Journal of Coastal Conservation 10: 53-64, 2004

Using earth observation data from multiple sources to map rare habitats in a coastal conservation area

Thackrah, Graham[1*]; Rhind, Peter[2,3]; Hurford, Clive[2] & Barnsley, Mike[4]

[1*]*Department of Geography, University College London, London, WC1E 6BT;* [2]*Countryside Council for Wales, Plas Penrhos, Ffordd Penrhos, Bangor, Gwynedd, LL57 2LQ, UK;* [3]*E-mail p.rhind@ccw.gov.uk;*
[4]*Department of Geography, University of Wales Swansea; E-mail m.barnsley@swan.ac.uk;*
[*]*Corresponding author; E-mail gthack@geog.ucl.ac.uk*

Abstract. Kenfig NNR (National Nature Reserve) is a coastal sand dune system in south Wales, UK. The site is an important location for the conservation of the fen orchid *Liparis loeselii*, a significant proportion of the UK population is found solely on the site. Approaches to the mapping and monitoring of the habitats at Kenfig NNR using EO (Earth Observation) methods are investigated.

Typical airborne EO missions over such sites produce more than a single source of EO data; these may include various optical imaging sensors with different spectral ranges, film cameras and ranging devices to measure topography. Conservation managers are thus presented with the problem of which sources of data to use when producing a land cover map of the site of interest.

Using a data set gathered over the Kenfig NNR site, we investigate land cover mapping methods for conservation. The land cover types of interest typically cover small areas within a much larger site so they present a hard problem for the EO data and associated classification methods to solve. Land cover classifications produced from the data sets provide a set of competing hypotheses of land cover type for the site.

Methods we use to resolve this competition between the data sets include voting methods, data fusion methods and a method utilising fuzzy logic to aggregate information. This paper is intended to act as an introduction to some of the issues involved in using EO data for habitat mapping in highly heterogeneous coastal dune environments and to present some preliminary results of the performance of each method.

Keywords: Dune slack; Kenfig NNR; *Liparis loeselii*; Vegetation survey.

Abbreviations: ATM = Airborne Thematic Mapper; CASI = CCW = Countryside Council for Wales; Compact Airborne Spectrographic Imager; EO = Earth Observation; cSAC = candidate Special Area for Conservation; NNR = National Nature Reserve; OWA = Ordered Weighted Average; SSSI = Site of Special Scientific Interest.

Introduction

As a cSAC (candidate Special Area for Conservation), there is a legal requirement for the monitoring of the reasons for notification of cSAC status at Kenfig NNR. This is under the European Union's Council Directive 92/43/EEC of 21 May 1992, more commonly known as the EU Habitats Directive. Furthermore, under the UK's Wildlife and Countryside Act (HMSO 1949, 1981) Kenfig has both NNR (National Nature Reserve) and SSSI (Site of Special Scientific Interest) status. Both of these designations also have legal requirements to monitor the status of the site in terms of its conservation value.

The use of EO data in mapping and monitoring coastal habitats to comply with conservation legislation has several potential advantages over traditional methods. The latter, typically ground based survey, have two main drawbacks, speed of execution and repeatability (McGwire 1992). EO methods solve both of these problems whilst introducing some of their own. It is for conservation managers and practitioners to decide if the benefits of using EO data outweigh the drawbacks.

Whilst not removing the need for ground based survey, EO data used for mapping and monitoring purposes do repay the initial effort with a fast repeatable method of land cover estimation over a site, free from problems associated with different human surveyors.

Typically providing complete coverage over a site, EO data are capable of producing a detailed land cover map over the area of interest. Due to the usual highly heterogeneous distribution of the land cover types in coastal systems there is a requirement for high spatial resolution data (better than 5 m in a UK coastal dune system) if a derived land cover map is to adequately represent the true nature of the coastal dune habitat mosaic. Until recently, space-borne sensors suitable for mapping land cover in natural and semi-natural areas are relatively coarse resolution (20 to 30 m).

More recently imaging platforms, such as the

IKONOS platform, offer data of 1 m resolution; the data are limited in their spectral resolution however. High spectral resolution is another requirement of accurate land cover mapping, particularly in coastal areas, as the land cover types are often similar spectrally and may only be distinguished using data with a high spectral resolution. For these reasons airborne sensors (capable of easily obtaining data better than 5 m resolution) were used in this study to produce land cover maps of Kenfig.

Due to the way in which natural and semi-natural areas develop throughout the growing season, it may be advantageous to acquire image data at more than one time in the year. Traditionally EO surveys of vegetation particularly have been undertaken during summertime, when higher plants are in flower. More information on land cover types may be acquired during the seasons of winter, spring or autumn

It is often easy to mount more than one imaging system on-board an aircraft when carrying out a survey of an area. Indeed, due to the high fixed costs involved in carrying out an airborne survey of an area, the extra cost of mounting more than a single imaging system on board an aircraft for such a purpose is negligible. After such a survey has been carried out there are therefore two or more data sets for analysis.

As a result of all this, quite often there may be multiple data sets available for analysis for a given land cover mapping exercise in a coastal area.

When one is faced with this situation, what should conservation managers and practitioners do with the separate EO data sources? It is possible to analyse just a single data source and produce a single land cover classification. This has the unfortunate consequence of ignoring all the information contained within the other sources available. It is also possible to analyse each data source available and produce a number of land cover classifications. This has an equally undesirable consequence, that of which land cover classification to utilize in a monitoring and mapping exercise. A third alternative is to analyse each data source together and produce a single, or consensus, land cover classification. The purpose of this paper is to investigate methods of achieving this latter option in the hope that an improved land cover classification will be better than any of the classifications produced from a single source.

The hypothesis investigated is therefore, that by combining data from multiple sources, from multiple times during the year, a better land cover classification may be derived than from any single source alone. Throughout the rest of this paper we will discuss combining data and information. Sources referred to will be image data from different sensors at different times of the year.

Site description

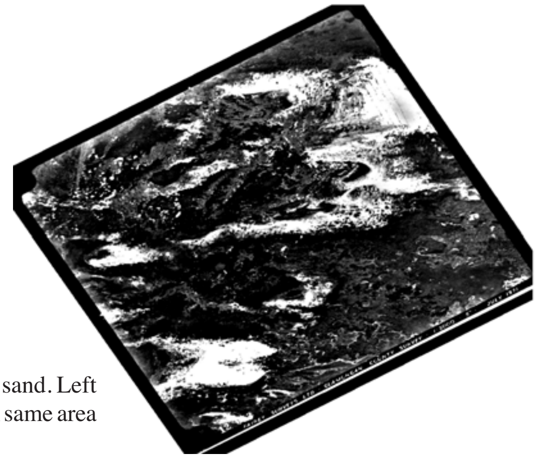
Kenfig NNR is a coastal sand dune system, on the south coast of Wales, between the cities of Cardiff and Swansea, in the mouth of the Bristol Channel. It is situated just to the south east of a large steel works at Port Talbot, and is bordered all along its north eastern edge by a major motorway. The southeastern boundary is a golf course and farmland and the southwestern edge is the coast.

The site is approximately 4 km long and 2 km wide, totalling over 600 ha. Most of the site is covered with relatively old, inactive dune formations. Habitats on the site are mostly typical dune types, some of the more mature dune systems on the most landward edge of the site are particularly mature woodland and bear little resemblance to coastal dune vegetation.

Consisting mostly of dune ridge and dune slack vegetation (dune slacks are the flat areas on the seaward side of a parabolic, landward moving, dune ridge) as well as much dune grassland, the dunes contain habitat representatives relatively rare on a UK and European scale. The younger humid dune slacks are typical of more active dune systems and their presence at Kenfig NNR is the major reason for its notification and protection under the various forms of legislation. Kenfig NNR is one of the best sites in Wales for the presence of these humid dune slacks (Natura 2000 code 2190: humid dune slacks). The site hosts around $^{3}/_{4}$ of the UK population of the fen orchid, *Liparis loeselii* and is an important site for the petalwort *Petalophyllum ralfsii* (under Annex 2 of the EU Habitats Directive). The site is managed by Bridgend borough Council, on behalf of CCW (The Countryside Council for Wales, Cyngor Cefn Gwylad Cymru).

Photographic evidence from the 1940s, 1950s, 1960s and 1970s (see Fig. 1) shows that there were significant areas of open sand (ca. 30-50%) during the 50 years before 1997. Due to the increasing maturity of the site since the 1940s, open sand no longer covers much of the site, an approximate estimate would be less than 5% by area. The general consensus amongst those responsible for site management is that it is likely that the amount of humid dune slack on the site has also decreased during this time.

Also present on the site is a large natural lagoon, the largest in south Wales. This is also a proposed feature for notification of cSAC status (as an example of Natura 2000 code 3140, hard oligo-mesotrophic waters with benthic vegetation of *Chara* spp. Running the whole seaward length of Kenfig NNR, just on the landward side of the principal dune ridge, is a derelict tarmac road. The fact that this road is *not* inundated with sand is a clear indication that the contemporary landward trans-



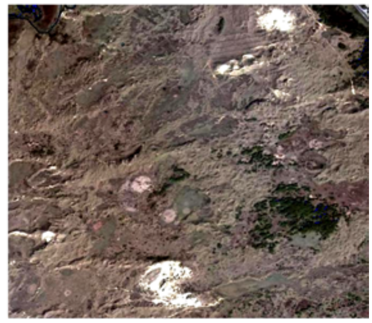


Fig.1. Comparison of area of open sand. Left image from 1971, right image from same area in 1997.

port of sand onto the site is minimal. This fact also suggests a reason for the lack of contemporary dune activity at the site.

The usual procession of vegetation types in the slack habitats is from bare sand to a low growing sparse assemblage of *Salix repens* (Creeping willow) with a covering of lichens, liverworts and moulds, giving a dark appearance to the sand substrate. During the winter these damp slacks flood with water, sometimes up to a several metres deep. This regular flooding during winter helps to keep the vegetation in this state. Over time, the species begin to change in their representation in the assemblage, giving way to more vascular plants with a denser cover and more soil development. This process is restarted with new dune formation as a result of storm damage to mature dunes or inundation with blown sand. It is assumed that both of these processes have been much reduced at the site since the late 1940s.

It is these processes of succession that threaten the rare habitats on the site. The condition of the humid dune slacks at the site is a current cause for concern for the managers of the site. It was this concern that was the stimulus for this current research.

Contemporary management at the site has concentrated on removal of the vegetation and top few centimetres of soil in the more mature dunes using earth excavation equipment. Agreements with the local steel works for the loan of caterpillar tracked bulldozers, and dumping of topsoil on the spoil heaps of the works has helped in this regard (C. Hurford pers. comm. 2004). Rhizomes of *Liparis loeselii* subsisting in the soil/sand substrate below this level then have a chance to regrow in the subsequent growing seasons.

In order to assess the success of this and other management at the site, some means of monitoring the site is required. We hope to show that EO data may be used as part of this process.

Data and data processing

The two optical data sets used in this study, CASI (Compact Airborne Spectrographic Imager) and ATM (Airborne Thematic Mapper) were acquired on-board the NERC (Natural Environment Research Council) aircraft. This platform is equipped with DGPS (Differential Global Positioning System) and an inertial navigation system to simultaneously acquire aircraft attitude and position information with the image data. NERC supply the raw image data, together with this platform positional data, and an application program to correct the image data for errors in platform attitude. The correction for attitude is, at the time these data were acquired, more accurate than the absolute positional accuracy. As such, geometric errors in the data are better corrected than absolute positional errors.

The CASI and ATM data were acquired on two separate flights, one in May and one in August 1997. Both sensors were flown on both flights, producing four separate image data sets.

The ATM instrument was mounted on a flexible mounting, allowing up to 15° of movement in the across track direction. The CASI instrument is rigidly mounted in the aircraft platform. This means that the ATM data initially has fewer geometric errors than the CASI data.

When inspecting both the CASI and ATM data initially it is clear that there are variations in the radiance values in the across track direction. This is due to the anisotropic reflectance of the land surface imaged by the sensor. The shape of the function describing this anisotropy (termed the BRDF, or Bi-Directional Reflectance Distribution Function) may be complex but in this case, the major variation is an increase in brightness in the across track direction away from nadir. This may be adequately characterized by a second degree polynomial function, and the gross effects therefore removed. This was done for the image data using the ENVI image

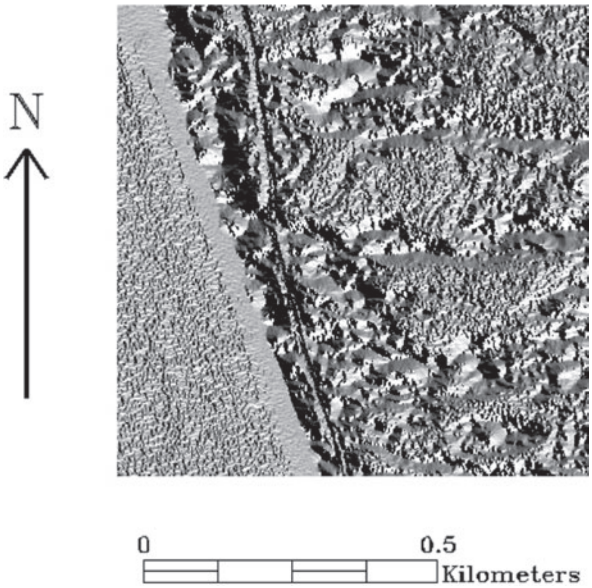


Fig. 2. Aspect, derived from LIDAR DEM.

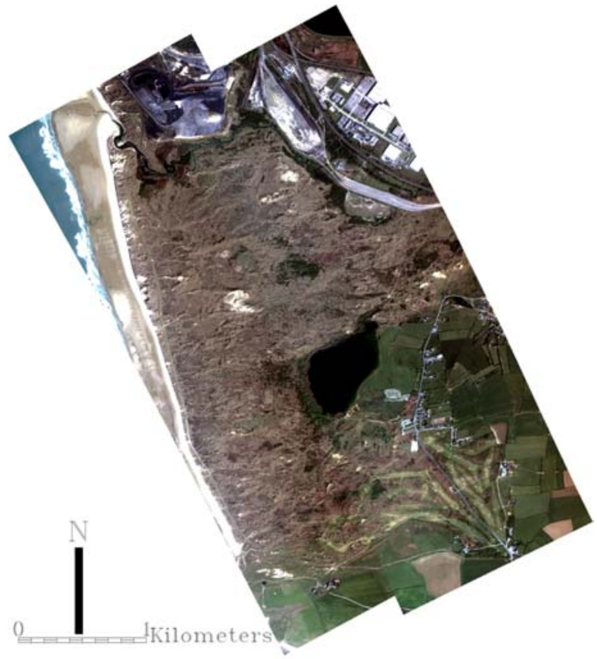


Fig. 3. TM mosaic of Kenfig NNR.

processing software. Although variations in image brightness are still apparent (see Fig. 3), they are somewhat reduced.

After this simple geometric and radiometric correction, the separate image strips covering Kenfig NNR were mosaiced together, producing a single image for each imaging sensor, for each date (four optical image data sets in total).

LIDAR (LIght Detection and Ranging) data were also available for the Kenfig NNR site. These data were made available to CCW by the UK Environment Agency (see Fig. 2, taken from midway down the coastal edge of the site). The LIDAR data are usually supplied either in the raw (x,y,z) triplet, or an already interpolated grid of data. In this case the LIDAR data for the site were supplied in the raw triplet form and were interpolated to a regular grid using the ArcView GIS software from ESRI.

To increase the utility of the elevation data, a land form classification was then produced. The method is after Wood (1996). The basic procedure is to classify each pixel into one of the following six classes (the extra band of information in the elevation data set is therefore a categorical variable); peak, ridge, pass, plane, channel and pit. The classification is done on the basis of the local surface neighbourhood of the input pixel. The partial derivatives of a polynomial fitted to the surface in the x and y directions are calculated, and the classification made on the basis of this and a set of rules. This and further processing to retrieve aspect and slope was carried out using the ENVI image processing software from Research Systems Incorporated. The resulting topographical data set consisted of a 4-band image, band 1 being the elevation data for each 2-m pixel, band 2 being aspect, band 3 was slope and band 4 topographic category. This data set is therefore a mixture of continuous and categorical data. The continuous data further being a mix of both approximately normally distributed data (elevation) and heavily skewed (aspect and slope).

All data were interpolated to a 2-m grid cell (pixel) size to ensure they, and any derived image products, could be directly compared to each other spatially.

The ground truth data were acquired during the period between 1997 and 2000. Where possible, DGPS positional data for the ground truth data were simultaneously acquired. The ground truth data acquired were floristic data recorded for a 1-m^2 quadrat. The recorders were all experienced botanists so that inter-recorder differences would hopefully be minimal. The final land cover map is therefore still subjective, but it is repeatably so. This is not the case with a land cover map derived entirely from land based survey as the initial data collection fully determine the final land cover map. This is in contrast to the case of using EO data to derive a land cover classification where the image data may be re-analysed using alternative ground truth data and the results compared.

Accurate co-registration of image data sets and ground truth data is critical for the purposes of using such data to train most image classifiers (Congalton 1991). Furthermore, for all the methods of data and information fusion outlined in the next section to be applied to these data correctly, each data set must be adequately registered to each other data set. This is to ensure that the land cover class for a given pixel in a given data set is representing the same area of the study site as the conjugate pixel in every other data set. The effect any misregistration has on the results of the experiments carried out here is discussed further in the Methods section.

Data set co-registration was carried out by firstly correcting each data set using the supplied correctional software. Subsequent to this, each data set was co-registered to the May ATM data using a network of ground control points. This procedure was followed as it meant the minimum of corrections applied to each image to achieve co-registration of every data set. Operators equipped with DGPS equipment often gathered the ground truth data. This facilitates the co-registration of these data with the image data. Both the quality of classifier training and the assessment of classifier accuracy may be adversely affected by an inability to accurately associate ground truth data with image data (Congalton 1991).

Methods

Methods to make use of information from multiple sources split easily into two main categories, data fusion and information fusion (Thackrah 2001). The former is the process by which data from multiple sources are combined in a single, synthetic, data set to be analysed singularly. Information fusion, by contrast, involves the separate analysis of each constituent data set, followed by some means of combining the results of each. In the case of land cover mapping using classified image data, the former case is realized by simply constructing a synthetic data set of the data values of all the sources. In the latter case the classifications of each data source are combined to produce a single, or consensus, classification.

Only one method of data fusion is investigated here, the synthetic data set method. In order for the elevation data to be utilized as well as the image data, a non-parametric classifier must be used however. A parametric classifier will perform poorly when given categorical data as well as continuous data, or when the data are not all distributed identically in the same multivariate model. The elevation data set consisted of various measurements, some of which are normally distributed, some of which were not. In this case, a typical parametric classifier, such as the maximum likelihood classifier, would not classify the data set well. The non-parametric classifier chosen for this was an artificial neural network. A simple multi layer perceptron ANN was used in this study, with error back propagation and fixed learning rate and momentum terms after Rumelhart et al. (1986). The ANN was trained using the same training data as the maximum likelihood classifier on the optical data.

As a first step in investigating methods of information fusion, the four optical data sets were classified separately, using the available ground truth data. This produced a set of constituent land cover classifications, one for each source of EO data. The maximum likelihood classifier was chosen for this purpose as it is widely available in many image processing software packages and is the most likely method to be utilized by conservation professionals undertaking such work. Both the final classifier output and the *a posteriori* probabilities of class membership (Foody et al. 1992) results were produced. These results are to be used as inputs to the various information fusion methods described next to derive a single consensus classification from the four constituent classifications. The elevation data set was also classified on its own, using the ANN method due to the presence of categorical data within it, and was also used as a constituent classification.

Data set registration

The main difficulty, with the use of EO and data or information fusion and with the use of ground truth data in classifier training is accurate data set registration. The constituent data sets *must* be registered accurately to each other. If this is not the case then errors in the resulting classification may be a result of misregistration, rather than the fault of the classification or fusion process (Congalton 1991). It is difficult to adequately geocorrect and co-register high resolution data sets acquired by airborne sensors. By adequate, it is essential that such data sets are co-registered to within one or two pixels. In the current case, the co-registration accuracy was of the order of 7 or 8 pixels.

To provide an approximate idea how this level of error would effect subsequent analyses, a sample classification was misregistered with itself by up to seven pixels. Using one as a base classification and the other as a 'ground truth' classification, the inter-classification accuracy was calculated. The results are shown in Table 1. As can be seen from the small number of examples in the table, the misregistration had a large effect on classification accuracy in each case. It must be remembered that this is the result of misregistering the *same* classification with itself and the best result was that the two classifications had 20 % of the pixels in error.

This effect is particularly severe in the coastal system studied here, as the distribution of land cover types

Table 1. Classification accuracy for misregistration experiment.

Iteration number	Offset (*x*/*y*)	Classification accuracy (%)	Khat
1	1/1	81.24	0.79
2	8/8	71.87	0.56
3	0/3	71.14	0.65
4	3/0	78.29	0.66
5	6/4	73.78	0.59

is highly heterogeneous with proportionally high numbers of edge pixels for each habitat type. This large number of boundary pixels leads to significant inter-classification error when misregistration occurs. More homogenous environments would be less affected by registration errors. With such a high degree of classification error expected purely from misregistration, it is hoped that there may still be quantifiable gains made by the various data and information fusion methods proposed.

By identifying if the methods of data and information fusion work at all well in a coastal system, more homogenous habitats, such as moorland or woodland areas, are expected to be more suited to the methods described.

There is a number of alternative methods for information fusion. Voting methods are perhaps the simplest. Averaging methods are another alternative and more advanced methods based in the field of fuzzy logic may also be employed. Voting methods are discussed in the next section.

Voting methods

Voting methods use the so-called 'crisp' output of a classifier (Foody 1999), that is the final class of a pixel as output from most classification methods. The available data sources are all classified individually and the class of each pixel recorded in an output image as an integer value. The classifications are then registered with each other resulting in a stack of images where for each pixel we have the suggested land cover type for each classified data source.

The two main voting methods are majority and plurality rule both of which are very simple and have very quick processing times given a set of input images. The plurality rule simple assigns a consensus class to a pixel that is the output of most of the individual classifications. Majority rule is identical to the plurality rule with the proviso that if the votes for the most popular class do not exceed 50%, a 'Not Classified' status is assigned to the pixel. The advantage of the latter case is that uncertainty (exemplified by more than $^1/_2$ the individual input classifications disagreeing) is explicitly propagated to the final classified product (albeit in a rather unsophisticated manner). The plurality rule ignores any uncertainty in the input classifications and produces a consensus class for every pixel in the output image.

A drawback with voting methods is their equal weighting of each members vote. The advantage of voting methods as decision making agents in a political democracy, that each vote counts the same amount, is a disadvantage in this case. We may know, for instance, that a particular source of information is more accurate, either overall, or in specific class cases, than another. In this case, the votes from that source are weighted exactly the same as all other sources, this is not a desirable feature of the method.

Ordered Weighted Average operators

OWA (Ordered Weighted Average) operators are a flexible class of operators that may replicate the operation of other aggregation operators (Yager 1988). Yager refers to the process as aggregation, though throughout this paper it has been referred to as fusion, throughout this section only, the term 'aggregation' will be used after Yager's usage. This class of operators are capable of performing aggregations of data organized in vectors in a similar manner to many other aggregation operations (Yager 1988, 1998). It is easy to show that an OWA operator may perform an aggregation in the same way as the median, mode and mean operators perform aggregations. This flexibility is seen as a strength of this class of operators. OWA operators may also be shown to simulate the AND and OR (which are also equivalent to the MIN and MAX operators respectively) logic operators on fuzzy sets. Furthermore they may also be constructed to show different degrees of AND-ness and OR-ness. Using this class of operators one may perform a variety of aggregations on the same vector of data.

$$(1,0,0,0) \times (3,1,1,1) \tag{1}$$

Shown in Table 2 is the example output from four separate classifications. In this case they are classifications of four different data sets by the same classifier. They may also be classifications of the same data set by four different classifiers, or four sets of training data.

The numbers represent the partial output that most classifiers produce. They may be the *a posteriori* class membership values produced from the maximum likelihood classifier or the output activation values of an ANN classifier for example. The output from the maximum likelihood classifier for the first case above would be class 2.

A simple way to use the method proceeds as follows. The test vector to be operated on is first ordered such that, $n_1 \geq n_2 \ldots \geq n_x$, and then multiplied by an aggregation vector of equal dimension to return a single value, thus performing an aggregation of the values in the test vector. The values in the aggregation vector are chosen depending on the aggregation desired. The example shown in Eq. (1) is the plurality rule (tied results are represented by a figure for each of the tied classes, giving a count of 1 for each of the other classes) for the data shown in Table 1. The value returned is class 4, as this is the class with the three votes from the system shown in Table 2.

A more complex way to use OWA operators is to form a vector for each class made up from the classifier outputs of each classifier for that class. These (in this case four) vectors are then multiplied by the chosen aggregation vector, producing a single vector with four members (one for each class, *not* classifier). This resulting vector is then multiplied by the MAX or OR vector – the first vector in Eq. (1) – and the result returned as the output of the system for that pixel. Again, the method is simple and proceeds quickly when applied to co-registered image data.

Using two consecutive MAX operators would return the class in the consensus image that had the highest single *a posteriori* value in the constituent classifications. Using the data from Table 1 this would be either class 2 or 4. Using a mean aggregation operator – Eq. (2) – followed by a MAX operator would return the class with the highest average *a posteriori* value in the constituent classifications. Again, using the data in Table 2 this would return class 4 as the consensus class.

Many different aggregation vectors can be used, the four most common are the mean, median, MAX (OR) and MIN (AND) operators, Vectors shown respectively below in Equations (2), (3), (4) and (5).

$$(0.25,0.25,0.25,0.25) \quad (2)$$
$$(0,0.5,0.5,0) \quad (3)$$
$$(1,0,0,0) \quad (4)$$
$$(0,0,0,1) \quad (5)$$

The values in the first two cases are $1/n$ (Eq. 2) when the number of classifiers is n and 1 or 0.5 when the number of classifiers are odd or even respectively. The size of the vectors is altered to be n when the number of classifiers is n.

In contrast to the case with the voting methods, this method of information aggregation does not assumes that each source is weighted equally. In this case there exists the possibility to differentially weight the sources however. It is possible to weight the members of the OWA vector according to the expected ability of each classifier to classify pixels into a particular class. This would be done on an *ad hoc* basis and there is no particular justification for any one given weighting strategy.

The OWA operators used in this study were the mean, median and MAX on the constituent vectors for each class in the system, followed by a MAX to return the consensus class.

Table 2. Example classifier output.

Class	Classifier 1	2	3	4
1	0.5	0.4	0.2	0.2
2	0.9	0.1	0.1	0.1
3	0.3	0.2	0.2	0.2
4	0.1	0.9	0.5	0.2

Fuzzy integral fusion

Fuzzy integrals, first proposed by Michio Sugeno (Sugeno 1974, 1977), provide an explicit means by which the different abilities of each constituent classification to identify each class may be taken into account in the information fusion process. The method will not be discussed in detail here, see Cho & Kim (1995), Lee & Lee-Kwang (1994); Tahani & Keller (1990); Thackrah et al. (1999, 2000) for details of its implementation in information fusion for image processing. The technique has successfully been used in image processing, particularly in target recognition applications. The use of it in other information fusion applications has been reviewed for the operational research community by Grabisch (1996) however its use in land cover mapping is a novel application.

In brief, the method proceeds as follows. The constituent data sets are each classified separately, as in the previous methods. The accuracy of each classifier in identifying each class is then determined using the available ground truth data. A fuzzy measure for each class, and a resulting score for each classifier, is then constructed using this information. Next, in a similar manner to the use of OWA operators in the previous section, the vector of classification output for each pixel is ordered. The fuzzy measure is then used, together with the information on class membership to arrive at a consensus class for each pixel in the image.

The fuzzy integral is a method by which the interaction between classifiers to be fused is modelled in a more appropriate manner (Grabisch 1996). Instead of the accuracies of each classifier summing proportionately for the combined system, they may sum disproportionately, particularly where two, or more, sources are expected not to add much more information over a single source. An example is the case with the optical data sets. We do *not* expect the accuracy of the emergent classification from the combination of two constituent classifications to be a simple sum of the accuracies of each constituent classification. More likely is a small, disproportionate, increase in overall accuracy. The elevation data, despite the poor ability of it to distinguish between all classes, may well increase the accuracy of the emergent system in a more disproportionate fashion than an initial inspection of its performance alone would suggest. In other words, the poorly performing classification of the elevation data may provide a disproportionate increase in the emergent classification accuracy than its performance alone would suggest.

The data sources used for the voting and OWA methods of information fusion were simply the optical data sources. Two separate experiments were carried out using the fuzzy integral for information fusion. The first utilized the classifications of the optical data only. The classification results of two optical sources, either the multi sensor data for the same time of year, or the multi temporal data for the same sensor, were used (making four combinations in total). The second experiment utilized these same two sources plus the ANN classification results of the optical and the elevation data (again making four combinations in total).

In exactly the same manner as the other methods already mentioned, this approach requires that each classification segments the scene into the same classes. In the case of the elevation data used here, this is not a straightforward task. The optical data has 12 or 14 bands (for ATM and CASI respectively). The elevation data set consists of four bands only. Each source of data was segmented into 16 land cover classes and it is not expected that a 4-band data set is able to segment a scene into 16 classes better than a 12 or 14 band source. Nevertheless it is expected that what little information the classification of the elevation data set brings to the final result will be effectively fused using the fuzzy integral.

An alternative way to proceed is to use the elevation data to filter the initial classification process. Elevation data for example may be used as an initial filter to reduce the potential number of classes that optical data must discriminate between. A good example in the current case is the use of elevation data to distinguish the location of dune slacks. A classification of the optical data must then only distinguish between the different dune slack habitats, rather than all the habitats present at Kenfig NNR. An exploration of this is left for further work.

Results and Discussion

The results for each class are too detailed to show here, but instead we show the total areas classified as each class for each classification and fusion method.

Shown in Fig. 4 is the output of the mean OWA operator. This is included to show that the classification of the digital image data segments the scene well and that the site is highly heterogeneous. It may also be noted that, by comparison with the image data in Fig. 3, the segmentation of the scene passes an initial visual assessment of accuracy.

Classification accuracy assessment

A detailed examination of each classification's performance was not possible, due to the limited availability of ground truth data. All the available data were used for training the classifiers. When all these data were also used for classification accuracy assessments, most came out greater than 95 % accurate. This is *not* expected to be a true measure of the classifiers accuracy, but instead more a reflection on how well the respective classifiers learned the training data. The unfortunate conclusion is that the training data were probably not representative of the area as a whole, which is one of the desirable properties of ground truth data. It was not rational in this case to split the training data into separate training and testing sets; too few data were available. If a representative training data set were available, then a reliable assessment of the overall accuracy of the different classifications could have been obtained by using the same set to test the resulting classification (Stehman 2000).

The usual method of assessing classification accuracy is to produce a confusion matrix. This is a matrix of values where the columns represent the ground truth classes and the rows the classifier output. A correctly classified pixel is added to the total in element (n,n) where n is representative of the class of interest. Incorrectly classified pixels are added to element (n,m) where n is the ground truth class (the correct class) and m is the actual output class. An inspection of the confusion matrix allows detailed analysis of how the classifier performs on a per-class basis. The Khat statistic (Congalton 1991) mentioned in Tables 1 and 6, takes these off diagonal members of the confusion matrix into account of its measure of classification accuracy in contrast to the raw accuracy figure that just calculates the percentage of correctly classified pixels.

The Khat statistic is a measure of classifier accuracy that takes into account misclassified, as well as correctly classified, pixels. As such the Khat statistic is a more realistic measure of classifier accuracy than the raw percentage of correctly classified pixels (Congalton 1991;

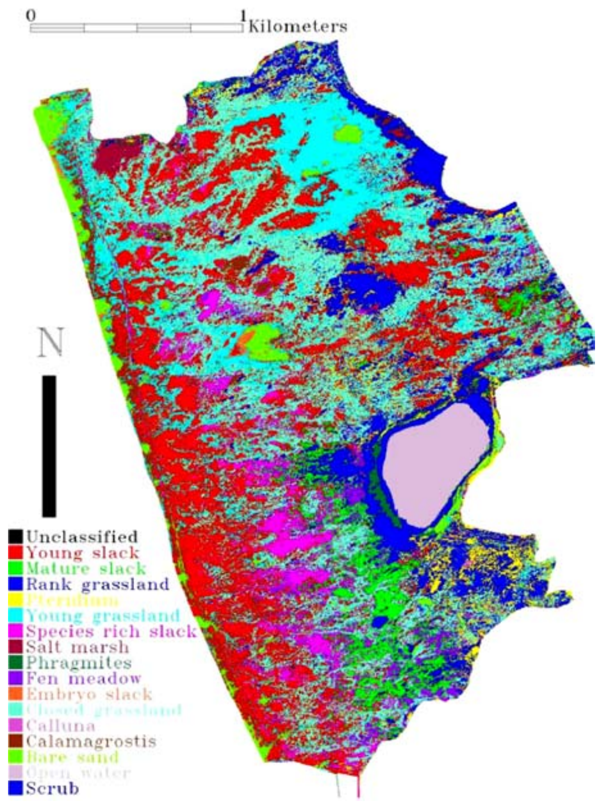


Fig. 4. Example classification output, mean OWA operator.

Congalton et al. 1983; Rosenfield & Fitzpatrick-Lins 1986). For a better assessment of classifier accuracy, the confusion matrices for each classification should be inspected. Space prohibits their printing here and the Khat statistic is included on the understanding that a single measure of classification accuracy is always going to be a compromise compared to an inspection of the confusion matrix. Despite these reservations on classification accuracy assessment, the calculated classifier accuracies are included here in Table 6.

Tables 3-5 show the proportions of each land cover class produced in the classifications. The classes are reasonably self explanatory, *Pteridium*, *Phragmites*, *Calluna* and *Calamagrostis* being either single species stands, or stands dominated by those genera respectively. The remaining classes were broad habitat types based on the site knowledge of CCW staff and the conservation objectives of the site, mainly the humid slack habitats. Table 3 shows the results of the voting method information fusion experiment, the OWA aggregations and the data fusion experiment. Table 4 shows the results of the 2 (optical) source fuzzy integral fusion experiment and Table 5 the 3 (optical plus elevation) source. Table 6 shows the overall classification accuracy and Khat statistic of each of the fusion experiments.

It may be seen from the first table that the land cover proportions of each class are relatively stable, even for the simpler information fusion methods. Approximately 18-23% of the site is covered by young slack habitat. The dominant habitat type appears to be closed grassland, with a wider range of likely coverage at between 20 and 36% of the site. The main land cover types on the site are thus young slack, young grassland, closed grassland and scrub. Young slack is the name given to a slack a successional stage beyond the humid dune slack and is of lower conservation value. The cover types, species rich slack and embryo slack are of high conservation value and have proportions of 5% and < 1% respectively.

The stable land cover proportions between each method also support the use of EO in a mapping exercise in such a heterogeneous semi-natural habitat as Kenfig NNR. The level of habitat classification exceeds that of a Phase 1 habitat survey, a nationally adopted level of biological survey within the UK (Anon. 1993), and provides a more detailed habitat map than the ground survey alone was capable of.

The critical habitat type that has high conservation value, and is essential for the long term presence of humid dune slacks at Kenfig NNR, embryo slack, shows up at between 0.5 and 1% of the site. It is expected that this value was much higher in the past, particularly during the 1940s and 1950s when the area of open sand at the site was much larger than presently. If nothing else, this last figure should provide evidence for the future designation of the site and increasing importance placed on its careful management.

The estimates of bare sand using these methods are typically approximately 2 to 4%. Aerial photography from the 1940s and 1950s shows subjectively at least 30% cover, it is clear that this proportion has decreased substantially over the intervening 40 to 50 years.

Tables 4 and 5 again show broadly similar proportions of the constituent land cover types. Perhaps this is to be expected as they are all based upon the same four input classifications. The land cover proportions of the individual classifications, whilst not included here, were similar. The accuracies of the individual classifications were also very similar to those of the fused classifications. Firm conclusions made on an analysis of the classifier accuracies should not be drawn due to the uncertainty that the ground truth data represent the site as a whole.

Subjectively at least however, the fused classifications displayed a much lower tendency to be affected by the brightness trends that affect classifications of single data sets. This is a valuable difference, these effects are certainly not a result of differences in land cover type, and are therefore an undesirable artefact in a

Table 3. Class total areas (%). Fusion = data fusion.

Class	Majority	Plurality	Mean	Median	Max	Fusion
Young slack	17.97	17.97	22.14	23.12	19.76	13.96
Mature slack	3.15	3.15	4.15	4.22	3.89	2.42
Rank grassland	2.76	2.76	3.94	3.99	3.86	8.01
Pteridium	1.95	2.02	3.42	2.92	4.65	1.42
Young grassland	9.41	9.59	10.50	10.47	10.85	12.56
Species rich slack	4.07	4.71	5.17	4.98	5.97	4.70
Salt marsh	1.58	2.23	2.30	1.90	3.81	1.07
Phragmites	0.57	0.83	0.94	0.76	1.48	6.27
Fen meadow	2.21	4.10	3.09	2.80	3.66	9.90
Embryo slack	0.37	0.43	0.53	0.41	0.82	2.70
Closed grassland	25.29	36.00	27.09	29.14	20.91	7.19
Calluna	0.04	0.42	0.16	0.05	0.40	2.42
Calamagrostis	2.65	7.56	3.33	3.33	3.87	11.54
Bare sand	2.11	3.31	2.65	2.17	3.39	4.34
Open water	3.85	3.85	3.44	3.53	3.54	5.63
Scrub	6.48	10.62	7.14	6.22	9.14	5.86
Unclassified	15.53	NA	NA	NA	NA	NA

Table 4. Class total areas (two sources) (%).

Class	May	August	ATM	CASI
Young slack	25.88	21.01	19.15	35.06
Mature slack	7.98	3.36	5.54	2.47
Rank grassland	2.80	4.25	3.80	3.94
Pteridium	1.81	5.79	5.35	1.85
Young grassland	8.53	5.93	5.76	8.65
Species rich slack	1.96	5.60	2.73	1.80
Salt marsh	1.36	1.16	1.10	0.99
Phragmites	0.68	0.43	0.44	0.45
Fen meadow	3.37	0.90	0.82	1.83
Embryo slack	0.13	0.31	0.15	0.30
Closed grassland	30.67	41.93	44.88	31.17
Calluna	0.04	0.00	0.27	0.01
Calamagrostis	7.02	1.09	2.32	3.83
Bare sand	1.01	1.30	1.11	1.37
Open water	3.53	3.53	3.55	3.53
Scrub	3.24	3.40	3.05	2.75

land cover classification. It may be seen from the classification in Fig. 4 (the result of the mean OWA aggregation of all four digital image classification results) that it shows no banding artefacts that could be associated with the brightness trends apparent in Fig. 3. The single, non fused, classification of the image data in Fig. 3 did show such obvious banding patterns. A more detailed and less subjective analysis of this is needed to draw firm conclusions, however this aspect of the fusion results is a promising one.

Due to the limited availability of adequate ground truth data in order to carry out a detailed assessment of the accuracy of each of the fused classifications, perhaps the most important conclusion to reach from these results is the similarity between the different methods. It may be seen that the general proportions of each land cover class are relatively stable between all the classification results. This is important information that can be used to good effect by the site managers. The status of the site may be monitored in this way and any reduction in the area of the dune slacks should be identified using these methods. Of course, the large expense of repeated twice yearly airborne missions over the site may preclude its operational deployment, however, a mission every few years would be enough to establish a suitable baseline such that field based monitoring of the important areas of the site is more easily accomplished.

It should be remembered that the use of EO data is not intended to replace the need for land survey in mapping and monitoring habitats, but rather as a supplement to such actions. As such, the identification of the proportions of each land cover type as shown here is a useful addition to the site managers' knowledge. It is unlikely that a similar number of land based surveys would have arrived at such similar figures for the different land cover types as the assorted methods of data and information fusion have done here. They would certainly have involved the consumption of a substantial amount of resources. Despite the high cost of the data acquisition, and the need for some ground truth data, the level of effort required to repeat the analysis included here is relatively low compared to even a single exhaustive land based survey.

We do not suggest here that the use of EO data made the identification of the small area of embryo slack possible, as it was well known before. What is important are the relative proportions of the other land cover types and their distribution; they are much harder to derive from land based survey. Previous land cover maps of the site derived from land based survey were only very broad categories and showed a very homogenous distribution of the areas of each type. It is clear that coastal dune systems are not homogenous (inspect Figs. 3 and 4) and high resolution (spatially and spectrally) EO approaches may go some way to mapping and monitoring habitats within them.

Differences between the ability of the ANN and maximum likelihood (ML) classification methods to classify categorical data are shown clearly in Table 6. The first two rows show the synthetic data set method of data fusion classified using the ML and ANN methods respectively. The ML row has the lowest classification accuracy of any of the methods carried out. ANN classification accuracy of the categorical data set was amongst the higher values of all the remaining methods. This result is expected as the ML method is highly unsuitable for classifying categorical data; however, it does show the utility of ANNs for classifying just this sort of data. A site managed for conservation for any length of time is likely to have categorical information relating to land cover in the form of the expert knowledge of the site

Table 5. Class total areas (three sources) (%).

Class	May	August	ATM	CASI
Young slack	16.47	21.24	23.93	22.60
Mature slack	3.87	3.15	6.56	2.32
Rank grassland	5.37	4.22	3.19	5.32
Pteridium	1.12	4.75	3.35	1.60
Young grassland	14.08	9.28	10.28	12.78
Species rich slack	1.38	7.29	4.17	2.90
Salt marsh	1.29	1.71	1.30	1.09
Phragmites	1.82	0.66	1.38	0.82
Fen meadow	5.95	2.61	6.86	4.01
Embryo slack	0.21	0.61	0.35	0.36
Closed grassland	28.02	28.87	25.39	28.81
Calluna	0.21	0.10	0.21	0.06
Calamagrostis	7.63	2.18	3.87	5.30
Bare sand	2.27	2.69	2.49	2.32
Open water	3.61	3.67	3.59	3.66
Scrub	5.70	6.97	6.11	6.05

Table 6. Summary of classification results.

Classifier	Accuracy (%)	Khat
Data Fusion (ML)	83.10	0.6959
Data Fusion (ANN)	96.60	0.9444
Majority	97.67	0.9581
Plurality	99.01	0.9642
OWA (Mean)	98.37	0.9708
OWA (Median)	98.29	0.9693
OWA (Max)	97.87	0.9617
Fuzzy ATM	96.83	0.9430
Fuzzy CASI	92.16	0.8590
Fuzzy May	97.11	0.9481
Fuzzy August	91.97	0.8559
Fuzzy ATM/Elev	97.83	0.9610
Fuzzy CASI/Elev	94.27	0.8972
Fuzzy May/Elev	97.56	0.9561
Fuzzy August/Elev	94.41	0.8996

managers. ANNs provide one means of integrating this information into an accurate and repeatable land cover classification method.

Finally, from an inspection of the results of the fuzzy integral fusion experiments, a very promising result is shown. It can be seen from Table 6 that the two source fusions had accuracies of between 92% (Khat 0.8590) and 97% (Khat 0.9481). When the elevation data were added to the sources to be combined, data whose individual classification accuracy was very low (61.61%, which is little better than random class allocation), these accuracies increased to between 94 and 98% (Khat of 0.8972 and 0.9561 respectively). Whilst we have commented before on the unreliability of the classifier accuracy assessment, this does suggest that the method successfully utilisized the elevation data and an improved classification was the result.

Conclusions

The aim of this study was an investigation of the suitability of data and information fusion methods, as applied to EO data, for the mapping and monitoring of a coastal dune system of high conservation status. It is hoped that the study has shown novel approaches that overcome some of the difficulties in traditional EO approaches to the problem. Geographical Information Systems (GIS) have had much application in the fusion of both data and information, however, nearly all of their data models for dealing with the problem assume each source counts equally in the final assessment. It is hoped that this study may go some way to make others question if this assumption is a valid one.

The success in mapping a highly heterogeneous site, such as Kenfig NNR, using EO data has been shown. In comparison to existing habitat maps of the site, the new maps are much more detailed and should provide site managers with a useful level of detail in a one-off mapping exercise. The potential for the use of these methods in long-term surveillance or monitoring, particularly with reference to the reasons for notification of cSAC, NNR and SSSI status is promising. The stable proportion of land cover classes illustrates this for each of the classification methods.

In terms of the effort involved in such an enterprise, the returns have been good. EO data acquisition, whilst not cheap, is not prohibitively expensive when compared to the cost of obtaining a land cover map of comparable complexity from more traditional, land survey based, methods.

Regarding the fusion techniques, we have seen how the various methods may be used to integrate data and information from various sources. Many of these are simple techniques, readily applied using commonly available GIS software. The increase in accuracy gained by using the more complex fuzzy integral approach is certainly worthwhile however it is not so easily achieved using existing software.

Acknowledgements. The authors wish to thank NERC for providing funding during this research, the staff of the Kenfig NNR reserve centre, particularly Rob Davies, for much assistance in the field over a number of years, and the Environment Agency of the UK for providing the LIDAR data. UCL is also thanked for its continuing support during the writing of this paper.

References

Anon. 1949. *National Parks and Access to the Countryside Act*. HMSO, London, UK.

Anon. 1981. *Wildlife and the Countryside Act*. HMSO, London, UK.

Anon. 1990. *Handbook for Phase 1 habitat survey – a technique for environmental audit*. JNCC, Peterborough, UK.

Cho, S.-B. & Kim, J.H. 1995. Combining multiple neural networks by fuzzy integral for robust classification. *IEEE Trans. Syst. Man Cybern.* 25: 380-384.

Congalton, R.G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* 37: 35-46.

Congalton, R.G., Oderwald, R.G & Mean, R.A. 1983. Assessing Landsat classification accuracy using discrete multivariate statistical techniques. *Photogramm. Engin. Remote Sens.* 49: 1671-1678.

Foody, G.M. 1999. The continuum of classification fuzziness in thematic mapping. *Photogramm. Engin. Remote Sens.* 65: 443-451.

Foody, G.M., Campbell, N.A., Trodd, N.M. & Wood, T.F. 1992. Derivation and applications of probabilistic measures of class membership from the maximum likelihood classification. *Photogramm. Engin. Remote Sens.* 58: 1335-1341.

Grabisch, M. 1996. The application of fuzzy integrals in multicriteria decision making. *Eur. J. Operat. Res.* 89: 445-456.

Lee, H.-M. & Lee-Kwang, H. 1994. Information aggregating networks based on extended Sugeno's fuzzy integral. In: Furuhashi, T. (ed.) *IEEE/Nagoya University Wisepersons Workshop*, pp. 56-66. Springer, Heidelberg, DE.

McGwire, K.C. 1992. Analyst variability in labelling unsupervised classifications. *Photogramm. Engin. Remote Sens.* 58: 1673-1677.

Rosenfield, G.H. & Fitzpatrick-Lins, K. 1986. A coefficient of agreement as a measure of thematic classification accuracy. *Photogramm. Engin. Remote Sens.* 52: 223-227.

Rumelhart, D.E., Hinton, G.E. & Williams, R.J. 1986. Learning internal representations by error propagation. In: Rumelhart, D.E. & McClelland, J.L. (eds.) *Parallel distributed processing: Explorations of the microstructure of cognition*, Ch. 8, pp. 319-362. MIT Press, Cambridge, MA, US.

Stehman, S.V. 2000. Practical implications of design-based sampling inference for thematic map accuracy assessment. *Remote Sens. Environ.* 72: 35-45.

Sugeno, M. 1974. *Theory of fuzzy integrals and its applications*. Ph.D. Thesis, Tokyo Institute of Technology, Tokyo, JP.

Sugeno, M. 1977. Fuzzy measures and fuzzy integrals: A survey. In: *Fuzzy automata and decision processes*, pp. 71-78. North Holland Publishing, Amsterdam, NL.

Tahani, H. & Keller, J.M. 1990. Information fusion in computer vision using the fuzzy integral. *IEEE Trans. Syst. Man Cybern.* 20: 733-741.

Thackrah, G.N. 2001. *Habitat mapping and biodiversity monitoring using Earth Observation and GIS*. Ph.D. Thesis, University of Wales, Swansea, UK.

Thackrah, G.N., Barnsley, M.J. & Pan, P. 1999. Merging land cover classifications from artificial neural networks using the Sugeno fuzzy integral. In: Barnsley, M.J. (ed.) *RSS99, From data to information,* pp. 71-78. RSS, Nottingham, UK.

Thackrah, G.N., Barnsley, M.J. & Pan, P. 2000. Evaluating the accuracy of merged land cover classifications made using a fuzzy integral. In: Sincak, P. (ed.) *Quo vadis computational intelligence?*, Ch. 4, pp. 186-191. Springer, Heidelberg, DE.

Wood, J.D. 1996. *The geomorphological characterisation of Digital Elevation Models*. Ph.D. Thesis, University of Leicester, UK.

Yager, R.R. 1988. On Ordered Weighted Averaging aggregation operators in multicriteria decision making. *IEEE Trans. Syst. Man Cybern.* 18: 183-190.

Yager, R.R. 1998. Including importances in OWA aggregations using fuzzy systems modelling. *IEEE Trans. Fuzzy Syst.* 6: 286-294.

Received 3 July 2003;
Revision received 19 January 2004;
Accepted 26 February 2004.
Co-ordinating Editor: F. Taveira Pinto