ANALYSIS OF THE TRANSFERABILITY OF SUPPORT VECTOR MACHINES FOR VEGETATION CLASSIFICATION

Henny Mills

Newcastle University, School of Civil Engineering and Geosciences, Cassie Building, Newcastle University, Newcastle upon Tyne, United Kingdom NE1 7RU

Commission, WGVII/4

KEY WORDS: Remote Sensing, Mapping, Classification, Accuracy, High resolution, Ikonos

ABSTRACT:

This paper reports on research analysing the potential of Support Vector Machines (SVMs) for mapping vegetation from high spatial resolution Ikonos imagery. The work investigated the utility of SVMs for mapping regional scale upland vegetation using limited ground data. Additionally, it analysed the ability of SVMs to be transferred as a classifier to pixels from remote geographical locations, which were not included in the training process. The classification and transferability of SVMs was investigated when varying their design and training. Overall, the classification and transferability results of SVMs showed very promising results, highlighting their capability and suitability for use in remote sensing classification.

1. INTRODUCTION

Support Vector Machines (SVMs) have gained a growing importance and recognition in remote sensing as a classification methodology. They have often outperformed other remote sensing classification methodologies, such as Artificial Neural Networks (ANN), and thereby underlined their potential for mapping from imagery. This paper reports on research which analysed the potential of SVMs for mapping vegetation from high spatial resolution Ikonos imagery at regional scale. It presents an investigation of the utility of SVMs for mapping regional scale upland vegetation from Ikonos imagery using limited ground data, transferring the methods applied to remote geographical locations. The classification and transferability potential was investigated and maximised when varying design and training, showing that pixels of remotely located areas can be classified with this advanced classification method. The study investigated the option of the optimal hyperplane of a SVM classification in relationship to generalisation, analysing if and how the generalisation of the hyperplane can be maximised to achieve high transferability classification with SVMs.

The application of remote sensing over large geographical areas is still limited, in either accuracy (spatial resolution) or algorithm performance. Transferability of a mapping approach is required to consistently map areas of national and regional scale and to overcome the lack of large amounts of ground data for most studies. Transferability would also enable an application to be more time-efficient and enable the mapping of a large area regularly (Woodcock et al., 2001). The transferability of classifiers over large geographical areas has, however, been found to be limited so far and is dependent upon the ability of the classifier to generalise (Benediktsson et al., 1990). The lack of transferable algorithms using remote sensing has limited its potential contribution for environmental studies and is still a disadvantage of many remote sensing algorithms (Foody et al., 2003).

This study, therefore, aimed to analyse the transferability potential of SVMs by applying them to map upland vegetation (Fukuda and Hirosawa, 2001; Gualtieri et al., 1999). SVMs require small training data sets, e.g. in comparison to ANN, as only the support vectors are used to locate the hyperplane between classes, resulting in a high generalisation ability (Huang et al., 2002). It highlights the potential of SVMs for remote sensing classification as sufficient training data are often expensive and difficult to obtain. The ability of SVMs to generalise using limited training data set offers an opportunity to achieve high classification accuracies. It was expected that SVMs would also be able to transfer knowledge gained during the training on one geographical area to classify pixels from unseen input data samples. The paper aims to highlight the advantages and potential of SVMs for remote sensing applications to the remote sensing community.

2. SUPPORT VECTOR MACHINES

The theory behind SVMs is only briefly described in this paper, as they have been explained in detail in (Foody and Mathur, 2004). SVMs aim to reduce the learning error between target and output data to zero by locating the optimal boundaries between classes, thereby finding the optimal hyperplane which minimises the probability of classification error and separates the data points of two classes. The optimal hyperplane maximises the margins from the closest data points to the hyperplane, the most difficult ones to classify, and is hence the one offering the highest generalisation in comparison to other separating hyperplanes (
Figure 1). The optimal hyperplane is defined as the decision surface to maximise the separation to be fulfilled by weight w_0, being normal to the hyperplane and bias b_0 (1).

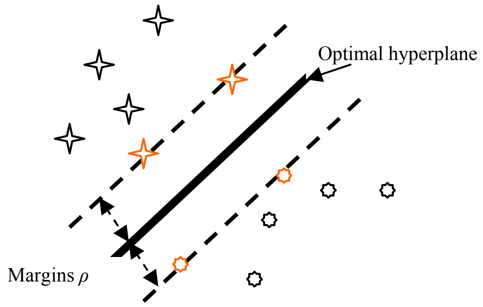

Figure 1: Hyperplane for linear separable patterns, showing the support vectors for the two classes on the hyperplane boundaries (dotted line) (Foody and Mathur, 2004).

Support vectors are part of the training data and are chosen by the algorithm to guarantee a stable representation of the data. As a result, the performance of SVMs would only be influenced if training samples removed are those used as support vectors.

$$\begin{array}{lll} w_0^T x_i + b_0 \geq 1 & for & t_i = 1 \\ w_0^T x_i + b_o \leq 1 & for & t_i = -1 \end{array} \quad (1)$$

The optimal hyperplane maximises the margins from the closest data points to the hyperplane and is, therefore, the hyperplane of the lowest capacity, resulting in best generalisation in comparison to other separating hyperplanes.

2.1 Support Vector Machines for Image Classification

For a remote sensing classification, data sets consist of nonseparable pattern, indicating that the classification error does in many cases not reach zero. The optimal hyperplane has therefore to be designed to minimise the probability of the classification error (Figure 2).

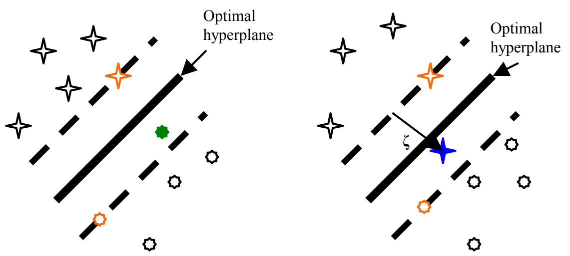

Figure 2: Classification of data points of inseparable data sets, with the filled round data point being within the region of separation but on the correct side (left) and the filled star data point being located within the wrong side of the hyperplane (Foody and Mathur, 2004).

The classification is supported by the introduction of a slack variable ζ. The slack variable is applied as upper bound on the number of training errors. The best separating hyperplane can be found with (2). The regularisation parameter C, based on the slack variable, controls the tradeoff between complexity and the number of nonseparable points and can be defined experimentally or analytically (Huang et al., 2002). A low C value might conclude in a high number of support vectors, whereas a large value for C can cause overfitting to the training data and thereby limit the SVM generalisation.

$$\Phi(w,\xi) = \frac{1}{2} w^T w + C \sum_{i=1}^{N} \xi_i \quad (2)$$

The application of SVMs for pattern classification consists of two stages: the nonlinear mapping into a high dimensional feature space Φ and the construction of the optimal hyperplane in that feature space as linear separation. An inner kernel function K is used for the transformation into the feature space. It enables the training and application of the SVM classifier in the high dimensional feature space without the explicit knowledge of the feature space mapping function (3) (Huang et al., 2002).

$$\sum_{i=1}^{N} \alpha_i d_i K(x, x_i) = 0 \quad (3)$$

Different learning machines can be applied: the two most common ones being the polynomial learning machine, with the power p defined by the user, and the Radial Basis Function (RBF). SVMs guarantee to find the global minimum of the error surface between target and output during the training process.

2.2 SVMs for multi-class classification

SVMs were developed as binary classification algorithms. To allow the application of SVMs for multiclass remote sensing classification, a combination of binary SVMs has to be used. A Multiclass SVMs can be carried out in two different ways: classification of training data for one class against the training data from all other classes (one vs. rest) or classification of all possible pairs of binary classification between two classes (one vs. one) (Gualtieri et al., 1999). The latter approach results in a higher number of classifications (4) and the pixels are given the class label of the class, which has gained most votes during the one vs. one classifications (Gualtieri et al., 1999).

$$\frac{n(n-1)}{2} \quad (4)$$

3. APPLICATIONS OF SVMS IN REMOTE SENSING

SVMs have been applied to pattern recognition and classification problems for a limited set of applications (Huang et al., 2002). In comparison to other classification approaches, such as ANNs, SVMs have proved to be superior in accuracy and stability for remote sensing problems (Fukuda and Hirosawa, 2001; Huang et al., 2002). Overall, the achieved classification accuracies of SVMs in comparison to ANNs showed an improvement of 1-4% and have highlighted the advantages and potential of SVMs for remote sensing classification. The SVM applications ranged from binary classification (Perkins et al., 2001) to multiple-class classification, of e.g. 13 classes (Fukuda and Hirosawa, 2001). For a SAR classification of 13 classes the one versus rest multiple classification approach was applied and showed a visually closer match to the reference map than the map of the ML classification (Fukuda and Hirosawa, 2001). The same strategy was applied for a three class (forest, non-forest and water) problem using simulated MODIS imagery (Huang et al., 2002). SVMs consisting of a polynomial kernel of a higher order presented the best performances. The study also confirmed that a large amount of training data does not necessarily improve the classification, because only the support vectors are used in the separation process between classes. A further SVM classification concluded in an improved accuracy of 1-2 % against ANNs (Huang et al., 2002). The ease of the SVM design was highlighted in a study using SAR and ATM imagery to map a site in Southern England into five classes. For different SVM designs accuracy between 89.4% and 91.6% were achieved, which were higher than the ones achieved with an ANN classification (Roli and Fumera, 2000). Similar classification accuracies, ranging between 87% and 96%, were attained for hyperspectral AVIRIS imagery, depending on the number of classes: six or 16 classes respectively (Gualtieri et al., 1999). Research showed that SVMs can be trained to high classification accuracies, even if only a small number of training samples is used, as long as those are defined as support vectors for the separation between two classes (Foody and Mathur, 2004). Further research into the optimal design and suitability of SVMs for different land cover applications has still to be carried out.

The acknowledgment of the suitability of SVMs for remote sensing applications could raise the recognition of the potential of remote sensing for many scientific areas and real world applications. SVMs offer a high generalisation ability which supported the decision to test their transferability and classification of multi-site upland vegetation in this study.

4. SVMS FOR MAPPING UPLAND VEGETATION

4.1 Study site and data

Upland vegetation is protected under British and European laws, which require information on the extent, condition and changes of upland vegetation species for managing and monitoring purposes. Previous research showed that traditional classification approaches resulted in varying accuracies of below 80% which were seen as in appropriate for public funded monitoring schemes. Given the published performance of SVMs for remote sensing classifications, SVMs were expected to offer a potentially improved classification tool for upland vegetation. A multispectral Ikonos Carterra™ Geo imagery was obtained, as recorded on 2 September 2002, covering an area of 156 km^2, in Northumberland National Park (UK). This is probably the optimum time for the acquisition of remote sensing imagery for upland areas, as most upland vegetation types are fully developed. The imagery was radiometrically and geometrically corrected using various Ground Control Points (GCP's), which were established using high precision geodetic Global Positioning System (GPS) receivers. Each point was corrected for relief displacement caused by the altitude variation of the area in the Ikonos image. The georectification was completed to an 'artificial' reference plane of 300 m (the average height above sea level of this area), resulting in an final horizontal accuracy of 2.5 m using 18 ground control points. Additionally, the Normalised Difference Vegetation Index was calculated to enhance the spectral separability between all

4.2 Design of SVMs

The design and training of SVMs was required to be carried in such a way that it would enable the SVMs to result in a high generalisation, even if this may decrease the classification accuracy of the training parameters. For the kernel function, two design parameters have to be chosen: the regularisation parameter C as the penalty parameter, and γ describing the characteristic of the kernel functions (Foody and Mathur, 2004). No rules yet exist for the optimal choice of these two parameters. Large values for both parameters might lead to overfitting of the SVMs to the training data and therefore to a limited generalisation ability (Cortez et al., 1997). Values of both parameters were varied for each kernel function, ranging between 0.01 and 600.

The multi-class SVM classifications were performed as one versus one classification, training the classifier on all possible pairs of binary classification between two classes. It resulted in 21 classifications (with seven classes). Despite the fact that this multi-class methodology consists of more binary classifications than the one versus rest classification approach, it terminates faster because of smaller data sets in each binary SVM and as a result requires less training data (Huang et al., 2002). The one versus one multi-class approach has also previously resulted in higher classification accuracy than the one versus rest SVM classification.

The additive Matlab toolbox LIBSVM was chosen for this study for the execution of SVMs because it enabled the application of multi-class classification.

4.3 Input and training data for SVMs

Firstly, SVMs were applied as an ordinary classifier, being trained on limited pixels of the study area and then applied to classify the remaining (majority) pixels of this area. For this part of the study, the SVMs (SVM_original) were trained using data from both test sites - the site A and site B and then applied to classify validation pixels from those sites. In the second part of this paper, the transferability of SVMs is evaluated when being trained on pixels of one area (site A) and then applied to second area (site B) of which no pixels were included in the training process (referred to as SVM_transfer). A third set of SVMs was developed which consisted additionally of a geographical label as input data set to boost the generalisation ability of the SVMs, resulting in six input parameters (referred to as SVM_label). The geographical label referred to the geographical location of the pixel: label 'one' was added as additional input information for pixels from site A and label 'two' was given to pixels from the site B. To allow appropriate training for this input data sets, training pixels of both sites had to be included in the training process.

The success of a classification algorithm depends on the quality, quantity and suitability of the selected training data. The chosen training data sets have to be representative as support vectors, only then will the position of the optimal hyperplane represent the characteristics of the vegetation and will be able to be used for transferability to classify vegetation of other geographical areas. An intensive field campaign was conducted at a time concurrent with image acquisition to provide training and validation data for the classification. Two randomly chosen field sites within the study site provided areas for the collection of field data, the larger one being named here Site A, while the second one was carried out at a site 8 km away (referred to as site B). Both areas are similar in terms of vegetation types and altitude, and are relatively flat, thus minimising spectral variation due to anisotropic reflectance effects. Mixed pixels were however excluded from the classification with SVMs.

Seven different upland classes were introduced with the SVM classification scheme, depending on the available ground data: *Calluna vulgaris*, mire (cotton grass), bogmire, bracken, acid grass, *Molinia caerula* and *Juncus* species. All data sets were separated into approximately 2/3 training data and 1/3 validation data. Each data set was normalised and rescaled between -1 and 1 to ease the execution of the SVM classifications.

The statistical properties of each pixel in the training data (e.g. standard deviation and mean) have been found to affect the classification accuracies of SVMs (Foody et al., 1995). The deviation from the mean was used as a guideline to include border pixels, covering the whole spectral range of each waveband for each class and identify the pixels which will be difficult to classify and may therefore characterise the support vectors.

5. RESULTS

The SVMs were applied in a supervised classification approach. Firstly, the SVMs, of each design, which resulted in the highest classification accuracy for the training data, will be described, followed by those which resulted in the highest classification accuracy for the validation data and, if applicable, the ones with the most successful transferability. Each of these SVMs was also evaluated against the other data sets and the resulting accuracies are described. Due to space limitations this paper only presents the results achieved with the SVMs based on an RBF kernel.

5.1 Classification with SVM_original classifier

At first, all SVMs were trained with data extracted from both geographical training locations (Site A and Site B).

5.1.1 SVM_original Classifiers Resulting in the Highest Accuracies for the Training Data: The SVMs classified the training data from both training sites with a maximum kappa coefficient of 99.89% (Table 1). The averaged classification accuracy of the training data over the best four performing SVM classifiers consisted of a kappa coefficient of 98.8%. The same selected networks were also applied to the validation data, consisting of pixels from both training sites. The SVM classifiers, which resulted in the highest kappa value for the training data (99.89%), resulted in a kappa coefficient of only 70.92% for the classified validation data (Table 1).

RBF	Kernel	Training	Validation
C	γ	data	data
2	500	99.05%	49.54%
50	100	99.37%	71.29%
100	**100**	**99.89%**	**70.92%**

Table 1: Kappa coefficient for the three best performing SVM_original classifiers, with SVM resulting in the highest kappa value shown in bold.

5.1.2 SVM_original Classifiers Resulting in the Highest Accuracies for the Validation Data: Similarly the SVM_original classifiers resulting in the highest accuracies for the validation data (unmixed pixels from both training sites) were analysed. The best SVM classifiers classified the validation data to a kappa coefficient of 81.22% (Table 2). In addition, the best performing SVMs for the validation data were applied to classify the training data. The SVM classification, which had resulted in the highest accuracy for the validation data, classified the training data to an accuracy of kappa values of 81.27% (Table 2).

RBF	Kernel	Training	Validation
C	γ	data	data
5	5	81.39%	80.85%
10	**3**	**81.27%**	**81.22%**
20	**2**	**81.48%**	**81.21%**

Table 2: Kappa coefficient for the best three performing SVM_original classifiers, with SVM resulting in the highest kappa value shown in bold.

5.2 Transferability of Support Vector Machines (SVM_transfer):

The second part of this paper examines the transferability potential of SVMs. It involved the training of SVMs only on pixels from site A followed by the application of those SVMs to previously unseen data from site B.

5.2.1 SVM_transfer Classifiers Trained on the site A: As expected, the accuracy with which the training data were classified was high, with the highest kappa coefficient being 96.91% (Table 3). On average, only considering the best four performing SVMs, 91.3% (kappa coefficient - 87.1%) of pixels were correctly classified with the SVM. The SVM classifier, which classified the training data to the highest accuracy, succeeded in the classification of the validation data only to a kappa value of 35.17% (Table 3).

RBF	Kernel	Training	Validation	Site B
C	γ	data	data	data
2	2	78.98%	85.35%	14.44%
1	200	94.07%	69.06%	0.83%
1	**600**	**96.91%**	**35.17%**	**0.62%**

Table 3: Kappa coefficient for the best three performing SVM_transfer classifiers, with SVM resulting in the highest kappa value shown in bold.

The ability of these SVMs to classify previously unseen data from a remote location was evaluated, when the classifiers were applied to the site B. The success of the classification was very limited: if the site B pixels were classified with these SVM_transfer classifiers, the classification was limited to a kappa coefficient of 14.44% or less for the best performing SVM classifiers.

5.2.2 SVM_transfer Classifiers Resulting in the Highest Accuracies for the Validation Data: The SVM_transfer classifiers which resulted in the highest accuracies for the validation data were analysed in the same way. The highest accuracies observed for the classification of the validation data was a kappa coefficient of 86.51%. On average, the validation data were classified with a kappa statistic of 85% when considering the four best performing SVMs (Table 4). The selected SVM_transfer classifiers of the highest accuracies for the validation data were also applied to the training data. A kappa value between 84.88% was calculated for the training data if classified with the SVM classifier, which produced the highest accuracy for the validation data (Table 4).

The kappa value of the site B pixels, classified with these selected SVM classifier, was 9.13% (Table 4).

RBF	Kerne l	Training	Validation	Site B
C	γ	data	data	data

0.5	2	76.12%	85.56%	23.43%
1	2	78.29%	85.35%	19.48%
1	**20**	**84.88%**	**86.51%**	**9.13%**

Table 4: Kappa coefficient for the best three performing SVM_transfer classifiers, with SVM resulting in the highest kappa value shown in bold.

5.2.3 Transferability of the SVM_transfer Classifiers: Finally, the SVM_transfer classifiers were assessed when classifying the pixels from the site B data set. The highest kappa value achieved for unseen pixels over all trained SVM_transfer classifier was 29.04% (Table 5). A high classification accuracy for the unseen pixels also stood for a high classification accuracy for the validation pixels from the site A, based on the generalisation of the SVM_transfer classifiers. The kappa values of the validation data were 82% for the SVM classifier, which resulted in the highest classification accuracy for the unseen data (Table 5). The accuracy of the training data, when classified with the chosen SVM_transfer classifiers showed lower kappa values of 73.75% for this SVM classifier.

RBF	**Kernel**	**Training**	**Validation**	**Site B**
C	**γ**	**data**	**data**	**data**
0.50	0.5	72.04%	79.88%	28.05%
0.50	**0.75**	**73.75%**	**82.04%**	**29.04%**
0.50	1.00	75.48%	83.46%	27.18%

Table 5: Kappa coefficient for the best three performing SVM_transfer classifiers, with SVM resulting in the highest kappa value shown in bold.

5.3 SVM_label Classifiers Enhanced with a Geographical Label

An enhancement to the Support Vector Machines was carried out by including a label related to the geographical location of each pixel (SVM_label).

5.3.1 SVM_label Classifier Resulting in the Highest Accuracy for the Training Data: The highest kappa values for the training data classified with the SVM_label classifiers was 92.60%, resulting in an averaged kappa value of 89.4% (Table 6). If the SVM_label classifiers, which classified the training data to the highest accuracies, were applied to the validation data, most of the accuracies exceeded 80% (Table 6).

RBF	**Kernel**	**Training**	**Validation**	**Site B**
C	**γ**	**data**	**data**	**data**
1	50	89.95%	85.47%	74.28%
1	**100**	**92.60%**	**81.64%**	**61.81%**
10	10	88.49%	84.00%	72.19%

Table 6: Kappa coefficient for the best three performing SVM_label classifiers, with SVM resulting in the highest kappa value shown in bold.

A big improvement was found for the classification accuracies of the validation pixels from the site B when classified with SVM_label classifiers. All selected SVM_label classifiers classified those pixels with a kappa value of over 60%. The best four SVMs for the training data resulted in average kappa value of 70.9%. The SVM classifier, which resulted in the highest accuracies for the training data, classified the pixels from the site B to 70.18% (kappa value 61.81%) (Table 6). The integration of the geographical label, therefore, seemed to support the generalisation of the SVMs, as is shown further below.

5.3.2 SVM_label Classifiers Resulting in the Highest Accuracies for the Validation Data: This study further investigated the SVM_label classifiers which resulted in the highest accuracy for the validation data, consisting of pixels from both test sites. A maximum kappa value for the validation data of 86.6% was achieved (Table 7). The selected SVM_label classifiers resulted in an average kappa value of 86.1% for the SVM for the validation data.

RBF	**Kernel**	**Training**	**Validation**	**Site B**
C	**γ**	**data**	**data**	**data**
0.5	**5**	**80.51%**	**86.61%**	**80.87%**
1	10	82.86%	86.43%	79.75%
3	5	84.14%	86.24%	78.57%

Table 7: Kappa coefficient for the best three performing SVM_label classifiers, with SVM resulting in the highest kappa values shown in bold.

The SVM which classified the validation data to highest values also classified the site B pixels to the highest accuracies, namely a kappa value of 80.87% (Table 7). The averaged accuracies for the pixel for the selected SVM_label were 79.5%. This showed that the training data class accuracies were lower than their maximum possible classification accuracy, when classified with SVM_label classifiers, but this decrease in classification accuracy, on the other hand, increased the generalisation ability of the SVMs.

6. DISCUSSION

The following section discusses the potential of SVMs for the classification and transferability of knowledge within images. The classification and transferability performance of SVMs is affected by different design and training parameters: SVMs design, number of input parameters and output classes. Due to the restricted length, however, this paper is not sufficient to analyse each of those parameters in detail and the number of output classes was not changed during this paper, leading to no conclusions regarding the influence of output classes on the performance of SVMs. Therefore findings are just presented as bullet points in this publication:

- The best choice of the *C* and γ parameter, to enable either the highest classification accuracy or the highest transferability ability, depends on trial and error to gain the optimal design of the SVMs. The recommendation of higher values for both parameters resulting in a higher classification potential can not be applied if a SVM with a high transferability ability is required (
- Figure **3**). Smaller values for the *C* and γ parameters are required to achieve the highest transferability ability, although this may lead to poorer accuracies for the training and validation data than theoretically achievable with SVMs.

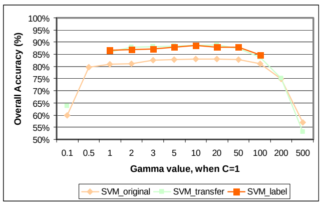

Figure 3: Influence of the design parameter γ for the classification accuracy of SVMs for the training data

- Intensive trial and error tests have to be carried out for the design of the SVMs prior to the training. SVMs are, consequently, more time intensive at this stage than ANNs.
- The use of ancillary input parameters, e.g. a geographical label, enhanced the generalisation performance of SVMs, resulting in higher accuracies for the validation data than without it. The analysis of the classification accuracy highlighted the capabilities of ancillary input variables for both the classification and transferability potential of SVMs.

6.1 Discussion of transferability of the classification methodologies

A classification algorithm consisting of a high generalisation ability is required to enable the transfer of knowledge obtained from one (or few) sites, to across a whole image, but also, ideally, across images, time and scale. An 'ideal' remote sensing classifier would be trained on a limited number of pixels from one training site within the area of interest and would be able to classify the remaining area to accuracies around 85%, the standard remote sensing classification accuracy. This paper reported on research into the suitability of SVMs for mapping vegetation at a remotely located area, requiring the classifier to be able to transfer the knowledge obtained during the training process. The following conclusion can be drawn:

- Caution has to be applied in choosing the design parameters for SVMs to achieve a high transferability, with large values for both design parameters decreasing the generalisation. Additionally, it showed that SVMs resulting in the highest kappa coefficient of the training data resulted in a lower kappa coefficient for the validation and transferability data. No overall rule could be defined for the best choice of those parameters.
- It was observed that the integration of even just a small number of pixels from the remotely located training area resulted in a improvement in classification accuracy for the pixels from this site, not encountered during the training process.
- The performance of SVMs is affected by the support vectors, therefore, the sample points allocated as support vectors are the most important points, when defining the optimal separating hyperplane to minimise the probability of the classification error. The training data is therefore required to be representative of the whole data set.

CONCLUSION

The potential of SVMs for the mapping of upland vegetation was investigated because the high generalisation ability of this classifier has recently been recognised in remote sensing. Overall, the classification and transferability results of SVMs showed very promising results, highlighting the capability and suitability of SVMs for use in remote sensing classification. On the other hand, certain issues of transferability in classification accuracy persist when using SVM classifications which may limit the utility for mapping from high spatial resolution remotely sensed data. Spectral variability within high spatial resolution pixels caused by different vegetation types can decrease the chance of finding the optimal hyperplane. Additionally, vegetation rarely exist within a monoculture, but rather occurs as a mixture of several species, resulting in mixed pixels for which classification SVMs are not suitable.

In general, the success of the transferability was limited when each classifier was applied to a location from outside the area from which the training data had been extracted. The transferability of SVMs could be greatly increased with ancillary input parameters, showing the opportunity for developing methodologies for within image, and potentially across image, time and scale classifications. It highlighted the potential and advantages of the use of SVMs for mapping, i.e. vegetation classes, from remotely sensed imagery. Further research into the optimal design and suitability of SVMs for different land cover applications has still to be carried out.

REFERENCES

Benediktsson, J.A., Swain, P.H. and Ersoy, O.K., 1990. Neural network approaches versus statistical methods in classification of multisource remote sensing data. IEEE Transactions on Geosciences and Remote Sensing, 28(4): 540-551.

Cortez, L., Durão, F. and Ramos, V., 1997. Testing some Connectionist Approaches for Thematic mapping of Rural Areas. In: I. Kanellopoulos, G.G. Wilkinson, F. Roli and J. Austin (Editors), Neurocomputation in Remote Sensing. Springer-Verlag, Berlin, pp. 142-150.

Foody, G.M., Boyd, D.S. and Cutler, M.E., 2003. Predictive Relations of Tropical Forest Biomass from Landsat TM Data and their Transferability between Regions. Remote Sensing of Environment, 85(4): 463-474.

Foody, G.M. and Mathur, A., 2004. Toward Intelligent Training of Supervised Image Classifications: Directing Training Data Acquisition for SVM Classification. Remote Sensing of Environment, 93: 107-117.

Foody, G.M., McCulloch, M.B. and Yates, W.B., 1995. The Effect of Training Set Size and Composition on Artificial Neural Network Classification. International Journal of Remote Sensing, 16(9): 1707-1723.

Fukuda, S. and Hirosawa, H., 2001. Polarimetric SAR Image Classification Using Support Vector Machines. IEICE Transactions on Electronics, E84-C(12): 1939-1945.

Gualtieri, J.A., Chettri, S.R., Cromp, R.F. and Johnson, L.F., 1999. Support Vector Machine Classifier as Applied to AVIRIS Data, Eighth JPL Airborne Earth Science Workshop.

Huang, C., Davis, L.S. and Townshend, J.R.G., 2002. An Assessment of Support Vector Machines for Land Cover Classification. International Journal of Remote Sensing, 23(4): 725-749.

Perkins, S., Harvey, N.R., Brumby, S.P. and Lacker, K., 2001. Support Vector Machines for Broad Area Feature Classification in Remotely Sensed Images, Proc. SPIE 4381, 9/04.

Roli, F. and Fumera, G., 2000. Support Vector Machines for Remote-Sensing Image Classification, EOS/SPIE Symposium on Remote Sensing, Barcelona, Spain.

Woodcock, C.E., Macomber, S.A., Pax-Lenney, M. and Cohen, W.B., 2001. Monitoring Large Areas for Forest Using Landsat: Generalization across Space, Time and Landsat Sensors. Remote Sensing of Environment, 78: 194-203.