Using mixed objects in the training of object-based image classifications

Hugo Costa, Giles M. Foody, and Doreen S. Boyd

School of Geography, University of Nottingham, Nottingham NG7 2RD, UK

Abstract

Image classification for thematic mapping is a very common application in remote sensing, which is sometimes realized through object-based image analysis. In these analyses, it is common for some of the objects to be mixed in their class composition and thus violate the commonly made assumption of object purity that is implicit in a conventional object-based image analysis. Mixed objects can be a problem throughout a classification analysis, but are particularly challenging in the training stage as they can result in degraded training statistics and act to reduce mapping accuracy. In this paper the potential of using mixed objects in training object-based image classifications is evaluated. Remotely sensed data were submitted to a series of segmentation analyses from which a range of under- to over-segmented outputs were intentionally produced. Training objects were then selected from the segmentation outputs, resulting in training data sets that varied in terms of size (i.e. number of objects) and proportion of mixed objects. These training data sets were then used with an artificial neural network and a generalized linear model, which can accommodate objects of mixed composition, to produce a series of land cover maps. The use of training statistics estimated based on both pure and mixed objects often increased classification accuracy by around 25% when compared with accuracies

obtained from the use of only pure objects in training. So rather than the mixed objects being a problem, they can be an asset in classification and facilitate land cover mapping from remote sensing. It is, therefore, desirable to recognize the nature of the objects and possibly accommodate mixed objects directly in training. The results obtained here may also have implications for the common practice of seeking an optimal segmentation output, and also act to challenge the widespread view that object-based classification is superior to pixel-based classification.

Keywords: OBIA; mixed pixels; under-segmentation; over-segmentation; scale parameter

1. Introduction

Information on the Earth's surface such as land cover and related environmental processes is of great importance for a plethora of applications, for example for decision-making on issues related to agriculture and food security (Fritz et al., 2013; Gardi et al., 2015), monitoring the distribution of species (Martin et al., 2013; Tuanmu and Jetz, 2014), and modelling of the Earth's climate (Luyssaert et al., 2014; Mahmood et al., 2014). For this reason, thematic mapping through a classification analysis is a very common application of remote sensing. Over the years substantial progress has been made in remote sensing-based mapping, and today there are many ways through which a classification analysis can be conducted (Lu and Weng, 2007; Momeni et al., 2016).

A key decision needed during a classification analysis is on which basic spatial unit to use. Considerable use of the pixel, the basic spatial unit of a digital image, and per-pixel based classification has been common for decades. However, grouping spatially connected pixels into objects by means of an image segmentation analysis, and using the object as the basic spatial unit has become very popular in recent years (Blaschke et al., 2014). The objects obtained from an image segmentation analysis may, in principle, form a more suitable spatial unit than the pixel for land cover mapping as they should relate to natural spatial units (e.g. fields) unlike pixels which are artificial units defined more by the sensing system than the properties of the ground. The use of objects comprising multiple pixels can also aid the calculation of potentially useful discriminatory variables such as descriptors of image texture (Laliberte and Rango, 2009).

There are, however, fundamental issues and assumptions of classification that often appear to be ignored or incompletely addressed in object-based image analyses. For example, it is common for the objects produced from the segmentation analysis to be routinely and unquestioningly used as if pure in the classification (e.g. Goodin et al., 2015; Shimabukuro et al., 2015; Uddin et al., 2015). However, this is often not the case, mainly for two reasons. First, remotely sensed data inevitably comprise a proportion of mixed pixels whatever the spatial resolution used (Addink et al., 2012; Cracknell, 1998; Fisher, 1997), which cannot be accommodated by traditional image segmentation. For example, Wu (2009) found that 40-50% of the pixels of an urban area represented in multispectral IKONOS data (4 m resolution) were mixed. Second, image segmentation often produces mixed objects as a result of under-segmentation error. This type of error corresponds to situations such as the failure of the image segmentation analysis to define a border splitting two land cover classes, thereby generating a single object containing more than one class (Clinton et al., 2010).

Failure to satisfy the assumptions of classification can greatly degrade the quality of the land cover map produced ultimately. In particular, the specific case of under-segmentation error (Gao et al., 2011; Hirata and Takahashi, 2011; Wang et al., 2004) is a problem throughout the classification process as mixed objects can degrade class training statistics, they cannot be appropriately allocated to a single class, and any such allocation must to some extent be erroneous (Heumann, 2011). Action is therefore needed to address the impact of these mixed units. That said, deviation from the assumptions of classification can, however, sometimes be made in each of the main stages of a classification analysis (e.g. Foody, 1999a). Specifically, impure units can be accounted for in training (Eastman and Laney, 2002; Foody, 1997; Hansen, 2012; Zhang and Foody, 2001), class allocation (Dronova et al., 2011; Foody, 1996; Wang, 1990), and testing a classification (Binaghi et al., 1999; Foody, 1995; Stehman et al., 2007). For example, van de Vlag and Stein (2007) generated objects based on remotely sensed data, classified them using fuzzy decision trees, and produced fuzzy error matrices in accuracy assessment. However, little research has been undertaken on the use of mixed units in training object-based image classifications.

Typically, the objects used in training are assumed to be pure (i.e. contain a single class), but a range of options are available if mixed objects are encountered. For example, the analyst could seek to simply ignore the problem, act to exclude the mixed cases, or adopt procedures that can accommodate the mixed nature of the units (Foody, 1999a, 1997). In object-based classification, the presence of mixed objects in training is sometimes addressed beforehand by deliberately favouring over-segmentation, that is, producing numerous small objects at the segmentation stage

(Boyden et al., 2013; Cánovas-García and Alonso-Sarría, 2015; Dronova et al., 2012; Van Coillie et al., 2008). However, this approach may be sub-optimal (Dorren, 2003; Gao et al., 2011; Hirata and Takahashi, 2011; Kim et al., 2009; Mishra and Crews, 2014) and is unlikely to remove all impure objects (Zhou et al., 2009; Zhou and Troy, 2008). Another solution sometimes adopted is the exclusion of mixed objects from the production of training statistics (Cai and Liu, 2013; Dean and Smith, 2003; Dronova et al., 2011; Güttler et al., 2016). In this way, the mixed units, which do not satisfy key assumptions of the analysis, are excluded so that the analysis can proceed with suitable data. Excluding mixed objects has, however, the consequence that the size of the training data sets will be reduced, and this could limit the quality of the resulting training statistics. This issue is particularly relevant in object-based classifications as the pool of potential training units is typically relatively small at the outset (Ma et al., 2015). Excluding mixed objects from the pool of selectable objects can exacerbate the challenge of finding a sufficient number of training objects (Mui et al., 2015; Wang et al., 2004).

Another issue to take into account while excluding mixed objects is the criteria according to which an object should be considered as mixed. It is unclear whether an object containing a very small fraction of pixels corresponding to a minority class should be excluded from training because there is the chance of those minority pixels having a negligible impact on the training statistics produced. For example, Cai and Liu (2013) excluded from training all objects whose dominant class occupied less than 90% of the objects’ area. The effect of issues such as threshold selection have not been investigated in detail (Li et al., 2016) and is most likely to be dependent on several factors, such as the remotely sensed data used and the land cover classes mixed (Dronova et al., 2011).

The assumptions of a classification also impact on the way training data sets should be used. For example, the training stage of a supervised classification should be designed in relation to the chosen classifier as different algorithms use the data differently. Specifically, with a standard statistical classifier, such as the maximum likelihood classification, it is important that each class is described appropriately which often requires a relatively large and representative training sample (Ediriwickrema and Khorram, 1997; Hagner and Reese, 2007; Paola and Schowengerdt, 1995; Richards and Kingsbury, 2014) while the use of a small sample of deliberately selected extreme and atypical samples may be more suited to non-parametric classifiers, such as a multilayer perceptron neural network, support vector machine, and classification tree (Foody, 1999b; Foody and Mathur, 2006; Hansen, 2012; Pal and Foody, 2012). Critically, the nature of the data used in training a classification should be acknowledged and addressed.

In this paper it is argued that it is not necessary, or even desirable, to exclude mixed objects from training an object-based image classification. In particular, it is possible to turn the apparent problem of mixed units into an asset, as with mixed pixels in per-pixel classification (Foody, 1997), recognizing that each individual mixed unit can be a source of training data on more than one class, and that mixed units can be used in training. Here, the potential of using mixed objects in training an object-based image classification is evaluated. A series of image segmentation analyses were undertaken from which training objects were selected, resulting in training data sets that varied in terms of size and proportion of mixed objects. The mixed objects generated at the segmentation stage and encountered at the training stage are included in the set of objects used to estimate training statistics, and the classification outputs produced by two classifiers are evaluated in relation to a conventional analysis using only pure objects. Thus, the work sets out to test the hypothesis that mixed objects may be used in the training of object-based image classifications to increase the accuracy with which land cover may be mapped from remotely sensed data.

2. Materials and methods

2.1 Study area and data sets

The analyses focused on a test site of approximately 45,000 ha in northern Portugal (Figure 1). The area corresponds to the downstream part of river Lima where the city of Viana do Castelo is settled. A diverse range of land cover types are present in the study area, and five land cover classes were defined: Artificial surfaces, Agricultural areas, Forest and semi-natural areas, Open spaces with little or no vegetation, and Wetlands and water bodies.

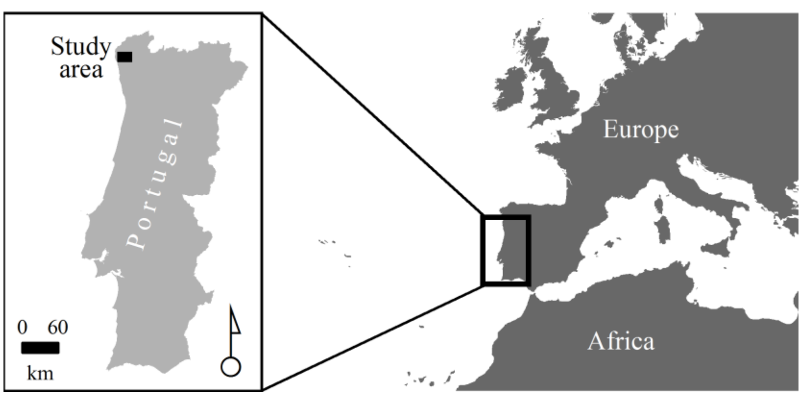

Figure 1. Location of the study area.

A Portuguese map, “Carta de Ocupação do Solo” of 2007 (COS2007), was used as reference data set (Figure 2a) in training and testing the object-based classifications. This map was produced by the Portuguese mapping agency (Direção-Geral do Território) through visual interpretation of aerial imagery and use of auxiliarydata such as field work and the national forest inventory. Land cover is represented according to a nomenclature similar to that used in this study in the third of a total of five hierarchical thematic levels used to map land cover with a minimum mapping unit of 1 ha (Caetano et al., 2010). As a guide to the thematic accuracy of the map, the overall accuracy is 96.82±1.01% at the 95% confidence level for the thematic detail used in this article, 5 classes, and the producer’s accuracy for each of the classes is >92%. This map provides the most accurate and detailed representation of the land cover that is available for the region and hence is suitable as reference data in the production and assessment of optimal image classifications in this study.

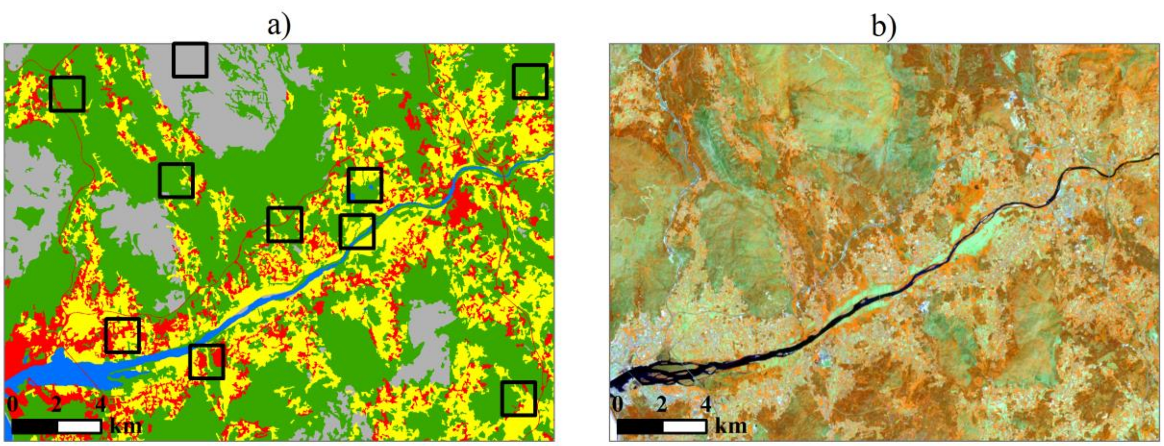

Figure 2. Data sets used: a) reference land cover map, COS2007, representing Artificial surfaces in red, Agricultural areas in yellow, Forest and semi natural areas in green, Open spaces with little or no vegetation in grey, and Wetlands and water bodies in blue; the square areas outlined in black are the training areas randomly located for the estimation of object-based training statistics. b) LISS-III image collected in Summer of 2006; RGB composition of data acquired in the near infrared, short wave infrared and red bands respectively.

Two images acquired during Spring and Summer of 2006 (Figure 2b) by the Linear Imaging Self Scanning Sensor (LISS-III) onboard IRS-P6 (also known as ResourceSat-1) were used. These two images are part of the IMAGE2006 European coverages provided by the European Space Agency (Müller et al., 2009). LISS-III is a multi-spectral camera operating in four spectral bands (green, red, near infrared, and short wave infrared) with a spatial resolution of 23 m in each. The two LISS-III images were orthorectified and resampled to 25 m spatial resolution using an SRTM-based digital elevation model (Müller et al., 2009). The four bands of the two images were stacked and thus formed an eight waveband data set.

2.2 Image segmentation

The LISS-III data were segmented using the multiresolution algorithm implemented in GeoDMA software (Körting et al., 2013), version 0.2.1, which is based on the popular algorithm of Baatz and Schäpe (2000). This is a region-based algorithm that uses spectral and shape properties of the objects being generated. Colour and shape parameters range within the interval 0-100 and are inversely proportional (i.e. colour=100-shape). In addition, the parameter shape depends on two other parameters, compactness and smoothness, also ranging within the interval 0-100 and inversely proportional. Finally, scale is a threshold of heterogeneity allowed within objects as regard colour, and shape properties. Essentially, the larger the scale parameter value the larger the heterogeneity allowed within objects, thus making larger and fewer objects (Körting et al., 2013).

The scale, colour, and shape parameter values are the most influential parameters (Luo et al., 2015) and were varied to obtain a series of different segmentations. First, five values were defined for the parameter colour, covering the entire range of possible values for this parameter: 1, 25, 50, 75, and 99. The importance of the spectral properties of the objects is positively related to the magnitude of the colour parameter. For simplicity, the parameter shape is not discussed hereafter as its value is automatically known for a given value of parameter colour. Then, for each of the five colour parameter values defined, the scale parameter was varied greatly as this is the most influential parameter (Luo et al., 2015). Specifically, eight values were defined from 10 to 80 in steps of 10. As a result, 40 segmentation outputs were obtained, ranging from over-segmented results mostly composed of small and possibly pure objects to under-segmented results mostly composed of large and possibly mixed objects. For the purposes of this paper an object was taken to be pure if more than 90% of its area was covered with a single class, similar to Cai and Liu (2013).

2.3 Training

A fraction of the study area was randomly selected for training purposes. Specifically, ten 225 ha square areas were selected randomly to provide training data. As a result, a total of 2250 ha, which is ~5% of the study area (Figure 2a), was available for training purposes. The training areas were intentionally defined as being small relative to the study area to simulate the limited availability of reference data that are typical of real-world applications. The objects of each of the segmentation outputs that intersected the training areas were selected for the production of training statistics. The objects generated via image segmentation are commonly used for estimating training statistics (e.g. Goodin et al., 2015). Thus, while the same geographical area was used in training each classification the set of objects used varied between the segmentation outputs. As a result, the training statistics varied between segmentation settings. In all cases, however, the representation of the land cover classes was constant and proportional to their abundance (Table 1) as the training areas defined were constant and randomly located.

Table 1. Proportion of area of the land cover classes mapped in the reference COS2007 land cover map in the training areas defined. The relative proportion of the classes is common to all training statistics estimated from the segmentation settings used.

Land cover class	Proportion of area (%)
Artificial surfaces	9.87
Agricultural areas	25.99
Forest and semi-natural areas	52.26
Open spaces with little or no vegetation	10.35
Wetlands and water bodies	1.53

The mean and standard deviation of the digital numbers of each object in the eight LISS-III spectral bands were used as training statistics, resulting in a total of 16 discriminating variables. The training objects were assigned reference class labels extracted from the COS2007 land cover map. The proportion of the area that each class occupied in a training object was estimated, ranging from 0.0 if the class was absent to the maximum value of 1.0 if the object was pure, with intermediate values for at least two classes if the object was of mixed class composition.

The remotely sensed data were classified using each of the segmentation outputs produced. Two scenarios were followed. First, following the traditional procedure of using only pure objects at the training stage. Specifically, each object intercepting the training areas was taken as pure and hence allowed to be a training object only for the class with which had the maximum membership based on the proportion of class area, which had to be superior to 90% (otherwise they were excluded from training). Second, all of the training objects, even if some were mixed, were used. The fractional coverage of the classes found in the objects was used as a measure of class membership, and training objects were allowed multiple and partial membership.

Because mixed objects were not excluded from training in the mixed training strategy, the size of the mixed training data sets was typically larger than when only pure objects were used. Since the size of the training set may impact the classification accuracy (Ma et al., 2015; Millard and Richardson, 2015) a series of analyses in which training set size was constant was also undertaken. For this additional analyses, reduced versions of the mixed training data sets were generated, with the size of the mixed data sets decreased to equal the size of the pure training data sets. The reduction of the size of the mixed training data sets was achieved by excluding randomly selected objects. Since the mixed training data sets may comprise both pure and mixed objects, this approach means that all objects, pure and mixed, had the same probability of being excluded. This allowed the size of the training data sets to be reduced without changing substantially the inherent ratio of pure to mixed objects. As the random exclusion of objects can result in numerous and different training data sets each of which with a potential different impact on the results, three reduced mixed training data sets were produced from each mixed training data set.

A series of classifications of the remotely sensed data using all 40 segmentation outputs was produced. With each segmentation output, classifications were produced that were trained using (i) pure training data sets, (ii) mixed training data sets, and (iii) the reduced (to same size as pure) training data sets.

2.4 Classification

In all analyses, multinomial regression models fitted by means of an artificial neural network with no hidden layer (Venables and Ripley, 2002) and a generalized linear model via penalized maximum likelihood (Friedman et al., 2010) were used for classification. These classifiers are available in the R programing language (R Core Team, 2016) from the packages ‘nnet’ and ‘glmnet’ respectively. Both classifiers allow fractional composition of the objects to be used in training in a manner similar to that explained in Foody (1997) for per-pixel classification.

The output of the classifiers is soft, indicating the probability of an object belonging to each class (Friedman et al., 2010; Venables and Ripley, 2002). However, traditional hard land cover maps were estimated by allocating each object the label of the class with which it had the greatest probability of membership. Although it may be beneficial to address the potential mixed nature of the objects at the class allocation stage, hard classification was adopted to confine the focus of the paper to the training stage. Each segmented image was thus used to produce hard land cover maps based on different training strategies: pure, mixed, and reduced mixed.

2.5 Accuracy assessment

The thematic accuracy of each classification produced was assessed. Confusion matrices comparing the land cover maps produced and the reference COS2007 land cover map were constructed through an operation of spatial intersection of the two data layers. Thus, instead of using a sample to estimate classification accuracy, the entire study area was used to assess the accuracy with which each of the 40 segmentation outputs produced was classified. Note, however, that the area associated with training (Figure 2a) was not used for accuracy assessment because that would artificially increase classification accuracy. Classification accuracy was expressed in terms of proportion of area correctly classified. Because the entire study area was used to estimate proportions of area correctly and incorrectly classified, the issues of selecting either pixels or objects as sampling units and producing estimates of accuracy which holds statistical uncertainty do not arise.

The accuracy of the 40 segmentation outputs generated was also assessed to provide a measure of under- and over-segmentation error, which is useful for analysis of the results. The method developed by Möller et al. (2013) and slightly refined by Costa et al. (2015) was used. This method belongs to a popular family of methods widely known as empirical discrepancy or supervised methods (Clinton et al., 2010; Zhang, 1996), and essentially compares the image segmentation output under evaluation to a reference data set (e.g. land cover map) to measure the geometric match between the objects that form them. Möller et al.'s (2013) method includes typical area-based and position-based metrics such as the ratio of overlapping area among generated and reference objects and the distance between the objects' centroid (Clinton et al., 2010; Whiteside et al., 2014) to detected and measure under- and over-segmentation error separately. The metrics are the basis for finding an optimal segmentation output that offsets the two types of error while informing on which type predominates when unbalanced, which is useful for this study. A summary of the segmentation accuracy is provided by metric M^g which measures the strength and type of segmentation error. Negative M^g values indicate that under-segmentation error dominates while positive M^g values represent the opposite case in which over-segmentation error dominates. Therefore, M^g~0 is considered indicative of optimal segmentation accuracy as both types of error are balanced (Möller et al., 2013). The reference data set used was the set of polygons of the COS2007 land cover map over the training areas defined. Thus, it was possible to determine whether the training data sets used were over-segmented, under-segmented, or balanced.

3. Results

The 40 segmentation outputs generated varied greatly in nature from over- to under-segmented, as expected, and two examples are shown in Figure 3 to highlight the different sets of objects obtained. The geometric accuracy of the objects that intersected the training areas, and hence used for training, was assessed, and the results are presented in Figure 4. Small values of the parameter scale produced over-segmented training objects (M^g>0) while large scale values yielded under-segmented outputs (M^g<0). For intermediate scale values, the type and magnitude of segmentation error became less evident. According to the Costa et al.'s (2015) adaptation of Möller et al.'s (2013) method, the scale value of 10, 30, 40, 50, and 70 were close to being optimal when the parameter colour was set at 1, 25, 50, 75, and 99, respectively, as M^g~0.

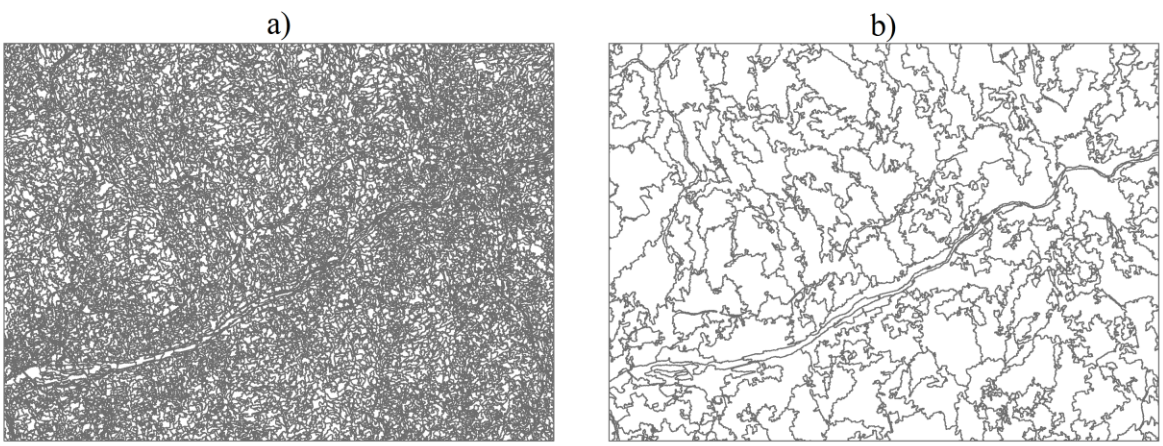

Figure 3. Segmentation results: a) Segmentation output produced with colour=75 and scale=10. b) Segmentation output produced with colour=75 and scale=80.

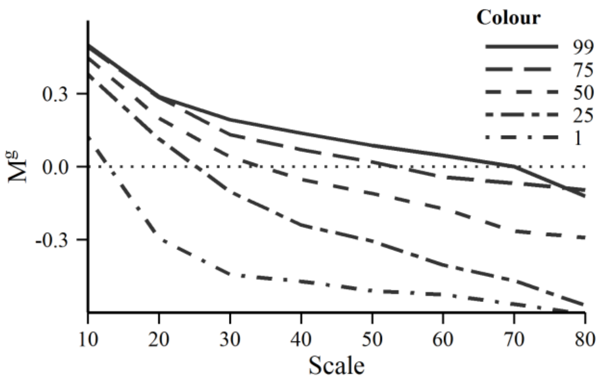

Figure 4. Segmentation accuracy based on Costa et al.'s (2015) adaptation of Möller et al.'s (2013) method. Dotted horizontal line corresponds to optimal accuracy.

The range of segmentation outputs generated resulted in training data sets of varying sizes (Figure 5). The number of training objects was large when over-segmentation was large (i.e. small values of parameter scale), and decreased as the level of under-segmentation increased. For example, when the parameter scale was set at 10 and 80, the training data sets generated comprised >430 and <100 training objects respectively. For a same value of parameter scale, larger training data sets were obtained when parameter colour was large. In all cases, some objects were mixed and hence the number of pure objects generated by a segmentation analysis was always less than the total number of objects generated. As such more objects were available for training when mixed rather than only pure objects were allowed. Specifically, 30-70% of the total number of the training objects was excluded when only pure objects were permitted in training. For example, the apparently near optimal segmentation output generated using colour=50 and scale=40 comprised 113 objects of which only 62 were pure (Figure 5). Thus, 51 objects (45% of the total) were of mixed class composition and hence excluded when only pure objects were allowed in training using an apparently near optimal segmentation. Mixed objects should, therefore, be expected to occur and even be common in object-based analyses.

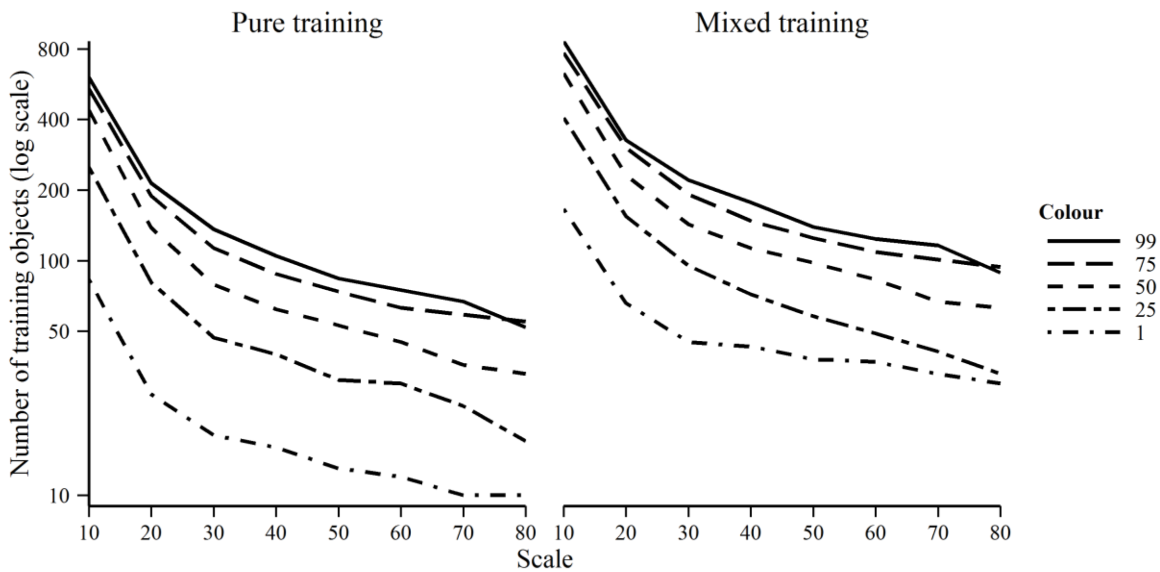

Figure 5. Size of training data sets.

The nature of the training data sets had a considerable impact on classification accuracy regardless of the classifier used. In general, the generalized linear model enabled classification to reach larger accuracy than the artificial neural network due to the regularization procedure of the former, but the results are consistent between them while excluding or allowing mixed objects in training. Critically, the magnitude of classification accuracy was consistently smaller for the classifications that excluded mixed objects in training relative to that which allowed mixed objects (Figure 6). For example, using the artificial neural network and segmentation output colour=50 and scale=40, the classification accuracy was 50.4 and 69.0% when the pure and mixed training data sets were used respectively (points A and B in Figure 6c; Figure 7a,b). The differences observed in terms of classification accuracy were so substantial that the smallest accuracy achieved with the use of the mixed training data sets (57.7%; point C in Figure 6b, using 33 training objects) was larger than the largest accuracy achieved with the use of only pure training data (54.7%; point D in Figure 6i, using 540 training objects) if the results obtained with colour=1 (Figure 6b and Figure 6f) are ignored. The minimum value of parameter colour was associated with somewhat atypical results. For example, the segmentation settings defined by colour=1 and scale=80 are associated with an increase in classification accuracy when trained with only pure objects, but the quality of the land cover maps is very small. Specifically, virtually the entire study area was classified as Forest and semi-natural areas class, and hence the classification accuracy tends to converge with the proportion of the study area covered by that class (~50%). The results obtained using colour=1 were caused mainly by the extremely small consideration of spectral information while generating the objects, and thereby relatively large values for the parameter colour are commonly used. The results obtained with this particular value of parameter colour are not referred or discussed hereafter for simplicity.

The reduced mixed training data sets also afforded larger classification accuracy than the pure training data sets. For example, the reduced mixed training data sets used to classify the segmentation output produced using colour=50 and scale=40 were enough for the accuracy of the artificial neural network to reach 61.6, 63.8, and 65.8% (points E in Figure 6c, Figure 7c,d), substantially more than with the use of pure training data (50.4%; point A in Figure 6c). There are a few cases in which the reduced mixed training data sets produced slightly larger classification accuracy than the full mixed training data sets, particularly when parameters colour and scale were large and small respectively (Figure 6e and Figure 5j). Thus, the difference in the accuracy achieved with the use of pure and mixed training sets is not simply an issue of training set size; mixed objects appear useful to produce valuable discriminatory information. The ability to increase classification accuracy through the use of mixed training objects should help the production of land cover maps that meet user needs and exceed appropriate target accuracy values (Foody, 2008). The key issue in relation to the hypothesis being tested in this article, however, is that the use of mixed objects can substantially increase classification accuracy relative to that achieved when only pure training objects are used. The difference in accuracy arising from the use of mixed and pure training sets varied with the specific parameter settings of the classifications, but was typically large. Specifically, difference in accuracy between classifications trained using pure and mixed training sets was up to 36% but typically in the order of 20% (Figure 6).

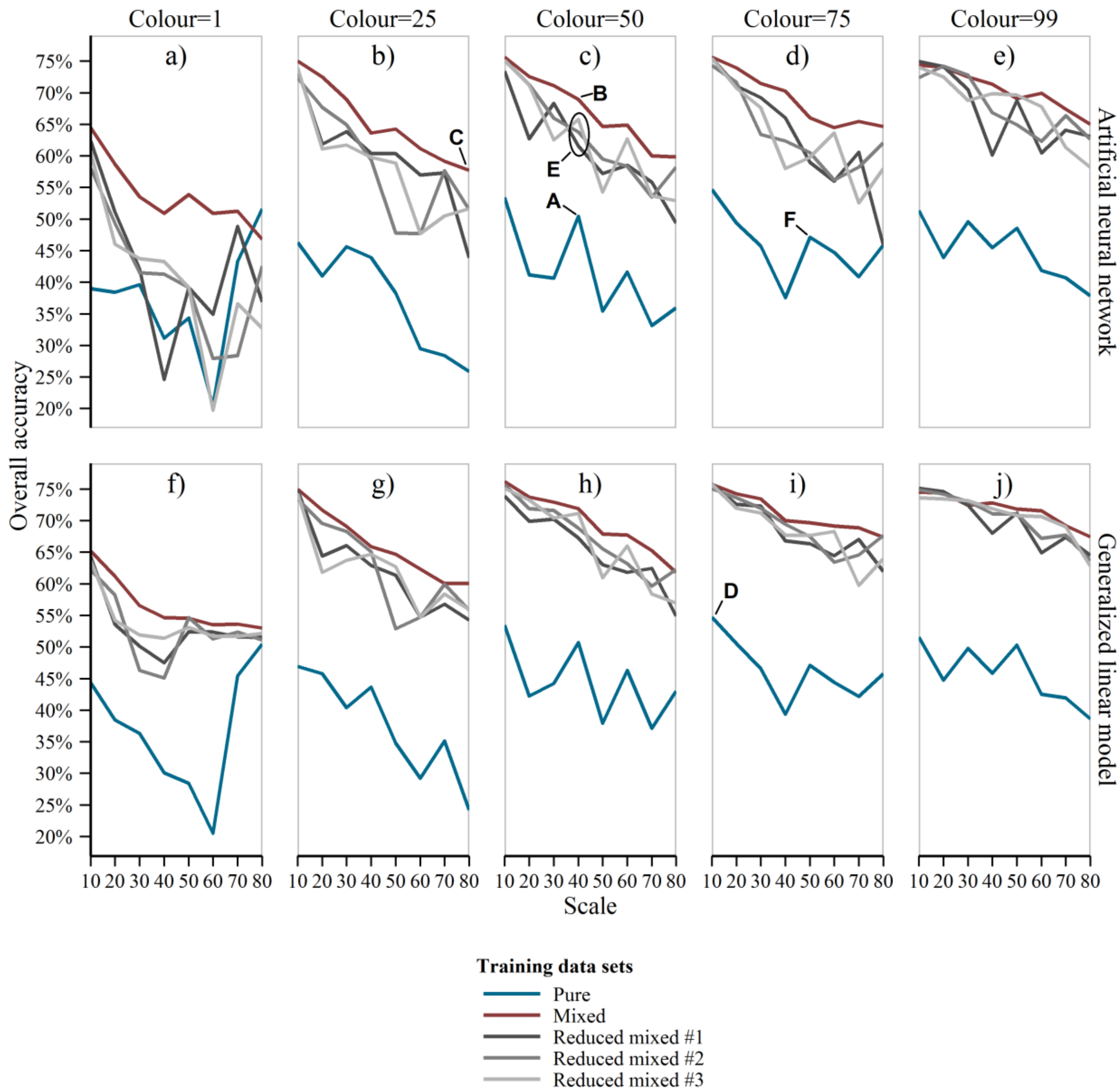

Figure 6. Classification accuracy as a function of the parameter scale with parameter colour set at 1, 25, 50, 75, and 99. The results obtained using the three reduced versions of the mixed training data sets are identified as #1, #2, and #3. Panel a) to e) and f) to j) refer to the artificial neural network and generalized linear model respectively.

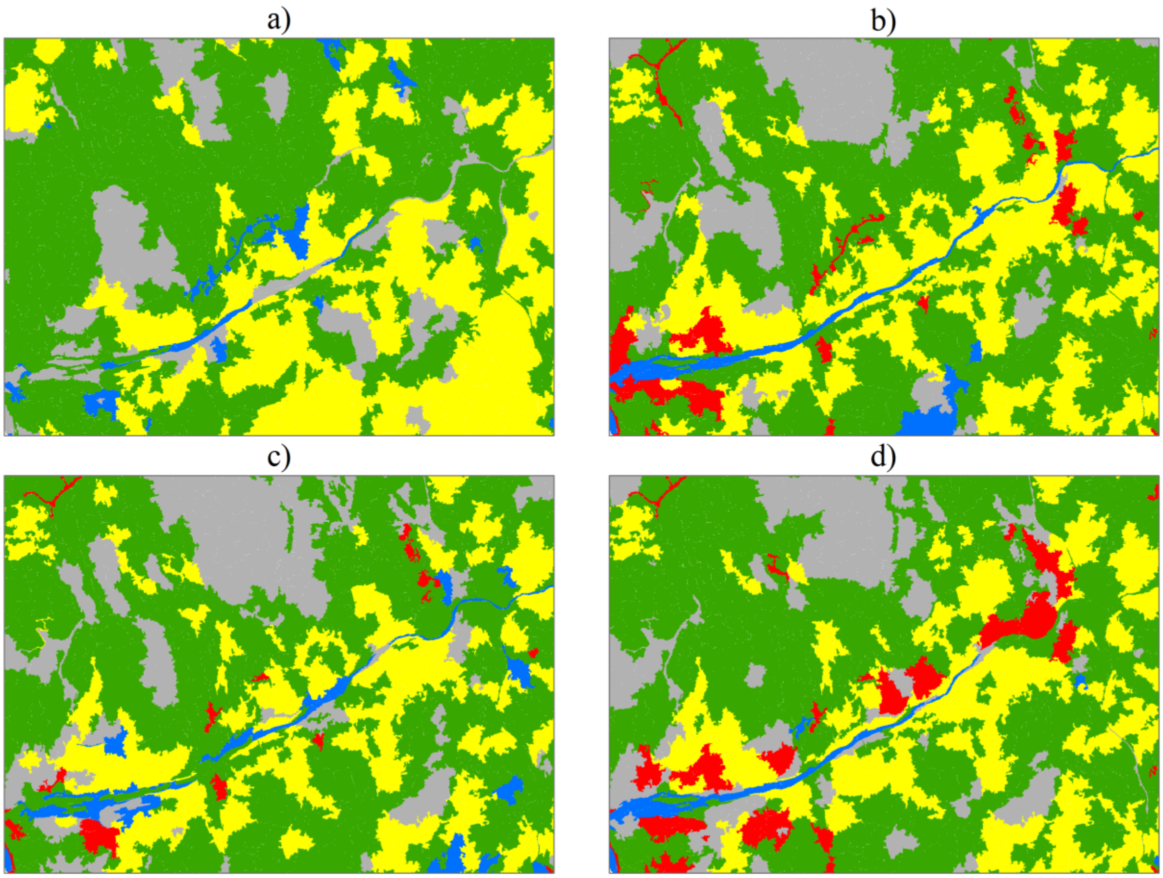

Figure 7. Land cover maps obtained using the segmentation output produced with colour=50 and scale=40, the artificial neural network, and different training strategies. a) Pure training (accuracy: 50.4%). b) Mixed training (accuracy: 69.0%). c) Mixed training with reduced samples #1 (accuracy: 61.5%). d) Mixed training with reduced samples #3 (accuracy: 65.8%). Colour legend as in Figure 2.

Beyond the evident difference between the classifications generated with pure and mixed, even if reduced, training data, there is a clear positive trend in classification accuracy with over-segmentation. The largest overall accuracies were reached using scale=10 regardless of the colour parameter values defined (25, 50, 75, or 99) and training strategy used (pure or mixed). However, there were differences in the way classification accuracy varied as a function of the parameter scale. Classification accuracy decreased continuously with increasing scale values when the entire mixed training data sets were used (smooth decreasing red lines in Figure 6); classification accuracy decreased variably when the reduced mixed training data sets were used (fluctuating decreasing grey lines in Figure 6), which is likely due to the random exclusion of specific training objects – objects with more or less impact on the training statistics could be excluded; finally, classification accuracy also tended to decrease as a function of the parameter scale when the pure training data sets were used, but marked variations are visible in Figure 6 (fluctuating blue lines). Overall accuracy sometimes peaked locally around the regions indicated as being close to balanced segmentation errors ($M^g \sim 0$), for example when the colour parameter was set at 50 and 75 and the artificial neural network was used (point A in Figure 6c and point F in Figure 6d). The local peaks around the regions indicated as being close to balanced segmentation errors agree with numerous studies reporting that segmentation results neither over- nor under-segmented are associated with land cover maps of larger thematic accuracy which have been produced via an image classification analysis trained with pure data (Dorren, 2003; Gao et al., 2011; Hirata and Takahashi, 2011; Kim et al., 2009; Kim and Warner, 2011; Mishra and Crews, 2014).

4. Discussion

Object-based classification of LISS-III data benefited from allowing training data sets to include mixed objects. An important advantage of mixed training is the opportunity to use relatively large training data sets since there is no need to exclude mixed objects. Furthermore, mixed objects allow efficiency as they give information on more than one class. It is well-known that the size of a training sample often positively influences classification accuracy (Ma et al., 2015;

Millard and Richardson, 2015). However, the size of the training data sets is not the only factor explaining the results obtained. When the parameter scale was set at small values, the pure training data sets were large, but the classification accuracy continued to be relatively small. On the contrary, mixed training afforded larger classification accuracy than pure training, even when the mixed training data sets were small. Indeed, the results obtained from the reduced mixed training data sets are closer to those obtained from the full mixed rather than the pure training data sets. Note that the difference in classification accuracy achieved with full and reduced mixed training data sets shrank with an increase in over-segmentation (Figure 6). This suggests that the size of the training data sets produced from extremely over-segmented outputs is not entirely needed, that is a smaller number of training objects may be sufficient to produce similar classification results as long as mixed objects are allowed in training.

Mixed objects provide useful discriminatory information, and this may be the main advantage of allowing mixed objects in training. Specifically, a representative sample of the objects generated is used, which includes objects of mixed in addition to pure class composition. Thus, classifiers can learn that the spectral properties of the objects may steam from the spectral signature of thematic classes as well as their mixture. Mixed units must convey information on more than one class and also will, in feature space, tend to lie between the classes involved. As a result of the latter the mixed units may be expected to lie close to the classification hyperplane that separates the classes. Mixed units, therefore, may have the potential to aid class separation, the central aim of a classification analysis. This potential can be exploited when the classifier can directly accommodate mixed responses in training and uses the training cases, rather than summary statistics, directly (Foody, 1999b; Foody and Mathur, 2006). Focus on separability rather than description of classes has been focus of innovative learning methods in recent years, such as active learning with mixed spectral responses (Samat et al., 2016). This may be especially relevant for classification as mixed objects may be common in an image segmentation output.

The use of mixed units is not without challenges. For example, due to intra-class spectral variation it is possible for units of slightly different thematic composition to have the exact same spectral response. However, the general trend is for mixed units to lie between the relevant classes in feature space. The exact position in feature space is a function of the mixing. A unit dominated by one class might be expected to lie relatively close to that class and still be distant from the other(s) involved while one of more equal mixing lies more centrally between the classes; exact details will depend on the specific classes and details (Foody, 2000; Hill et al., 2007; Lee and Lathrop, 2006; Zhu et al., 2013). Another difference between the pure and mixed training strategies relates to the fluctuating classification accuracy observed across the range of segmentation levels used. The accuracy of the classifications that used only pure training data was highly variable as a function of the parameter scale. These results were possibly caused by the imbalanced number of training objects per class (Table 1) and the relatively small size of the pure training data sets (Ma et al., 2015). Imbalanced training sets and the limited spectral resolution of the data used may have set limits to the achievable accuracy and larger accuracy could potentially be achieved with a larger, more balanced, training set and hyperspectral data (Carrão et al., 2008). On the contrary, by setting parameter colour at relatively large values benefited classification accuracy irrespective of the training strategy followed because the spectral content of the remotely sensed data gained importance for the generation of the objects.

The fluctuating classification accuracy associated with pure training sometimes showed a peak with segmentation settings which were near to optimal (e.g. point A in Figure 6c and point F in Figure 6d). This may suggest that larger accuracy of the image segmentation analysis offers larger classification accuracy and justifies the common practice for searching for optimal image segmentation results (Dorren, 2003; Gao et al., 2011; Hirata and Takahashi, 2011; Kim et al., 2009; Kim and Warner, 2011; Mishra and Crews, 2014). Indeed, since image segmentation and object-based classification have become popular methods for land cover mapping, the body of literature dedicated to methods for parameterization and accuracy assessment of image segmentation has grown (Clinton et al., 2010; Whiteside et al., 2014; Yang et al., 2015). Typically, these methods are focused on finding a segmentation result considered as being optimal in the sense that over- and under-segmentation error are minimal and balanced. However, a comprehensive analysis of the results shows that over-segmentation is associated with larger classification accuracy, and thus the assessment of segmentation accuracy is not necessarily informative for the prediction of an accuracy object-based classification (Figure 4 and Figure 6; Li et al., 2016; Ma et al., 2015). A number of studies (Belgiu and Drăguţ, 2014; Räsänen et al., 2013; Verbeeck et al., 2012) have observed that classification accuracy and segmentation accuracy, the latter at least as defined by empirical discrepancy methods, may not be positively related.

It was notable that the results showed a positive trend of classification accuracy with over-segmentation. In the situation of over-segmentation the size and spectral content of the objects generated become close to those of the pixels, and thus the results suggest that classification accuracy might possibly reach a maximum if per-pixel or near to per-pixel (Dronova et al., 2012; Ju et al., 2005) classification had been undertaken. Comparing object-based and per-pixel image classification has received much attention, and some publications have reported similar or larger accuracy of per-pixel classification as compared to object-based classification (Cai and Liu, 2013; Goodin et al., 2015; Robertson and King, 2011). However, the majority of the literature actually appears to hold the contrary view, that is that object-based classification achieves larger accuracy than per-pixel (Estoque et al., 2015; Goodin et al., 2015; Memarian et al., 2013; Whiteside et al., 2011). Apparently, the view that object-based classification is superior to per-pixel classification has become widespread and commonly unquestionable. The results presented above emphasize the need for more research in this respect. For example, typical comparisons between object-based and per-pixel classifications have relied on pure training, while the use of mixed training, which is also beneficial for per-pixel classification (Eastman and Laney, 2002; Foody, 1997; Hansen, 2012; Zhang and Foody, 2001), should be considered in comparative studies. Furthermore, it should be taken into account that the suitability of per-pixel and object-based classifications may depend on scale issues related to the land cover patterns on the ground, and the spatial resolution of the remotely sensed data and classification nomenclature used. For example, the appropriateness of a segmentation level varies as a function of the land cover classes mapped (Castilla et al., 2014; Kim and Warner, 2011; Laliberte and Rango, 2009) not least because the definition of categorical classes is a scale dependent issue (Ju et al., 2005).

Finally, this study used an artificial neural network and generalized linear model able to accommodate mixed objects which is a fundamental aspect to take into account if training using impure units is to be undertaken. Alternative parametric classifiers, such as the maximum likelihood classification provided in commonly used software packages, may be less appropriate as relatively large and representative samples formed by pure training units are needed to describe the classes; unless the training statistics are rectified. For example, the component parts of a mixed object can be unmixed and used to estimate the signal of the object as it would be if pure (Foody and Arora, 1996). The use of mixed objects for training a classification is, therefore, possible for a range of classifiers and may facilitate land cover mapping from remote sensing. The mixed nature of the spatial units used may also be addressed at the class allocation stage as partial and multi class membership are estimated, which can then be assessed based, for example, on the fuzzy confusion matrix (Binaghi et al., 1999; Stehman et al., 2007). This paper focused on the training stage, but a fully fuzzy classification approach may be implemented for thematic mapping (Foody, 1997; Zhang and Foody, 2001).

5. Conclusions

An implicit assumption made typically in object-based image classification is that the objects are pure. This is often not the case, and in this paper it was shown that mixed objects can be accommodated into the training stage of an object-based image classification. Contrary to common practice, it may be, therefore, not necessary to remove mixed objects from the training stage of a supervised image classification. Rather, an analysis of the effects of allowing mixed objects in training should be considered, which also affords an increase in the size of the training data set, and may contribute to an increase in classification accuracy. For example, by using mixed objects in this study often the overall accuracy increased by around 25% relative to that achieved using pure objects only. Furthermore, the results suggest that it may not be necessary to follow common practice and seek an optimal segmentation output. Specifically, deliberate over-segmentation may be a suitable strategy for generating objects for optimal training.

Acknowledgements

The authors are grateful to Markus Möller from the University of Halle (Saale) for providing R code, the LISS-III data used was provided by the European Space Agency, and the research was supported by the PhD Studentship SFRH/BD/77031/2011 from the "Fundação para a Ciência e Tecnologia" (FCT), funded by the "Programa Operacional Potencial Humano" (POPH) and the European Social Fund. The quality of the original manuscript was improved based on valuable suggestions of anonymous reviewers and the Associate Editor.

References

Addink, E.A., Van Coillie, F.M.B., De Jong, S.M., 2012. Introduction to the GEOBIA 2010 special issue: From pixels to geographic objects in remote sensing image analysis. Int. J. Appl. Earth Obs. Geoinf. 15, 1–6. doi:10.1016/j.jag.2011.12.001

Baatz, M., Schäpe, A., 2000. Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation, in: Strobl, J., Blaschke, T., Griesebner, G. (Eds.), Angewandte Geographische Informationsverarbeitung XII. Beiträge Zum AGIT-Symposium

Salzburg 2000. Herbert Wichmann Verlag, Heidelberg, pp. 12–23.

Belgiu, M., Drăguţ, L., 2014. Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery. ISPRS J. Photogramm. Remote Sens. 96, 67–75. doi:10.1016/j.isprsjprs.2014.07.002

Binaghi, E., Brivio, P.A., Ghezzi, P., Rampini, A., 1999. A fuzzy set-based accuracy assessment of soft classification. Pattern Recognit. Lett. 20, 935–948. doi:10.1016/S0167-8655(99)00061-6

Blaschke, T., Hay, G.J., Kelly, M., Lang, S., Hofmann, P., Addink, E.A., Queiroz Feitosa, R., van der Meer, F., van der Werff, H., van Coillie, F., Tiede, D., 2014. Geographic Object-Based Image Analysis – Towards a new paradigm. ISPRS J. Photogramm. Remote Sens. 87, 180–191. doi:10.1016/j.isprsjprs.2013.09.014

Boyden, J., Joyce, K.E., Boggs, G., Wurm, P., 2013. Object-based mapping of native vegetation and para grass (*Urochloa mutica*) on a monsoonal wetland of Kakadu NP using a Landsat 5 TM Dry-season time series. J. Spat. Sci. 58, 53–77. doi:10.1080/14498596.2012.759086

Caetano, M., Nunes, A., Dinis, J., Pereira, M. d. C., Marrecas, P., Nunes, V., 2010. Carta de uso e ocupação do solo de Portugal Continental para 2007: Memória descritiva. Instituto Geográfico Português, Lisbon.

Cai, S., Liu, D., 2013. A comparison of object-based and contextual pixel-based classifications using high and medium spatial resolution images. Remote Sens. Lett. 4, 998–1007. doi:10.1080/2150704X.2013.828180

Cánovas-García, F., Alonso-Sarría, F., 2015. A local approach to optimize the scale parameter in multiresolution segmentation for multispectral imagery. Geocarto Int. 30, 937–961. doi:10.1080/10106049.2015.1004131

Carrão, H., Goncalves, P., Caetano, M., 2008. Contribution of multispectral and multitemporal information from MODIS images to land cover classification. Remote Sens. Environ. 112, 986–997. doi:10.1016/j.rse.2007.07.002

Castilla, G., Hernando, A., Zhang, C., McDermid, G.J., 2014. The impact of object size on the thematic accuracy of landcover maps. Int. J. Remote Sens. 35, 1029–1037. doi:10.1080/01431161.2013.875630

Clinton, N., Holt, A., Scarborough, J., Yan, L., Gong, P., 2010. Accuracy assessment measures for object-based image segmentation goodness. Photogramm. Eng. Remote Sensing 76, 289–299.

Costa, H., Foody, G.M., Boyd, D.S., 2015. Integrating user needs on misclassification error sensitivity into image segmentation quality assessment. Photogramm. Eng. Remote Sensing 81, 451–459. doi:10.14358/PERS.81.6.451

Cracknell, A.P., 1998. Synergy in remote sensing-what's in a pixel? Int. J. Remote Sens. 19, 2025–2047.

Dean, A.M., Smith, G.M., 2003. An evaluation of per-parcel land cover mapping using maximum likelihood class probabilities. Int. J. Remote Sens. 24, 2905–2920. doi:10.1080/01431160210155910

Dorren, L., 2003. Improved Landsat-based forest mapping in steep mountainous terrain using object-based classification. For. Ecol. Manage. 183, 31–46. doi:10.1016/S0378-1127(03)00113-0

Dronova, I., Gong, P., Clinton, N.E., Wang, L., Fu, W., Qi, S., Liu, Y., 2012. Landscape analysis of wetland plant functional types: The effects of image segmentation scale, vegetation classes and classification methods. Remote Sens. Environ. 127, 357–369. doi:10.1016/j.rse.2012.09.018

Dronova, I., Gong, P., Wang, L., 2011. Object-based analysis and change detection of major wetland cover types and their classification uncertainty during the low water period at Poyang Lake, China. Remote Sens. Environ. 115, 3220–3236. doi:10.1016/j.rse.2011.07.006

Eastman, J.R., Laney, R.M., 2002. Bayesian soft classification for sub-pixel analysis: a critical evaluation. Photogramm. Eng. Remote Sens. 68, 1149–1154.

Ediriwickrema, J., Khorram, S., 1997. Hierarchical maximum-likelihood classification for improved accuracies. IEEE Trans. Geosci. Remote Sens. 35, 810–816. doi:10.1109/36.602523

Estoque, R.C., Murayama, Y., Akiyama, C.M., 2015. Pixel-based and object-based classifications using high- and medium-spatial-resolution imageries in the urban and suburban landscapes. Geocarto Int. 30, 1113–1129. doi:10.1080/10106049.2015.1027291

Fisher, P., 1997. The pixel: A snare and a delusion. Int. J. Remote Sens. 18, 679–685. doi:10.1080/014311697219015

Foody, G.M., 2008. Harshness in image classification accuracy assessment. Int. J. Remote Sens. 29, 3137–3158. doi:10.1080/01431160701442120

Foody, G.M., 2000. Estimation of sub-pixel land cover composition in the presence of untrained classes. Comput. Geosci. 26, 469–478. doi:10.1016/S0098-3004(99)00125-9

Foody, G.M., 1999a. The continuum of classification fuzziness in thematic mapping. Photogramm. Eng. Remote Sensing 65, 443–451.

Foody, G.M., 1999b. The significance of border training patterns in classification by a feedforward neural network using back propagation learning. Int. J. Remote Sens. 20, 3549–3562. doi:10.1080/014311699211192

Foody, G.M., 1997. Fully fuzzy supervised classification of land cover from remotely sensed imagery with an artificial neural network. Neural Comput. Appl. 5, 238–247. doi:10.1007/BF01424229

Foody, G.M., 1996. Approaches for the production and evaluation of fuzzy land cover classifications from remotely-sensed data. Int. J. Remote Sens. 17, 1317–1340. doi:10.1080/01431169608948706

Foody, G.M., 1995. Cross-entropy for the evaluation of the accuracy of a fuzzy land cover classification with fuzzy ground data. ISPRS J. Photogramm. Remote Sens. 50, 2–12. doi:10.1016/0924-2716(95)90116-V

Foody, G.M., Arora, M.K., 1996. Incorporating mixed pixels in the training, allocation and testing stages of supervised classifications. Pattern Recognit. Lett. 17, 1389–1398. doi:10.1016/S0167-8655(96)00095-5

Foody, G.M., Mathur, A., 2006. The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. Remote Sens. Environ. 103, 179–189. doi:10.1016/j.rse.2006.04.001

Friedman, J., Hastie, T., Tibshirani, R., 2010. Regularization Paths for Generalized Linear Models via Coordinate Descent. J. Stat. Softw. 33, 1–22. doi:10.1359/JBMR.0301229

Fritz, S., See, L., You, L., Justice, C., Becker-Reshef, I., Bydekerke, L., Cumani, R., Defourny, P., Erb, K., Foley, J., Gilliams, S., Gong, P., Hansen, M., Hertel, T., Herold, M., Herrero, M., Kayitakire, F., Latham, J., Leo, O., McCallum, I., Obersteiner, M., Ramankutty, N., Rocha, J., Tang, H., Thornton, P., Vancutsem, C., van der Velde, M., Wood, S., Woodcock, C., 2013. The need for improved maps of global cropland. Eos, Trans. Am. Geophys. Union 94, 31–32. doi:10.1002/2013EO030006

Gao, Y., Mas, J.F., Kerle, N., Pacheco, J.A.N., 2011. Optimal region growing segmentation and its effect on classification accuracy. Int. J. Remote Sens. 32, 3747–3763. doi:10.1080/01431161003777189

Gardi, C., Panagos, P., Van Liedekerke, M., Bosco, C., De Brogniez, D., 2015. Land take and food security: assessment of land take on the agricultural production in Europe. J. Environ. Plan. Manag. 58, 898–912. doi:10.1080/09640568.2014.899490

Goodin, D.G., Anibas, K.L., Bezymennyi, M., 2015. Mapping land cover and land use from object-based classification: An example from a complex agricultural landscape. Int. J. Remote Sens. 36, 4702–4723. doi:10.1080/01431161.2015.1088674

Güttler, F.N., Ienco, D., Poncelet, P., Teisseire, M., 2016. Combining transductive and active learning to improve object-based classification of remote sensing images. Remote Sens. Lett. 7, 358–367. doi:10.1080/2150704X.2016.1142678

Hagner, O., Reese, H., 2007. A method for calibrated maximum likelihood classification of forest types. Remote Sens. Environ. 110, 438–444. doi:10.1016/j.rse.2006.08.017

Hansen, M., 2012. Classification trees and mixed pixel training data, in: Remote Sensing of Land Use and Land Cover, Remote Sensing Applications Series. CRC Press, pp. 127–136. doi:doi:10.1201/b11964-12

Heumann, B.W., 2011. An Object-Based Classification of Mangroves Using a Hybrid Decision Tree—Support Vector Machine Approach. Remote Sens. 3, 2440–2460. doi:10.3390/rs3112440

Hill, R., Granica, K., Smith, G.M., Schardt, M., 2007. Representation of an alpine treeline ecotone in SPOT 5 HRG data. Remote Sens. Environ. 110, 458–467. doi:10.1016/j.rse.2006.11.031

Hirata, Y., Takahashi, T., 2011. Image segmentation and classification of Landsat Thematic Mapper data using a sampling approach for forest cover assessment. Can. J. For. Res. 41, 35–43. doi:10.1139/X10-130

Ju, J., Gopal, S., Kolaczyk, E.D., 2005. On the choice of spatial and categorical scale in remote sensing land cover classification. Remote Sens. Environ. 96, 62–77. doi:10.1016/j.rse.2005.01.016

Kim, M., Madden, M., Warner, T.A., 2009. Forest type mapping using object-specific texture measures from multispectral Ikonos Imagery: Segmentation quality and image classification issues. Photogramm. Eng. Remote Sensing 75, 819–829.

Kim, M., Warner, T., 2011. Multi-scale GEOBIA with very high spatial resolution digital aerial imagery: scale, texture and image objects. Int. J. Remote Sens. 32, 2825–2850. doi:10.1080/01431161003745608

Körting, T.S., Garcia Fonseca, L.M., Câmara, G., 2013. GeoDMA—Geographic Data Mining Analyst. Comput. Geosci. 57, 133–145. doi:10.1016/j.cageo.2013.02.007

Laliberte, A.S., Rango, A., 2009. Texture and scale in object-based analysis of subdecimeter resolution unmanned aerial vehicle (UAV) imagery. IEEE Trans. Geosci. Remote Sens. 47, 1–10. doi:10.1109/TGRS.2008.2009355

Lee, S., Lathrop, R.G., 2006. Subpixel analysis of landsat ETM + using Self-Organizing Map (SOM) neural networks for urban land cover characterization. IEEE Trans. Geosci. Remote Sens. 44, 1642–1654. doi:10.1109/TGRS.2006.869984

Li, M., Ma, L., Blaschke, T., Cheng, L., Tiede, D., 2016. A systematic comparison of different object-based classification techniques using high spatial resolution imagery in agricultural environments. Int. J. Appl. Earth Obs. Geoinf. 49, 87–98. doi:10.1016/j.jag.2016.01.011

Lu, D., Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance. Int. J. Remote Sens. 28, 823–870. doi:10.1080/01431160600746456

Luo, H., Wang, L., Shao, Z., Li, D., 2015. Development of a multi-scale object-based shadow detection method for high spatial resolution image. Remote Sens. Lett. 6, 59–68. doi:10.1080/2150704X.2014.1001079

Luyssaert, S., Jammet, M., Stoy, P.C., Estel, S., Pongratz, J., Ceschia, E., Churkina, G., Don, A., Erb, K., Ferlicoq, M., Gielen, B., Grünwald, T., Houghton, R.A., Klumpp, K., Knohl, A., Kolb, T., Kuemmerle, T., Laurila, T., Lohila, A., Loustau, D., McGrath, M.J., Meyfroidt, P., Moors, E.J., Naudts, K., Novick, K., Otto, J., Pilegaard, K., Pio, C.A., Rambal, S., Rebmann, C., Ryder, J., Suyker, A.E., Varlagin, A., Wattenbach, M., Dolman, A.J., 2014. Land management and land-cover change have impacts of similar magnitude on surface temperature. Nat. Clim. Chang. 4, 389–393. doi:10.1038/nclimate2196

Ma, L., Cheng, L., Li, M., Liu, Y., Ma, X., 2015. Training set size, scale, and features in Geographic Object-Based Image Analysis of very high resolution unmanned aerial vehicle imagery. ISPRS J. Photogramm. Remote Sens. 102, 14–27. doi:10.1016/j.isprsjprs.2014.12.026

Mahmood, R., Pielke, R.A., Hubbard, K.G., Niyogi, D., Dirmeyer, P.A., McAlpine, C., Carleton, A.M., Hale, R., Gameda, S., Beltrán-Przekurat, A., Baker, B., McNider, R., Legates, D.R., Shepherd, M., Du, J., Blanken, P.D., Frauenfeld, O.W., Nair, U.S., Fall, S., 2014. Land cover changes and their biogeophysical effects on climate. Int. J. Climatol. 34, 929–953. doi:10.1002/joc.3736

Martin, Y., Van Dyck, H., Dendoncker, N., Titeux, N., 2013. Testing instead of assuming the importance of land use change scenarios to model species distributions under climate change. Glob. Ecol. Biogeogr. 22, 1204–1216. doi:10.1111/geb.12087

Memarian, H., Balasundram, S.K., Khosla, R., 2013. Comparison between pixel- and object-based image classification of a tropical landscape using Système Pour l’Observation de la

Terre-5 imagery. J. Appl. Remote Sens. 7, 73512. doi:10.1117/1.JRS.7.073512

Millard, K., Richardson, M., 2015. On the importance of training data sample selection in random forest image classification: a case study in peatland ecosystem mapping. Remote Sens. doi:10.3390/rs70708489

Mishra, N.B., Crews, K. a., 2014. Mapping vegetation morphology types in a dry savanna ecosystem: integrating hierarchical object-based image analysis with Random Forest. Int. J. Remote Sens. 35, 1175–1198. doi:10.1080/01431161.2013.876120

Möller, M., Birger, J., Gidudu, A., Gläßer, C., 2013. A framework for the geometric accuracy assessment of classified objects. Int. J. Remote Sens. 34, 8685–8698. doi:10.1080/01431161.2013.845319

Momeni, R., Aplin, P., Boyd, D., 2016. Mapping complex urban land cover from spaceborne imagery: The influence of spatial resolution, spectral band set and classification approach. Remote Sens. 8, 88. doi:10.3390/rs8020088

Mui, A., He, Y., Weng, Q., 2015. An object-based approach to delineate wetlands across landscapes of varied disturbance with high spatial resolution satellite imagery. ISPRS J. Photogramm. Remote Sens. 109, 30–46. doi:10.1016/j.isprsjprs.2015.08.005

Müller, R., Krauß, T., Lehner, M., Reinartz, P., Forsgren, J., Rönnbäck, G., Karlsson, Å., 2009. IMAGE 2006 European coverage, methodology and results.

Pal, M., Foody, G.M., 2012. Evaluation of SVM, RVM and SMLR for accurate image classification with limited ground data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 5, 1344–1355. doi:10.1109/JSTARS.2012.2215310

Paola, J.D., Schowengerdt, R. a., 1995. A detailed comparison of backpropagation neural network and maximum-likelihood classifiers for urban land use classification. IEEE Trans. Geosci. Remote Sens. 33, 981–996. doi:10.1109/36.406684

R Core Team, 2016. R: a language and environment for statistical computing.

Räsänen, A., Rusanen, A., Kuitunen, M., Lensu, A., 2013. What makes segmentation good? A case study in boreal forest habitat mapping. Int. J. Remote Sens. 34, 8603–8627. doi:10.1080/01431161.2013.845318

Richards, J., Kingsbury, N., 2014. Is there a preferred classifier for operational thematic mapping? IEEE Trans. Geosci. Remote Sens. 52, 2715–2725. doi:10.1109/TGRS.2013.2264831

Robertson, L.D., King, D.J., 2011. Comparison of pixel- and object-based classification in land cover change mapping. Int. J. Remote Sens. 32, 1505–1529. doi:10.1080/01431160903571791

Samat, A., Li, J., Liu, S., Du, P., Miao, Z., Luo, J., 2016. Improved hyperspectral image classification by active learning using pre-designed mixed pixels. Pattern Recognit. 51, 43–58. doi:10.1016/j.patcog.2015.08.019

Shimabukuro, Y.E., Miettinen, J., Beuchle, R., Grecchi, R.C., Simonetti, D., Achard, F., 2015. Estimating burned area in Mato Grosso, Brazil, using an object-based classification method on a systematic sample of medium resolution satellite images. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8, 4502–4508. doi:10.1109/JSTARS.2015.2464097

Stehman, S. V., Arora, M.K., Kasetkasem, T., Varshney, P.K., 2007. Estimation of fuzzy error matrix accuracy measures under stratified random sampling. Photogramm. Eng. Remote Sensing 73, 165–173. doi:10.14358/PERS.73.2.165

Tuanmu, M.-N., Jetz, W., 2014. A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. Glob. Ecol. Biogeogr. 23, 1031–1045. doi:10.1111/geb.12182

Uddin, K., Shrestha, H.L., Murthy, M.S.R., Bajracharya, B., Shrestha, B., Gilani, H., Pradhan, S., Dangol, B., 2015. Development of 2010 national land cover database for the Nepal. J. Environ. Manage. 148, 82–90. doi:10.1016/j.jenvman.2014.07.047

Van Coillie, F.M.B., Verbeke, L.P.C., De Wulf, R.R., 2008. Semi-automated forest stand delineation using wavelet based segmentation of very high resolution optical imagery, in: Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications. pp. 237–256. doi:10.1007/978-3-540-77058-9_13

van de Vlag, D.E., Stein, A., 2007. Incorporating Uncertainty via Hierarchical Classification Using Fuzzy Decision Trees. IEEE Trans. Geosci. Remote Sens. 45, 237–245. doi:10.1109/TGRS.2006.885403

Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S. Springer, New York.

Verbeeck, K., Hermy, M., Van Orshoven, J., 2012. External geo-information in the segmentation of VHR imagery improves the detection of imperviousness in urban neighborhoods. Int. J. Appl. Earth Obs. Geoinf. 18, 428–435. doi:10.1016/j.jag.2012.03.015

Wang, F., 1990. Fuzzy supervised classification of remote sensing images. IEEE Trans. Geosci. Remote Sens. 28, 194–201. doi:10.1109/36.46698

Wang, L., Sousa, W.P., Gong, P., 2004. Integration of object-based and pixel-based classification for mapping mangroves with IKONOS imagery. Int. J. Remote Sens. 25, 5655–5668. doi:10.1080/014311602331291215

Whiteside, T.G., Boggs, G.S., Maier, S.W., 2011. Comparing object-based and pixel-based classifications for mapping savannas. Int. J. Appl. Earth Obs. Geoinf. 13, 884–893. doi:10.1016/j.jag.2011.06.008

Whiteside, T.G., Maier, S.W., Boggs, G.S., 2014. Area-based and location-based validation of classified image objects. Int. J. Appl. Earth Obs. Geoinf. 28, 117–130. doi:10.1016/j.jag.2013.11.009

Wu, C., 2009. Quantifying high-resolution impervious surfaces using spectral mixture analysis. Int. J. Remote Sens. 30, 2915–2932. doi:10.1080/01431160802558634

Yang, J., He, Y., Caspersen, J., Jones, T., 2015. A discrepancy measure for segmentation evaluation from the perspective of object recognition. ISPRS J. Photogramm. Remote Sens. 101, 186–192. doi:10.1016/j.isprsjprs.2014.12.015

Zhang, J., Foody, G.M., 2001. Fully-fuzzy supervised classification of sub-urban land cover from remotely sensed imagery: Statistical and artificial neural network approaches. Int. J. Remote Sens. 22, 615–628. doi:10.1080/01431160050505883

Zhang, Y.J., 1996. A survey on evaluation methods for image segmentation. Pattern Recognit. 29, 1335–1346. doi:10.1016/0031-3203(95)00169-7

Zhou, W., Huang, G., Troy, A., Cadenasso, M.L., 2009. Object-based land cover classification of shaded areas in high spatial resolution imagery of urban areas: A comparison study. Remote Sens. Environ. 113, 1769–1777. doi:10.1016/j.rse.2009.04.007

Zhou, W., Troy, A., 2008. An object-oriented approach for analysing and characterizing urban landscape at the parcel level. Int. J. Remote Sens. 29, 3119–3135. doi:10.1080/01431160701469065

Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., Samanta, A., Piao, S., Nemani, R.R., Myneni, R.B., 2013. Global data sets of vegetation Leaf Area Index (LAI)3g and Fraction of Photosynthetically Active Radiation (FPAR)3g derived from Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) for the period 1981 to 2. Remote Sens. 5, 927. doi:10.3390/rs5020927

LIST OF FIGURE CAPTIONS

Figure 1. Location of the study area.

Figure 2. Data sets used: a) reference land cover map, COS2007, representing Artificial surfaces in red, Agricultural areas in yellow, Forest and semi natural areas in green, Open spaces with little or no vegetation in grey, and Wetlands and water bodies in blue; the square areas outlined in black are the training areas randomly located for the estimation of object-based training statistics. b) LISS-III image collected in Summer of 2006; RGB composition of data acquired in the near infrared, short wave infrared and red bands respectively.

Figure 3. Segmentation results: a) Segmentation output produced with colour=75 and scale=10. b) Segmentation output produced with colour=75 and scale=80.

Figure 4. Segmentation accuracy based on Costa et al.'s (2015) adaptation of Möller et al.'s (2013) method. Dotted horizontal line corresponds to optimal accuracy.

Figure 5. Size of training data sets.

Figure 6. Classification accuracy as a function of the parameter scale with parameter colour set at 1, 25, 50, 75, and 99. The results obtained using the three reduced versions of the mixed training data sets are identified as #1, #2, and #3. Panel a) to e) and f) to j) refer to the artificial neural network and generalized linear model respectively.

Figure 7. Land cover maps obtained using the segmentation output produced with colour=50 and scale=40, the artificial neural network, and different training strategies. a) Pure training (accuracy: 50.4%). b) Mixed training (accuracy: 69.0%). c) Mixed training with reduced samples #1 (accuracy: 61.5%). d) Mixed training with reduced samples #3 (accuracy: 65.8%). Colour legend as in Figure 2.