A fuzzy classification of sub-urban land cover from remotely sensed imagery

J. ZHANG
School of Remote Sensing and Geo-Informatics, Wuhan Technical University of Surveying and Mapping, 39 Luoyu Road, Wuhan 430070, People's Republic of China; e-mail: jxz@hp01.wtusm.edu.cn

G. M. FOODY
Department of Geography, University of Southampton, Highfield, Southampton, SO17 1BJ, England, UK; e-mail: g.m.foody@soton.ac.uk

(Received 21 October 1996; in final form 16 October 1997)

Abstract. Fuzzy methods in remote sensing have received growing interest for their particular value in situations where the geographical phenomena are inherently fuzzy. A fuzzy approach is investigated for the classification of sub-urban land cover from remote sensing imagery and the evaluation of classification accuracy. Under the fuzzy strategy, fuzziness, intrinsic to both remotely sensed data and ground data, is accommodated and usefully explored. For comparative purposes, hard and fuzzy classifications were produced and tested using hard and fuzzy evaluation techniques. The results show that the fuzzy approach holds advantages over both conventional hard methods and partially fuzzy approaches, in which fuzziness in only the remotely sensed imagery is accommodated. It was found that Kappa coefficients were more than doubled when applying the fuzzy evaluation technique as opposed to the hard evaluation technique. Furthermore, the fuzzy approach paves the way towards an integrated handling of remotely sensed data and other spatial data.

1. Introduction

An important part of the provision of environmental information by remote sensing is achieved through image classification. The classification of remotely sensed imagery relies on the assumptions that the study area is composed of a number of unique, internally homogeneous classes and that classification analysis based on reflectance data and ancillary data can be used to identify these unique classes with the aid of ground data (Lillesand and Kiefer 1994, Townshend 1981). However, such assumptions are often not valid in areas with significant fuzziness. Fuzziness often occurs due to the presence of mixed pixels (mixels), which are not completely occupied by a single, homogeneous category. Mixels occur because the pixel size may not be fine enough to capture detail on the ground necessary for specific applications (Campbell 1987). They may also occur where the ground properties, such as vegetation and soil types, vary continuously (Wood and Foody 1993).

Conventional classification methods such as maximum likelihood classification are often incapable of performing satisfactorily in the presence of mixels. With a maximum likelihood classification, the classified output for each pixel comprises only the code of the class with which it has the highest probability of membership. This kind of classification technique is termed 'hard'. Intermediate data about
membership strength or similarity, usually obtained through computationally intensive procedures, are generally not provided to end users, although they may be very informative. For a mixed pixel the single class allocation provided by a hard classification must, to some extent, also be erroneous. Furthermore, as a mixed pixel displays a composite spectral response it may not even be allocated to one of its component classes. Mixed pixels therefore tend to degrade image classification accuracy.

Classification accuracy is usually evaluated with the aid of a confusion or error matrix. This shows the correspondence between the predicted and the actual classes of membership for a set of testing pixels. From the error matrix, it is possible to derive several measures of classification accuracy, such as the overall classification accuracy and the Kappa coefficient of agreement (Janssen and van der Wel 1994). Such measures may, however, only be usefully derived for hard classifications. Furthermore, these measures provide no information on accuracies below the individual class level, which may suggest spatially invariant accuracy within the individual units labelled as the corresponding classes. This is usually not true, especially when one recognizes the heterogeneity characteristic of many geographical phenomena (Bailey 1988). Furthermore, it is difficult to examine closely the source and nature of errors in classifications using hard methods.

To adapt to the fuzziness prevalent in natural phenomena, fuzzy approaches have been proposed (Wang 1990). These approaches are fuzzy in the sense that they allow for the multiple and partial class membership properties (note that the approaches are often based more on soft computing than fuzzy logic). Fuzzy approaches allow more information on the relative strengths of the class membership at pixel level to be made available to end users. Thus, for instance, both data producers and users can be made aware of the potential areas vulnerable to misclassification. The information on per-pixel class membership may also be used for post-processing of image classifications (e.g. Pathirana and Fisher 1991, Barnsley and Barr 1996). Moreover, areal data of different land covers at the sub-pixel level may be obtained, because sub-pixel component land cover proportions may be strongly related to their corresponding class membership values (Foody and Cox 1994, Foody 1996).

However, the fuzzy approaches generally used do not take into account fuzziness that may be characteristic of the ground data (Foody 1995, 1996). That is, these fuzzy approaches deal with fuzziness in the classification of remotely sensed data, but not in the ground data. Therefore, they are considered to be ‘partially fuzzy’ approaches. In order to evaluate the accuracy of a partially fuzzy classification, mixed pixels are often deliberately avoided. As the majority of pixels may be mixed, such a selective processing may result in an inaccurate estimation of classification accuracy (Foody 1995, 1997). Partially fuzzy approaches may therefore not be suitable when fuzziness is also a characteristic of the ground data, which is often the situation. Foody (1995) identified the need for using fuzzy ground data to evaluate the accuracy of a fuzzy classification in some circumstances. However, the fuzzy ground data used were sub-pixel component land cover proportions, which may arguably be only feasible when a pixel represents an area on the ground composed of a set of discrete polygons of classes. Since this may not be the case or the acquisition of sufficiently detailed ground data may be difficult, other methods for deriving fuzzy ground data are required.

This work was inspired by the capability and flexibility of fuzzy approaches to the classification of remotely sensed imagery and the evaluation of their classified
data products, as demonstrated by Foody (1995) and Gopal and Woodcock (1994). Section 2 will elaborate on the derivation of fuzzy classifications and fuzzy ground data. Different methods are described and discussed with respect to their relative usefulness. That is followed by a description of the methods for analysing the derived fuzzy classification and fuzzy ground data. Attention is then given to the empirical study carried out in the context of sub-urban land cover mapping, with reference to the data acquisition and the results obtained. Graphic and quantitative results will show the suitability of the fuzzy approach to geographical studies, in particular, of fuzzy phenomena, and what the fuzzy approach may hold for the future.

2. Deriving the fuzzy classification and fuzzy ground data

The concept of the fuzzy set is central to a fuzzy classification (Bezdek et al. 1984). A pixel can have partial and multiple memberships to the candidate classes. The relative strengths of class membership are termed as fuzzy membership values (FMVs) (Lowell 1994). FMVs generally range from 0 to 1, and are positively related to the strength of class membership of a pixel to the candidate classes. The FMVs of a pixel generally sum to 1 across all the candidate classes.

There are a number of ways of deriving the fuzzy membership values, depending on the specific classification techniques used. Assuming normality, maximum likelihood classification is based on an estimated pdf density function (pdf), as expressed by equation (1), for each of the reference classes under consideration:

\[
p(x|i) = \frac{1}{(2\pi)^{n/2} |V_i|^{1/2}} \exp\left(- \frac{1}{2} D_i^2 \right)
\]

where \( p(x|i) \) is the pdf for a pixel \( x \) as a member of class \( i \), \( n \) is the number of bands, \( |V_i| \) is the determinant of the variance and covariance matrix of class \( i \), and \( D_i^2 \) is the Mahalanobis distance from pixel \( x \) to the centroid of \( i \) (Fisher and Pathirana 1989, Foody et al. 1992).

A class label is usually derived for a pixel by assessing the \( a \) posteriori probabilities of membership on the assumption that it belongs to one of the pre-defined classes. The \( a \) posteriori probability of a pixel \( x \) belonging to class \( i \), \( L(i|x) \), may be determined from the equation:

\[
L(i|x) = \frac{p(i)p(x|i)}{\sum_{j=1}^{c} p(j)p(x|j)}
\]

where \( c \) is the total number of classes and \( p(i) \) the \( a \) priori probability of class \( i \). The probabilistic measure as calculated by equation (2) in combination with equation (1) is as attractive as a FMV, because, given that each pixel belongs to one of the classes, the \( a \) posteriori probabilities of a pixel sum to 1 across all classes.

Another type of probabilistic measure may be derived from referring spectral distance values such as the Mahalanobis distance to the \( \chi^2 \) distribution with \( n \) degrees of freedom; the value \( n \) is the number of spectral bands comprising the reflectance data. This kind of probability is understood as the proportion of pixels a distance further than \( D_i^2 \) from the mean of the class \( i \) belonging to the class \( i \) and indicates the typicality of class membership (Campbell 1984, Foody et al. 1992).

Artificial neural networks (ANNs) are also attractive for use in the classification of remotely sensed imagery and have gained increasing popularity in remote sensing, particularly as they do not rely on distribution assumptions and are able to integrate
ancillary data acquired at a low level of measurement precision (Foody 1997). When a neural network (NN) is used for classification, the strength of class membership can be measured by the activation level of the network output units (Foody 1996). FMVs can also be derived from a range of other classifiers including, for example, the so-called non-parametric classification approach, such as that reported by Skidmore and Turner (1988) and the fuzzy $c$-means algorithm (Bezdek et al. 1984). Finally, fuzzy classification can be performed in both supervised and unsupervised modes. For instance, FMVs can be derived from the fuzzy $c$-means clustering algorithm in either mode (Bezdek et al. 1984).

To assess the accuracy of a classification of remotely sensed data, ground data, typically derived from photogrammetry, field surveying or existing maps, are used. In such ground data, the spatial variations evident in reality are commonly obscured. For example, the extraction of land cover information from rigorously orientated stereo models from aerial photographic pairs will generally result in a map with discrete polygons of classes, with the boundary fuzziness and interior heterogeneity filtered out. The ground data set formed will thus be hard, with one class associated with each unit mapped.

One method of deriving fuzzy ground data is to ‘put back’ the heterogeneities within each mapping unit. It is then possible to use the proportions of different classes within a polygon or other mapping unit, such as a pixel’s equivalent area on the ground, as probabilities (Foody 1995). Furthermore, to allow gradual variations within mapping units to be mapped, interpolation methods may be used. One can label the inner parts of individual polygons with a probability of 100% or 0 to the class associated with that polygon. This probability will decrease when moving towards the boundaries (at the boundary the probabilities may be 0.5/0.5 for a simple two-class situation), until it reaches 0% or 0.0 in the locus of the centroids of the opposite polygons. The changing pattern of class probabilities along a transect may be modelled by some function (Wang and Hall 1996). This function may well be based on the distances between the points with known class probabilities and the points whose probabilities with respect to the candidate classes are to be interpolated. For example, Lowell (1994) suggested that fuzzy maps can be generated by first delineating relatively well defined boundaries or features and then inferring those fuzzier features based on some distance-based functions.

It is, however, worth noting that, for a discrete class such as a lake, the probability will vary not continuously but abruptly as one moves from one side of the boundary to the other, and will be relatively stable within the polygon of the class. In this case, distance-based interpolations may not be appropriate. Even for continuous classes, it is unlikely that the probability of finding a particular class at a point would be merely a simple function of the distances between the points of known class probabilities in a local neighbourhood and the point to be evaluated. It is likely that the spatial distribution of points of known probabilities in relation to the point to be evaluated would have equally important, if not dominant, effects (Bierkens and Burrough 1993a). Thus, it seems that the simple interpolation method discussed above cannot be suitably applied to derive fuzzy ground data due to a lack of both theoretical and empirical evidence. Other methods need to be considered.

Geostatistical methods have been developed to deal with categorical (or qualitative) attribute data as well as continuous (or quantitative) attribute data. Geostatistics typically uses semi-variograms to quantify spatial correlations and to guide spatial interpolations. Indicator kriging estimates the conditional (posterior) probability
distribution without making assumptions about the form of the prior distribution functions. Its application for estimating the probability of a class being located at a point is briefly described below.

Suppose \( c \) mutually exclusive classes can be found over a region and \( s \) points of known probabilities (termed as observations or classified samples) are available, as shown in figure 1, where solid rectangles represent observations. Each observation is classified as a member of one of the possible classes. For each class \( i \) \((i=1, \ldots, c)\) under consideration, the \( s \) observations are transformed into binary data (i.e. an observation is transferred to 1·0 if it is classified as class \( i \), 0·0 otherwise). These binary data are denoted as \( z_i(x_k), k=1, 2, \ldots, s \), for a particular class \( i \).

Indicator kriging may be used to estimate the probabilities of finding individual classes \( i \) \((i=1, \ldots, c)\) at a point \( x_0 \) using:

\[
\hat{z}_i(x_0) = \sum_{k=1}^{s} \lambda_k z_i(x_k)
\]

where \( z_i(x_k) \) represents a binary variable at an observation \( x_k \) \((k=1, 2, \ldots, s)\) as described above and \( \lambda_k \) is the weight associated with the observation \( x_k \). These weights (i.e. \( \lambda_k \) values) are calculated using a suitable semi-variogram model (Bierkens and Burrough 1993 a, b). For each point, the probability of each class occurring may therefore be derived, forming a fuzzy ground data set.

3. Classification accuracy assessment

Once the fuzzy classification and fuzzy ground data are derived, it is possible to evaluate the classification accuracy and undertake other analyses which make use of the wealth of information provided by the fuzzy classification approach.

Before further discussion, it is necessary to denote the vectors \( \mathbf{F} \) and \( \mathbf{P} \) vectors:

\[
\mathbf{F}(x) = (f_1(x), \ldots, f_c(x))
\]

\[
\mathbf{P}(x) = (p_1(x), \ldots, p_c(x))
\]

where \( f_j(x) \) and \( p_j(x) \) \((j=1, \ldots, c)\) are pixel \( x \)’s fuzzy membership values (derived from the remotely sensed data) and its corresponding ground element’s probability

Figure 1. A diagram showing the process of estimating the occurrence of classes at a point \( x_0 \) in the neighbourhood within the search radius \( r \), given that \( n \) observations (as shown by solid rectangles) are available.
scores (derived from the ground data) to the classes under consideration (ordered from class 1 to class \(c\)).

To generate a conventional maximum likelihood classification, the \(\mathbf{F}(x)\) vectors of all pixels are subjected to a ‘maximization’ process, by which each pixel is labelled as belonging to the class having the maximum value. In other words, pixel \(x\) is to be classified into class \(j\) on the condition expressed by:

\[
f_j = \text{maximum} (f_1, \ldots, f_c) \quad j = 1, \ldots, c
\]

(5)

The accuracy of this classification may be assessed using conventional measures such as the percentage correct allocation or Kappa coefficient of agreement calculated from a classification confusion matrix. The value of the maximum FMV can also be used to construct a companion band of the imagery to depict the spatial variability of confidence a user can place in the classified data at the pixel level.

When evaluating a classification using conventional confusion matrix based methods, the pixels are assumed to be pure. This is often not the case. Sometimes, pixels that can be considered to be pure pixels or pixels dominated by a single class need to be identified and selected in order to locate samples for assessing the accuracy of a particular classification. This selection process can be done via a ‘slicing’ operation, by which the maximum FMV (denoted by \(f_{\text{max}}\)) from vector \(\mathbf{F}(x)\) of a pixel is examined with reference to a pre-determined threshold \(\tau\). Specifically, this processing is performed such that a pixel \(x\) is selected if the value of \(f_{\text{max}}\) is not less than \(\tau\). Clearly, the purity of the pixels selected will rise the higher the threshold, but this will be associated with a decrease in the number of selected pixels. In this context the term pure pixels is therefore used only relatively, and indicates pixels dominated by a particular class. Fuzzy ground data, as denoted by \(\mathbf{P}(x)\) vectors, can be similarly processed in a maximization and a slicing operation.

The relative magnitude of the FMV vector of a pixel \(\mathbf{F}(x)\) may also be analysed. For example, the \(f_j(x)\) values \((j = 1, 2, \ldots, c)\) of a pixel \(x\) can be arranged descendingly. That is, actually sorting out the most likely down to the least likely classes of membership for each pixel. The sorted classes may be denoted by a vector \(\mathbf{O}\) \((c_1, c_2, \ldots, c_c)\). Similar processing can be applied to the corresponding ground data set for pixel \(x\). The resulting sorted sequence is denoted by vector \(\mathbf{O}\) \((C_1, C_2, \ldots, C_c)\). A ‘soft’ comparison may then be performed by comparing the most likely, the second most likely, down to the least likely classes as indicated by the vectors \(\mathbf{O}\) and \(\mathbf{O}\). That is by examining if \(c_a = C_b, a, b = 1, 2, \ldots, t\), where \(t\) specifies the tolerance set for a soft comparison (no more than \(c\)). The greater the tolerance, the more likely a match is to be obtained, and hence the larger the error. When \(t\) is set to 1 and if \(c_1 = C_1\), it is said that a perfect match is achieved because the most likely classes are found to be the same for the test data and the ground data. This is what happens when an agreement is reached during a hard classification. When the tolerance is set to the second most likely class, a soft comparison is performed based on the set of conditions as expressed below:

\[
c_1 = C_1 \text{ or } c_1 = C_2 \text{ or } c_2 = C_1 \text{ or } c_2 = C_2
\]

(6)

While the accuracy of a hard classification can be evaluated by means of an error matrix, which can be used to calculate measures such as the overall classification accuracy and Kappa coefficient, other measures have to be used to indicate the accuracy of a classification when the classified data or ground data are fuzzy. Gopal and Woodcock (1994) suggested the use of fuzzy measures to derive a range
of indicators of classification performance when the ground data are fuzzy. The methods, however, do not allow comparisons of different classifications. Entropy may be used to indicate the accuracy of a fuzzy classification, as entropy describes the variations in class membership probabilities associated with each pixel. This is, however, only suitable when ground data are hard. When both classified data and ground data are fuzzy, a more appropriate index of accuracy may be based on cross-entropy (Foody 1995). Other possible indices of classification accuracy may be based on correlation analysis and distance measures, including the Euclidean distance between the representations of the land cover in the image classification and ground data (Foody 1996, Foody and Arora 1996).

It is anticipated that the possibilities which fuzzy approaches will open up are significant, and will undoubtedly be promising for a full analysis of remotely sensed imagery and integral handling of remotely sensed data and GIS spatial data, as uncertainties and fuzziness intrinsic to GIS spatial databases have received growing interest from GIS communities (e.g. Goodchild 1989, Veregin 1995). To illustrate the potentials and advantages of a fuzzy approach, an empirical study is presented below.

4. An empirical study

4.1. The study area and data acquisition

An area of \(\sim 2\text{km}^2\), located within the city of Edinburgh, around Blackford Hill (figure 2) was selected. A sub-urban area was chosen for its richness in geographical diversity and thus its appropriateness for fuzzy classification. There is a variety of urban thematic and topographic features: a wooded valley, residential, commercial and academic buildings, road networks, footpaths, recreational areas, a small lake, worked allotments, hills and flat ground. The residential areas are compact. The roads, the pavements, the roofs, the walls and the hedges exist in complex spatial arrangements. The dispersed individual trees or groups of trees blend into neighbouring land cover types. Shrubs dominate the west end of the Blackford Hill area and grassland cover tends to exist in those areas not covered by concrete, bare ground, tall trees or shrubs.

SPOT HRV and Landsat TM data of the site were used and for illustrative purposes an extract of the SPOT HRV data is shown in figure 3. Ground data were derived based on the 1:24 000 scale aerial photographs (in natural colour) (figure 2). The Landsat TM data and 1:24 000 scale aerial photographs were both acquired in the summer of 1988, and can be safely assumed to be free of significant temporal changes in land cover. The SPOT HRV image was acquired in the summer of 1985, but no major changes in land cover were observed and so the aerial photographs were used to derive ground data for comparison against both the SPOT HRV data and the Landsat TM data.

For the purpose of this study, the USGS land use and land cover classification system for use with remote sensing data was used, with the following five classes appropriate to the scene defined: grass (grassland, including parkland), built (built-up land, including barren land), wood (wooded land, no distinction was made between deciduous and coniferous woodland), shrub (shrubland, including open wooded land), and water (water bodies, including lakes and water works).

4.2. Deriving the fuzzy classification and fuzzy ground data

To derive fuzzy classified remotely sensed data, a fuzzy c-means clustering algorithm, programmed in FORTRAN 77 on VAX/VMS, was used. The algorithm is a
Figure 2. An enlarged portion of a 1:24000 scale aerial photograph showing the study area (in black and white). Note the scanned data used in the research had a pixel size of approximately 4 m on the ground.

Figure 3. The SPOT HRV sub-scene used (all the three XS bands are used, copied in black and white).
non-hierarchical clustering technique which iteratively minimizes a least-squares error term and has been used widely in remote sensing investigations (e.g. Cannon et al. 1986). This program was based on the fuzzy c-means clustering algorithm published by Bezdek et al. (1984) and, due to the unsatisfactory results from the unsupervised mode, this algorithm was applied in a supervised mode. In the analyses, the algorithm's weighting exponent $m$ (Bezdek et al. 1984) was set at 2.5 and bands 1, 2 and 3 were used for the SPOT HRV image, and bands 3, 4 and 5 were used for the Landsat TM image (Zhang 1996). For this, training data were acquired by identifying, for each class, blocks of representative pixels from the imagery. For SPOT HRV data, the numbers of training pixels selected were 633, 1369, 332, 616 and 544 for grass, built, wood, shrub and water classes respectively. For Landsat TM data, the numbers of training pixels selected were 72 for grass, 109 for built, 39 for wood, 47 for shrub and 16 for water. The program was used to calculate the FMVs for each pixel in each of the five classes using class statistics derived from the training data. The resulting fuzzy classified data were attached as extra bands (one for each class) to their corresponding imagery.

To generate fuzzy ground data, two methods were tested. Firstly, sub-pixel land cover proportions, as relevant to SPOT HRV data (10m) and Landsat TM data (30m), were derived from the aerial photographs. For this, photogrammetric digitizing of land cover from the aerial photographs was carried out based on the reconstituted stereo model on an analytical plotter. This process resulted in a polygonal land cover ground data set. These data were then rasterized at fine cell size (1 m) and the fine grid data were then aggregated in accordance with the SPOT HRV and Landsat TM pixel sizes independently. Data for the proportions of sub-pixel component land cover were calculated on a pixel-by-pixel basis with respect to the five land cover classes. To facilitate subsequent analysis, the fuzzy ground data derived from sub-pixel component land cover proportions were also stored as a five-band image.

Secondly, indicator kriging was used to spatially interpolate class membership (Deutsch and Journel 1992). For this, a set of classified samples was identified from screen-displayed photogrammetric data. These sample points were carefully selected to ensure each could be considered as a pure point and thus has full membership (100%) of the named class with zero membership values to the other classes. They were then transformed to a grid coordinate system with grid cell size equal to $2.5 \times 2.5 \text{m}^2$ and the semi-variograms were calculated. The kriging procedure was eventually run with the output grids cell sizes equal to SPOT HRV and Landsat TM data pixel sizes. Again, the fuzzy ground data derived from indicator kriging were stored as a five-band image.

4.3. Classification evaluation

Hard classifications of the remotely sensed and ground data set were derived by using the maximization operation. This resulted in the production of conventional data sets used, or assumed, in remote sensing studies. In addition, for the purpose of a later slicing operation, the maximum FMVs were stored as extra bands of their corresponding imagery of fuzzy classifications. To enable a conventional evaluation of classification accuracy, both the fuzzy image classification and fuzzy ground data were hardened and the overall percentage correct allocation and Kappa coefficient of agreement derived. Results for both SPOT HRV data and Landsat TM data are listed in table 1. It was apparent that the classification accuracies were low, with misclassified cases distributed across the test site (figure 4).
Table 1. Summary of the classification accuracies using different approaches to compare the image classification and ground data.

<table>
<thead>
<tr>
<th>Types of comparisons</th>
<th>Number of pixels</th>
<th>Overall classification accuracies (%)</th>
<th>Kappa coefficients (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Hard comparisons between hardened fuzzy classified data and hardened fuzzy ground data derived from indicator kriging</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPOT HRV data</td>
<td>7113</td>
<td>49.9</td>
<td>34.8</td>
</tr>
<tr>
<td>Landsat TM data</td>
<td>777</td>
<td>50.2</td>
<td>39.0</td>
</tr>
<tr>
<td>2. Selective hard comparisons between pure pixels and pure ground data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPOT HRV data</td>
<td>1837</td>
<td>74.7</td>
<td>61.6</td>
</tr>
<tr>
<td>Landsat TM data</td>
<td>213</td>
<td>72.8</td>
<td>59.4</td>
</tr>
<tr>
<td>3. Soft comparisons between fuzzy classified data and fuzzy ground data derived from indicator kriging</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPOT HRV data</td>
<td>7113</td>
<td>92.8</td>
<td>90.0</td>
</tr>
<tr>
<td>Landsat TM data</td>
<td>777</td>
<td>93.2</td>
<td>90.1</td>
</tr>
</tbody>
</table>

Using a fuzzy classification of the remotely sensed data and the fuzzy ground data, it is possible to extend the analysis. Firstly, the accuracy of a fuzzy classification can be assessed by selectively comparing the pure pixels with pure ground data remembering that these so-called pure pixels are not pure in absolute terms but are the purest pixels in the data sets. Technically, this is done by slicing and then comparing the pixels, which are considered pure above a specific threshold, as described in §3. In this research, the maximum FMVs of Landsat TM data were sliced at a threshold of 41%, with their corresponding ground probabilities sliced at 72%. The two thresholds for SPOT HRV data were set at 54% and 77% respectively. These thresholds were set at the average values of the maximum FMVs stored from the analyses. Not surprisingly, the classification accuracies derived were higher than those derived from the conventional hard approach applied to all pixels (table 1) as only ‘pure’ pixels were used. This highlights the need to accommodate for mixed pixels when evaluating the accuracy of a classification using techniques which assume pure pixels.

Furthermore, based on the fuzzy classified data and fuzzy ground data, a soft comparison down to the second most likely classes level was undertaken by using the condition (6) described in §3. This resulted in substantial increases in the classification accuracies. In particular, the classification accuracies derived from the soft comparison method are almost double those obtained by the conventional hard comparisons, while Kappa coefficients are more than double. This large increase in the agreements explains, in another perspective, that a considerable amount of misclassification occurs due to the fuzziness and complexity which exists in the real world. The soft comparison method described here accounts for fuzziness by taking account of both the most likely and the second most likely class memberships. In fact, in situations in which the underlying geographical entity is fuzzy, the second most likely class memberships may be very important (Zhang 1996). Finally, the
spatial patterns of misclassification after soft comparisons may be shown as maps. The result for Landsat TM data is shown in figure 5 as an example. Not surprisingly, the majority of misclassifications occur near boundaries, where transition zones and geographical complexities are encountered, though the relationship is not simple.

4.4. Quantitative analysis

For a fuzzy classification, it is possible to calculate entropy measures on a pixel basis to describe the partitioning of membership between the classes. Results are
Figure 5. The distribution of misclassifications after a soft comparison between fuzzy classified Landsat TM data and fuzzy ground data: (a) classified data; (b) misclassified data.

presented as histograms in figure 6 for both SPOT HRV and Landsat TM data. The means of the entropy measures for SPOT HRV and Landsat TM data were 1.56 and 1.64 respectively, implying that pixels are significantly mixed. It is therefore difficult to interpret the accuracy of classification based on entropy measures, as pixels with a low or high entropy measure may accurately represent the ground situation (Foody 1995).

Cross-entropy allows the closeness of a fuzzy classification to a fuzzy ground data set to be measured, and so indicates accuracy (Foody 1995, 1996). The closer
the classification to the ground data, the lower the cross-entropy and the higher the classification accuracy. This measure may also be used when hard ground data are used. Figure 7 indicates the distributions of cross-entropy values based on fuzzy classified SPOT HRV and Landsat TM data and their corresponding hard ground data, for which the means of the cross-entropy measures are 4.31 and 3.81 respectively. Figures 8 and 9 show the cross-entropy values derived when fuzzy ground data derived using the two different methods as described in §4.3 are used. When using sub-pixel component proportions as fuzzy ground data, the means of cross-entropy measures for SPOT HRV and Landsat TM data are 3.67 and 2.49, which are much smaller than those based on hard ground data. The means of the cross-entropy measures are further reduced to 2.67 for SPOT HRV data and 1.89 for Landsat TM data when the fuzzy ground data are derived from indicator kriging (figure 9).

So far, it has been seen that accommodating for fuzziness in the production and evaluation of image classification is more appropriate than the use of conventional hard (in which no accommodation for fuzziness is made) and partially fuzzy approaches (in which fuzziness is accounted for in the image classification but not the ground data). A better performance in the example studied is also achieved through the use of indicator kriging to derive fuzzy ground data rather than using sub-pixel component proportions as fuzzy ground data. Furthermore, the relative suitability of using indicator kriging as opposed to the sub-pixel proportions of the
component classes within mapping units to generate fuzzy ground data can be evaluated by the strengths of correlations between the FMVs derived from a fuzzy classification and the probability data of corresponding fuzzy ground data. The correlation analysis was performed for each class and consistently higher correlation coefficients were obtained when indicator kriging was used to derive the fuzzy ground data sets for both SPOT HRV and Landsat TM data (table 2), although the magnitudes of the differences were variable, suggesting that using probabilistic data generated by geostatistical approaches is the more suitable of the two approaches, at least for the purposes of this study. One of the reasons why the use of sub-pixel component land cover proportions may be less suitable than interpolated data from kriging is that sub-pixel land cover component proportion data are used on the assumption that the component land covers can be accurately represented as mixtures of discrete polygons of the classes. This assumption is unlikely to be true of areas with significant fuzziness.

5. Conclusions

This paper has presented a case study, in which hard and fuzzy classifications were performed and tested using hard and fuzzy evaluation techniques. The results show that a fuzzy classification strategy may enable a suitable and effective
Figure 8. Histograms of cross-entropy measures calculated based on: (a) fuzzy classified SPOT HRV data and fuzzy ground data derived from sub-pixel component proportions; (b) fuzzy classified Landsat TM data and fuzzy ground data derived from sub-pixel component proportions.

Table 2. Correlation coefficients ($r$) for the relationships between fuzzy classified data and fuzzy ground data as derived using: (i) sub-pixel component cover proportions and (ii) indicator kriging. ($^m$ denotes statistically significant at the 95% level of confidence, while $^\times$ denotes otherwise.)

<table>
<thead>
<tr>
<th>Classes</th>
<th>SPOT HRV data</th>
<th>Landsat TM data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7113 pixels</td>
<td>777 pixels</td>
</tr>
<tr>
<td>Grass</td>
<td>$0.43_m$</td>
<td>$0.49_m$</td>
</tr>
<tr>
<td></td>
<td>$-0.04 \times$</td>
<td>$0.06 \times$</td>
</tr>
<tr>
<td>Built</td>
<td>$0.58_m$</td>
<td>$0.70_m$</td>
</tr>
<tr>
<td>Wood</td>
<td>$0.52_m$</td>
<td>$0.52_m$</td>
</tr>
<tr>
<td>Shrub</td>
<td>$0.25_m$</td>
<td>$0.30_m$</td>
</tr>
<tr>
<td>Water</td>
<td>$0.72_m$</td>
<td>$0.68_m$</td>
</tr>
<tr>
<td></td>
<td>$0.28_m$</td>
<td>$0.34_m$</td>
</tr>
<tr>
<td></td>
<td>$0.28_m$</td>
<td>$0.34_m$</td>
</tr>
<tr>
<td></td>
<td>$0.10_m$</td>
<td>$0.45_m$</td>
</tr>
</tbody>
</table>

classification of remotely sensed imagery depicting inherently fuzzy phenomena and their evaluations, and allows for locational and quantitative examinations of the misclassification in classified data. To use the fuzzy approaches detailed ground data are required. While this inevitably increases the cost and complexity of an investi-
J. Zhang and G. M. Foody

Figure 9. Histograms of cross-entropy measures calculated based on: (a) fuzzy classified SPOT HRV data and indicator-kriged fuzzy ground data; (b) fuzzy classified Landsat TM data and indicator-kriged fuzzy ground data.

The benefits accrued, particularly in terms of improved representation and accuracy, must be considered. It is also worth stressing that detailed ground data may be required for conventional classification analyses, as ground data for a pixel are supposed to describe the class membership properties of the area on the ground represented by the pixel; conventional classification approaches have the same requirement for ground data detail as the fuzzy approaches (Foody 1999). Furthermore, it is relatively easy to accommodate for fuzziness in the various stages of the classification process (Foody and Arora 1996). For example, a number of approaches to derive and evaluate fuzzy classifications have been reviewed and these may be integrated into standard classification approaches if desired.

This paper concludes by reinforcing the importance of accommodating fuzziness in the last two stages of the supervised classification approach. This may be extended by accounting for fuzziness in the training stage, resulting in a fully fuzzy classification (Foody 1997). More importantly, this paper confirms that the information richness offered by a fuzzy classification strategy can be usefully explored in order to achieve improved and integrated handling of remotely sensed data with other spatial data (Zhang 1996), as uncertainties in spatial databases have been identified as being of significant importance on the research agenda (Bailey 1988, Goodchild 1989, Wang and Hall 1996).
Acknowledgments

Advice and contributions from Dr R. Kirby and Dr N. Stuart (University of Edinburgh) and Professor M. Goodchild (NCGIA) are gratefully acknowledged. The majority of the work reported in this paper was undertaken while JZ was at the University of Edinburgh and GMF at the University of Salford.

References


Fuzzy classification of sub-urban land cover


