there is a lack of image processing tools that can compensate for the inevitable problem present in land cover/use classes including heterogeneous spectral classes. This is very often the case with habitat mapping, where small patch sizes and high land cover heterogeneity further deteriorate the classification accuracies (Smith et al., 2002). This study demonstrates the use of advanced image processing methods for classifying multispectral very high spatial resolution data acquired by the IKONOS-2 satellite.

The three automatic methods presented here are the kernel re-classification algorithm (KRC), RBF-NNs, and SVMs. The KRC algorithm has overall performed fairly. Rather weak results were obtained for class boundaries from homogeneous to spectrally variable areas (sometimes referred to as 'the boundary effect'; Gong, 1994) as well as along the river. The latter is due to the absence of training clusters because of the object size and shape. The corresponding average similarity index (see Eq. (2)) is 80%, which can be considered satisfactory for such a complex scene. For training the RBF-NN and SVM classifiers, the same samples were extracted from the original 4-band image. The best classifiers produced by the two methods provided accuracies of 68.8 and 71.7%, respectively, and clearly outperformed the best KRC classifier. Visual interpretation of the output maps also shows the clear advantage of the RBF-NN and SVM classifiers in terms of noise within the classes. It is also worth noting the improved thematic representation of the SVM and RBF-NN maps compared to KRC map, since the classified objects have similar shapes to the ground

truth map, including fewer unclassified and misclassified areas within the objects. As far as robustness is concerned regarding the design parameters, the RBF-NN exhibited accuracies ranging from 59.09% to 68.81% depending mainly on the NN structure, and to a lesser extend on the window size. The SVM method was found less robust especially for large window sizes and produced classifiers with accuracies ranging from 54.37% to 71.68%.

In addition, it should be emphasized that there is a clear difference in the computational requirements of the three methodologies (seconds for SVM vs. minutes for RBF-NN vs. hours for KRC) in a typical Pentium IV PC workstation. As a result, the SVM classification is strongly recommended for tackling similar problems involving very high spatial resolution satellite images.

As a general conclusion one can derive that the EUNIS nomenclature scheme is very complex to be automatically classified, even by the three advanced classifiers tested, given the fine structure of the classes (in terms of high land cover mixture within the classes) and the spatial resolution of the satellite data used. Although the overall accuracy was rather high especially for the SVM classifier, there are specific classes, for which the accuracy was low (i.e. G1.1 was classified with an accuracy of 40.54%). We strongly believe that a classification aiming at such a detailed level of the EUNIS scheme (level 5 in the above example) requires higher spatial resolution and multi-date imagery. This would result in higher inter-class spectral variability, which can be better tackled by pixel window classifiers. In addition, some elements of the classes considered in the present case study (Table 1) were not resolved by the sensor acquiring 4-m spatial resolution imagery and therefore could not be identified by the classification methods used.

The limitations described above have an inevitable effect on any subsequent calculation of landscape metrics. Therefore, although in principle automated habitat mapping from satellite images can overcome the problem of photo-interpreters' objectivity, there is still a requirement to standardise the methodology used, in order to have meaningful results when studying the temporal changes of a protected area.

Acknowledgements

Application of KRC to very high spatial resolution satellite images was funded by the SPIN (Spatial Indicators for European Nature Conservation) research project supported by the European Commission under the Fifth Framework Programme. Contract no.: EVG 1-CT-2000-019 (project web site: www.spin-project.org). The authors would also like to express their thanks to the anonymous reviewers for their helpful comments and suggestions that substantially improved the content of this article.

References

Abuelgasim, A.A., Gopal, S., Irons, J.R., Strahler, A.H., 1996. Classification of ASAS multiangle and multispectral measurements using artificial neural networks. Remote Sensing of Environment 57 (2), 79–87.

Barnsely, M.J., Barr, S.L., 1996. Inferring urban land use from satellite sensor images using kernel-based spatial re-classification. Photogrammetric Engineering and Remote Sensing 62 (8), 949–958.

Bock, M., 2003. Remote sensing and GIS-based techniques for the classification and monitoring of biotopes. Case examples for a wet grass- and moor land area in Northern Germany. Journal of Nature Conservation 11 (2), 145–155.

Camps-Valls, G., Bruzzone, L., 2005. Kernel-based methods for hyperspectral image classification. IEEE Transactions on Geoscience and Remote Sensing 43 (6), 1351–1362.

Chang, C.C., Lin, C.J., 2001. LIBSVM: a Library for Support Vector Machines. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm (accessed December 20, 2005).

Chettri, S.R., Cromp, R.F., Birmingham, M., 1992. Design of neural networks for classification of remotely sensed data. NASA 1992 Goddard Conference on Space Applications of Artificial Intelligence, Greenbelt, Maryland, pp. 137–150.

Darken, C., Moody, J., 1990. Fast adaptive *k*-means clustering: some empirical results. IEEE INNS International Joint Conference on Neural Networks; Proceedings, vol. 2, pp. 233–238.

Decatur, S.E., 1989. Application of neural networks to terrain classification. International Joint Conference on Neural Networks, pp. 283–288.

Dietterich, T.G., 1998. Approximate statistical test for comparing supervised classification learning algorithms. Neural Computation 10 (7), 1895–1923.

Dunn, J.C., 1973. A fuzzy relative to the ISODATA process and its use in detecting compact, well-separated clusters. Journal of Cybernetics 3 (1), 32–57.

EEA, European Environmental Agency, 2002. EUNIS Habitat Classification Web Application. Electronic Source: http://mrw.wallonie.be/dgrne/sibw/EUNIS/home.html (accessed December 20, 2005).

Foody, G.M., Mathur, A., 2004a. Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification. Remote Sensing of Environment 93 (1–2), 107–117.

Foody, G.M., Mathur, A., 2004b. A relative evaluation of multiclass image classification by support vector machines. IEEE Transactions on Geoscience and Remote Sensing 42 (6), 1335–1343.

Gong, P., 1994. Reducing boundary effects in a kernel-based classifier. International Journal of Remote Sensing 15 (5), 1131–1139.

Gong, P., Howarth, P.J., 1992. Land-use classification of SPOT HRV data using a cover-frequency method. International Journal of Remote Sensing 13 (8), 1459–1471.

Hepner, G.F., Logan, T., Richter, N., Bryant, N., 1990. ANN classification using a minimal training set: comparison to conventional

supervised classification. Photogrammetric Engineering and Remote Sensing 56 (4), 469–473.

Kanellopoulos, I., Varfis, A., Wilkinson, G.G., Megier, J., 1990. Land-cover discrimination in SPOT-HRV imagery using artificial neural network a 20-class experiment. International Journal of Remote Sensing 13 (5), 917–924.

Keramitsoglou, I., Kontoes, H., Elias, P., Sifakis, N., Fitoka, E., Weiers, S., 2003. Kernel-based reclassification algorithm applied on very high spatial resolution satellite imagery of complex ecosystems. Proceedings of the 10th International Symposium of Remote Sensing, Remote Sensing for Agriculture. Ecosystems and Hydrology V, 8–10 September 2003, Barcelona, Spain, vol. 5232, pp. 276–283.

Kerr, J.T., Ostrovsky, M., 2003. From space to species: ecological applications for remote sensing. Trends in Ecology and Evolution 18 (6), 299–305.

Kim, K.I., Kim, J.H., Jung, K., 2002. Face recognition using support vector machines with local correlation kernels. International Journal of Pattern Recognition and Artificial Intelligence 16 (1), 97–111.

Kontoes, C.C., Raptis, V., Lautner, M., Oberstadler, R., 2000. The potential of kernel classification techniques for land use mapping in urban areas using 5 m-spatial resolution IRS-1C imagery. International Journal of Remote Sensing 21 (16), 3145–3151.

Leonard, J.A., Kramer, M.A., 1991. Radial basis function networks for classifying process faults. IEEE Control Systems Magazine 11 (3), 31–38.

Moody, J., Darken, C., 1989. Fast learning in networks of locally-tuned processing units. Neural Computation 1 (2), 281–294.

Sarimveis, H., Alexandridis, A., Tsekouras, G., Bafas, G., 2002. A fast and efficient algorithm for training radial basis function neural networks based on a fuzzy partition of the input space. Industrial and Engineering Chemistry Research 41 (4), 751–759.

Schölkopf, B., Smola, A.J., 2002. Learning with Kernels—Support Vector Machines, Regularization, Optimization and Beyond. MIT Press, Cambridge.

Smith, J.H., Wickham, J.D., Stehman, S.V., Yang, L., 2002. Impacts of patch size and land-cover heterogeneity on thematic image classification accuracy. Photogrammetric Engineering and Remote Sensing 68 (1), 65–70.

Tou, J.T., Gonzalez, R.C., 1974. Pattern Recognition Principles. Addison-Wesley Publishing Company, Reading, Massachusetts.

Vapnik, V.N., 1995. The Nature of Statistical Learning Theory. Springer Verlag, New York.

Vapnik, V.N., 1998. Statistical Learning Theory. Springer Verlag, New York.

Zhang, Q., Wang, J., Gong, P., Shi, P., 2003. Study of urban spatial patterns from SPOT panchromatic imagery using textural analysis. International Journal of Remote Sensing 24 (21), 4137–4160.