Available online at www.sciencedirect.com
SCIENCE DIRECT®
Forest Ecology and Management 183 (2003) 31–46
Forest Ecology and Management
www.elsevier.com/locate/foreco

Improved Landsat-based forest mapping in steep mountainous terrain using object-based classification

Luuk K.A. Dorren[a,*], Bernhard Maier[b], Arie C. Seijmonsbergen[a]

[a]*Institute for Biodiversity and Ecosystem Dynamics—Physical Geography, Universiteit van Amsterdam, Nieuwe Achtergracht 166, NL-1018WV Amsterdam, The Netherlands*
[b]*Stand Montafon—Forstfonds, Montafonerstr. 21, A-6780 Schruns, Austria*

Received 17 December 2001; received in revised form 26 November 2002; accepted 3 February 2003

Abstract

The accuracy of forest stand type maps derived from a Landsat Thematic Mapper (Landsat TM) image of a heterogeneous forest covering rugged terrain is generally low. Therefore, the first objective of this study was to assess whether topographic correction of TM bands and adding the digital elevation model (DEM) as additional band improves the accuracy of Landsat TM-based forest stand type mapping in steep mountainous terrain. The second objective of this study was to compare object-based classification with per-pixel classification on the basis of the accuracy and the applicability of the derived forest stand type maps. To fulfil these objectives different classification schemes were applied to both topographically corrected and uncorrected Landsat TM images, both with and without the DEM as additional band. All the classification results were compared on the basis of confusion matrices and kappa statistics. It is found that both topographic correction and classification with the DEM as additional band increase the accuracy of Landsat TM-based forest stand type maps in steep mountainous terrain. Further it was found that the accuracies of per-pixel classifications were slightly higher, but object-based classification seemed to provide better overall results according to local foresters. It is concluded that Landsat TM images could provide basic information at regional scale for compiling forest stand type maps especially if they are classified with an object-based technique.

Keywords: Forest mapping; Segmentation; Landsat TM; Topographic correction; Mountain forest

1. Introduction

Intensification of land use in mountainous terrain demands integrated watershed planning, of which forest management is an important part. Forest mapping and forest ecological research form an important contribution to management strategies and plans. Since the 1970s, many authors have been investigating how remote sensing could contribute to forest mapping (Peterson et al., 1987; Ardö, 1992; Congalton et al., 1993; Gemmell, 1995; Martin et al., 1998; Kilpeläinen and Tokola, 1999; Pax-Lenney et al., 2001; Tokola et al., 2001) and to forest ecological research (Ekstrand, 1994; He et al., 1998; Lucas and Curran, 1999; Coops and Culvenor, 2000). The usefulness of Landsat TM imagery as an aid in forest management is generally agreed if TM imagery is combined with field data and high-resolution imagery (Gemmell, 1995; Trotter et al., 1997; Salvador and Pons, 1998; Kilpeläinen and Tokola, 1999; Hyyppä

* Corresponding author. Tel.: +31-20-525-7420; fax: +31-20-525-7431.
E-mail address: l.dorren@science.uva.nl (L.K.A. Dorren).

0378-1127/$ – see front matter
doi:10.1016/S0378-1127(03)00113-0

et al., 2000; Bebi et al., 2001). However, the resolution of Landsat TM imagery is in many cases too low to derive forest parameters required by foresters such as timber volume, basal area and tree height. Consequently, TM imagery is generally used in forestry to provide contemporary and historic forest stand type maps (Anderson level II; see Anderson et al., 1976) and forest cover maps for larger regions. These maps could be used for assessing large-area regeneration, (re-)growth and causes of distinct changes in forests (Fiorella and Ripple, 1993).

The problem with Landsat TM-based forest stand type maps of mountainous terrain, derived with traditional per-pixel classification methods, is that the accuracy is fairly low, even if a digital elevation model (DEM) is added as extra band (Frank, 1988; Itten and Meyer, 1993). Accuracies of forest stand type maps of less-rugged or flat terrain are generally higher (see Mickelson et al., 1998). One of the reasons for this is that both land cover and topography determine the spectral values in remote sensing imagery, especially in case of steep or high relief energy areas. Strong variability in the reflectance from canopies of similar forests and direct shadows as well as cast shadows are indissoluble results of the topography in such areas, which complicate the classification. Topographic correction solves this problem according to many authors (Leprieur et al., 1988; Civco, 1989; Ekstrand, 1996; Tokola et al., 2001), although others do not agree on this (Carpenter et al., 1999). Despite topographic correction, traditional per-pixel methods still encounter problems with deriving forest stand type maps from Landsat TM images of mountainous terrain, because of the large variability in reflectance within forest stands and spectral confusion of species (Meyer et al., 1993).

Several authors have already referred to classifying segmented or merged pixels of Landsat TM images as an alternative to per-pixel methods (Ton et al., 1991; Woodcock and Harward, 1992; Stuckens et al., 2000). Image segmentation has the potential to merge pixels into objects that correspond to forest stands (Woodcock and Harward, 1992). Accordingly, this method is called object-based classification. This method could improve Landsat TM-based forest mapping in steep mountainous terrain. Therefore, the first objective of this study was to assess whether topographic correction of Landsat TM bands and adding a DEM as additional band results in improved forest stand type mapping in steep mountainous terrain. The second objective was to evaluate which classification method performed best: traditional per-pixel classification or object-based classification. These objectives raise the following research questions:

1. What is the effect of topographic correction on the accuracy of Landsat TM-based forest stand type maps?
2. Does the classification of a Landsat TM image improve when a DEM is used as additional band during the analysis?
3. Does object-based classification improve the accuracy and the applicability of Landsat TM-based forest stand type maps compared to per-pixel classification?

To answer these questions, we firstly prepared the used data, which also included the extraction of a forest mask. This mask was used for both classification methods. Subsequently, we performed the per-pixel classification and interpreted the results. After that we carried out the object-based classification and compared those results with the per-pixel classification. The set-up of this paper is identical to our research procedure, which is shown in the flow diagram in Fig. 1.

2. Materials

2.1. Study area

The Montafon region, which covers approximately 530 km^2, served as a test site for this study. It is located in the southern part of Vorarlberg, the westernmost province of Austria (between 47°80′and 46°50′ latitude and 9°41′and 10°90′ longitude, see Fig. 2). As an alpine region its altitude ranges from 600 m above sea level (a.s.l.) at the valley bottom up to over 3000 m a.s.l. The average relief energy is approximately 250 m/km^2. The geology is determined by sedimentary rocks in the northern part and by metamorphic and crystalline rocks in the southern part (GBA, 1998). The Montafon valleys are open to the northern windward side of the Alps and therefore under strong oceanic influence (mean annual temperature approximately 8 °C, mean annual precipitation at the bottom

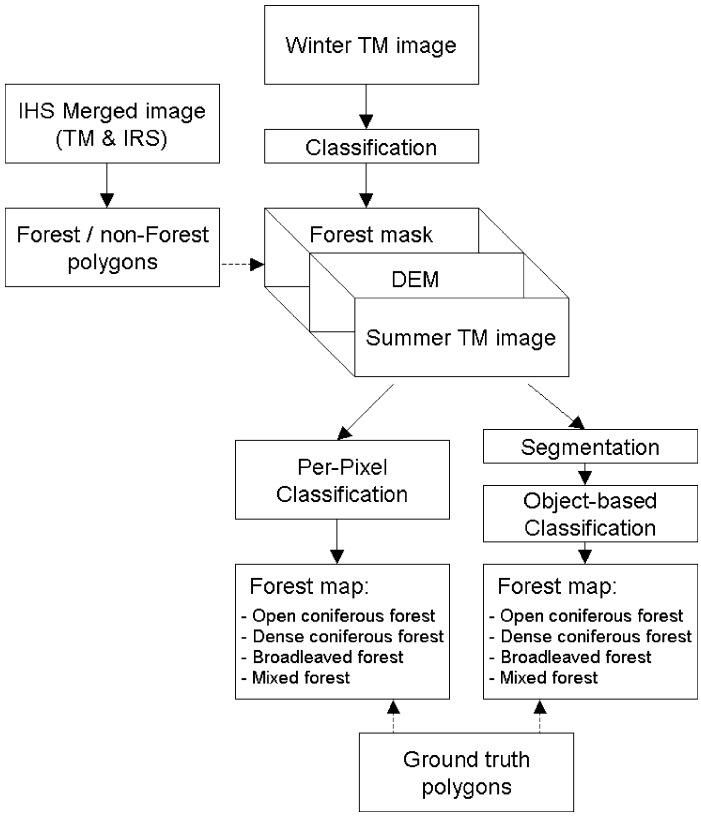

Fig. 1. Flow diagram of the research procedure (the dotted arrows refer to accuracy assessment).

of the main valley approximately 1300 mm). The forests reach up to about 1800 m a.s.l. and cover a total area of approximately 12,500 ha. Thirty-three percent of the forested area is steeper than 45°. The forests provide essential protection for villages and infrastructures against snow avalanches and rockfall. In addition, these forests are important for, amongst others, timber production and recreation. Therefore, the forests are managed in a multifunctional sustainable way. A local forest organisation is responsible for forests management in this region. For this they need information on the spatial coverage of the three main forest types: broadleaved forest, mixed forest and deciduous forest. The broadleaved (*Fagus silvatica*, *Acer pseudoplatanus*, *Tilia cordata*, *Fraxinus excelsior*) and mixed forests (*Picea abies*, *F. silvatica*, *Abies alba*) are mainly found in the valleys and up to 1000 m a.s.l. Spruce (*P. abies*) forests predominate above 1000 m a.s.l., though the mixed forests reach up to 1500 m a.s.l. *Larix decidua* and *Pinus cembra* are only found in fragmental stands close to the timberline (Maier, 1993). The distribution of forest stand types at regional scale is mainly determined by altitude, but at slope scale other factors such as parent material, relief, microclimate, humus forms, light and disturbances caused by natural hazards also become important (van Noord, 1996). Generally, the canopy density decreases with increasing altitude due to the changing climatic conditions, e.g. heavier snow loads. Because uneven-aged forests predominate, tree distribution, stem density, tree height and basal area vary strongly between forest stands. As typical for mountain forests, these varied forest stands are distributed in a patchy mosaic-like structure (Maier, 1993).

2.2. Satellite imagery and digital elevation data

A summer and a winter Landsat TM5 scene (respectively, recorded 9 September 1998 and 28 January 1998, both at 9:40 CMT) and an IRS-1C (Indian

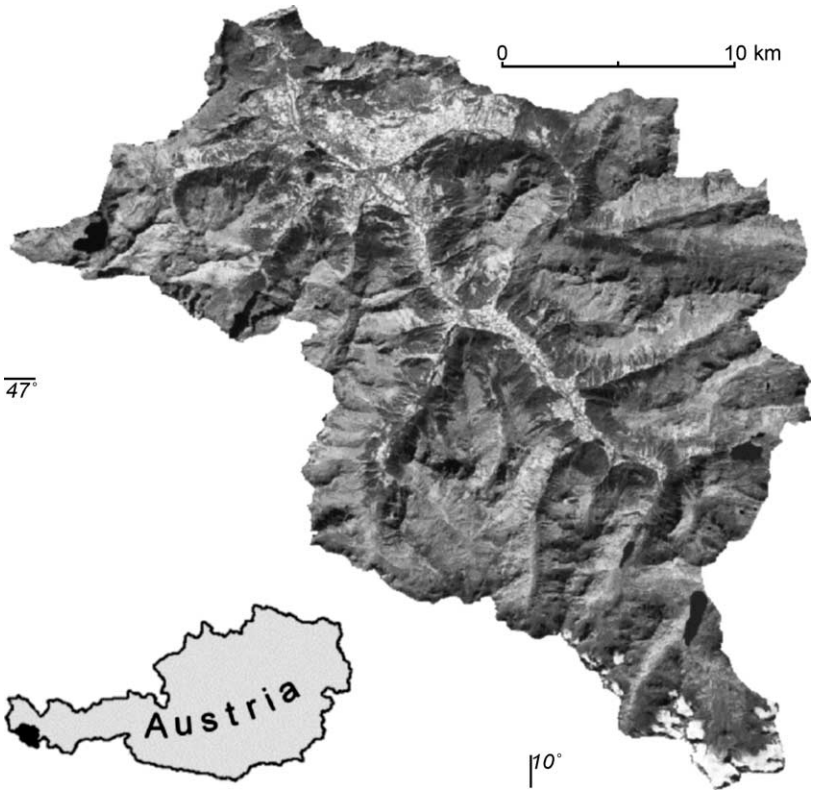

Fig. 2. Study area: the Montafon region (near-IR band of the September 1998 Landsat TM image).

Remote Sensing Satellite) panchromatic scene (recorded at 25 September 1997, 10:35 CMT) were used in this study. The resolution of a Landsat TM scene is 30 m × 30 m; the resolution of a panchromatic IRS-1C scene is 6 m × 6 m. A DEM produced by the National Austrian Mapping Agency was available for this study. This DEM is arranged as a regular grid of surface height points at 25 m spacing. According to the product information (BEV, 2001), this DEM was created by interpolation of photogrammetric height measurements at a ground distance of 50 m, enhanced and supplemented with prominent terrain structures. The modelled elevation accuracy is given as ±1 to ±3 m in open and flat land and ±5 to ±20 m in forests and high-alpine terrain.

2.3. Validation and training data

Local foresters digitised ground truth polygons on the basis of digital colour-infrared (CIR) orthophotos, which covered different areas with: open coniferous forest, dense coniferous forest, broadleaved forest and mixed forest. Mixed forests are stands where neither broadleaved nor coniferous trees account for more than 75% of the tree crown area, in accordance with the definition of UN-ECE/FAO (2000). The CIR orthophotos were taken in the period July–September 1996 and have a resolution of 0.25 m × 0.25 m. Ground truth polygons were created in areas where tree species, tree distribution and stem density were comparable in 1996 and 1998. The ground truth polygons were converted into a raster map with a cell size of 25 m × 25 m. Then, this map was randomly divided in training areas and validation areas. The validation areas were used for the accuracy assessment. The accuracies of all the classifications were assessed on the basis of confusion or error matrices (Congalton, 1991; Richards and Jia, 1999) and kappa statistics (KHAT) (Cohen, 1960; Hudson and Ramm, 1987; Congalton, 1991; Congalton and Plourde, 2000).

3. Data preparation methods

3.1. Radiometric and geometric pre-processing

Radiometric correction was applied to all the bands of the two Landsat TM scenes by multiplying by the gain and by adding the offset, which are both provided in the image header files. The images used for this study were extracted from the corrected scenes and covered approximately 35 km × 35 km. Fortunately, no detector errors or clouds were present in the used images.

The topography of mountainous terrain is responsible for geometric distortions (Itten and Meyer, 1993; Richter, 1997; Sandmeier and Itten, 1997), but in the case of steep mountainous terrain these distortions increase, therefore approximately 50 ground control points were used for orthorectifying both Landsat TM images and the IRS image. The RMS error of the rectification was 0.57 pixels (17.1 m) for the January 1998 image and 0.43 pixels (12.9 m) for the September 1998 image. Both images were resampled to 25 m× 25 m resolution using a nearest neighbour algorithm. The IRS-1C Panchromatic image was orthorectified with an RMS error of 0.26 pixels (1.56 m) and resampled to 5 m × 5 m resolution using a cubic convolution algorithm, since this image has only been used for visual interpretation and not for classification.

3.2. Extracting a forest mask

Forest masks were derived using maximum likelihood and Parallelepiped classification of the bands 1–5 of the January 1998 Landsat TM image and the bands 2, 4 and 5 of the September Landsat TM 1998 image. The January image was tested because firstly shadows are not a problem in this image, since scatter light enlightens shadowed forested areas due to the snow cover, even on steep north facing slopes, which show limited reflection in summer images. Secondly, there is no confusion between shrubs and broadleaved forests, because small shrubs are covered under a thick snow pack. For each of the classification results, confusion matrices and KHAT values were calculated on the basis of forest/non-forest polygons. These polygons were digitised on screen on the basis of a merged IRS-TM image, which was created with an Intensity-Hue-Saturation transformation (Carper et al., 1990) using the September 1998 Landsat TM image and the IRS-1C Panchromatic image. maximum likelihood classification of the bands 1–5 of the January 1998 Landsat TM image provided the most accurate forest mask (overall accuracy: 0.94, KHAT value: 0.73).

3.3. Topographic correction

To account for large variations in surface illumination that existed between areas that receive direct sunlight and areas in complete shade, the September 1998 image was corrected for topographic effects. Currently, a wide range of algorithms for topographic correction exists (Smith et al., 1980; Teillet et al., 1982; Colby, 1991; Conese et al., 1993; Meyer et al., 1993; Franklin and Giles, 1995; Sandmeier and Itten, 1997; Degui et al., 1999). The most applied algorithms are the Cosine Correction, the C-correction and the (extended) Minnaert Correction, which is based on the Minnaert constant (Minnaert, 1941) to account for non-Lambertian behaviour of vegetated surfaces (Jansa, 1998). Gu and Gillespie (1998) developed the Sun-Canopy-Sensor-correction (SCS-correction) specifically for removing topographic effects in Landsat TM images of forested areas. The SCS-correction, the Cosine Correction, the C-correction and the Minnaert Correction were evaluated by plotting reflectance values of the masked (using the forest mask) topographically corrected September 1998 TM band 4 versus the incidence angles (Meyer et al., 1993; Richter, 1998). These scatterplots indicated that the SCS-correction performed best in reducing the correlation between incidence angle and reflectance value. Therefore the SCS-correction was used in this study, of which a subset of the resulting image is shown in Fig. 3. The SCS-correction is based on the following algorithm:

$$\mathrm{DN}_{\mathrm{corr}} = \mathrm{DN}_{\mathrm{in}} \frac{\cos(\beta)\cos(\mathrm{sz})}{\cos(i)} \tag{1}$$

where $\mathrm{DN}_{\mathrm{corr}}$ is the corrected DN value, $\mathrm{DN}_{\mathrm{in}}$ the uncorrected DN value, β the slope angle (°), sz the solar zenith angle (°), i the incident angle (°), which is the angle between the direction normal to the surface and the solar beam and could be calculated following (Holben and Justice, 1980):

$$\cos(i) = \cos(\mathrm{sz})\cos(\beta) + \sin(\mathrm{sz})\sin(\beta)\cos(\mathrm{az\text{-}aspect}) \tag{2}$$

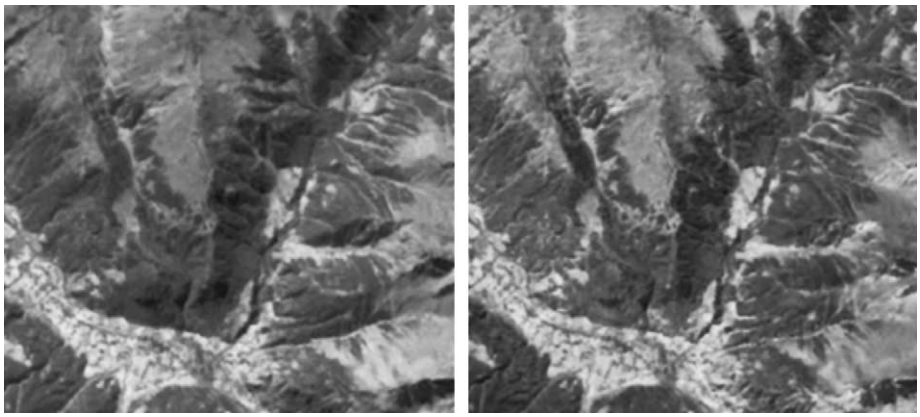

Fig. 3. Left image: subset of the raw September 1998 TM band 4; right image: subset of the topographic corrected September 1998 TM band 4 using the SCS-correction (black pixels are slopes with incidence angles larger than 90°).

where az is the solar azimuth angle (°), aspect the aspect of the slope (°).

4. Per-pixel classification methodology

After pre-processing, extracting the forest mask and applying topographic correction to the September 1998 Landsat TM image, per-pixel classifications were carried out with different datasets. All the datasets consisted of TM band 5, TM band 4 and a synthetic band (TM band 4 − TM band 2). The synthetic band was used for separating broadleaved and mixed forest (Itten et al., 1992). Spectral histograms and bi-spectral plot analyses of the signatures for the training areas indicated that the use of this combination of bands would optimise the classification since the separability of the classes was optimal within these bands.

Either topographically corrected or uncorrected TM bands were classified during each classification test. Additionally, each set of bands was classified with and without the DEM as additional band, to evaluate its effect on the classification result. This technique has been called the 'logical channel approach' (Strahler et al., 1978; Hutchinson, 1982). Here, the DEM was stretched to obtain values ranging from 0 to 255. All the different datasets were masked during classification by using the forest mask. Each classification produced four forest classes (open coniferous forest, dense coniferous forest, broadleaved forest and mixed forest) as well as a set of unclassified pixels.

It was decided to use maximum likelihood for the per-pixel classifications. The justification for this decision was based on KHAT values of preliminary results of classification tests with various unsupervised and supervised classification techniques (Richards and Jia, 1999) and additional tests with classifying principal components.

5. Per-pixel classifications: results and discussion

5.1. The effect of topographic correction

Table 1 presents the overall accuracy and the KHAT values of the different per-pixel classifications. These were calculated from the error matrices in which the classification results were compared with the validation data. This table shows that the accuracies of per-pixel classifications using topographic corrected bands are higher than those during which uncorrected bands were used. In the first case, the overall accuracy increased with 0.10 and the KHAT value increased with 0.11 when comparing classification numbers 1 and 2 (Table 1). In case of classification numbers 3 and 4, the overall accuracy increased with 0.19 and the KHAT value increased with 0.22 (Table 1). Another distinct effect of topographic correction is shown by the distribution of errors over the incidence angles (Fig. 4). This figure shows that the classification errors in the case topographic corrected images were used are more or less regularly distributed over all

Table 1
Summary of the per-pixel classification results

Classification number	Used dataset	Overall accuracy ($n = 7411$)	KHAT ($n = 7411$)
1	DEM, TM4, TM5, TM4 – TM2	0.73	0.63
2	DEM, TM4[a], TM5[a], TM4 – TM2[a]	0.63	0.52
3	TM4, TM5, TM4 – TM2	0.64	0.51
4	TM4[a], TM5[a], TM4 – TM2[a]	0.45	0.29

[a] Bands not corrected for topography.

incidence angles. In contrast, the classification errors in the case uncorrected images were used are more concentrated in faint illuminated pixels. These results show that topographic correction improves the accuracy of Landsat TM-based forest stand type maps in steep mountainous terrain. These findings agree with Meyer et al. (1993), Richter (1998) and Tokola et al. (2001).

5.2. Optimal dataset for per-pixel classification

As shown in Table 1, per-pixel classification of the topographic corrected TM bands 4, 5 and the synthetic band in combination with the DEM produced the highest overall accuracy (0.73) and KHAT value (0.63). Per-pixel classifications of the other datasets produced lower accuracies. The importance of the DEM for forest mapping in steep mountainous terrain could be explained by the fact that the distribution of forest stand types on regional scale is mainly determined by altitude. Visual evaluations of the classification results showed that classifications of TM bands without the DEM as additional band resulted in unrealistic distributions of forest stand types, such as broadleaved forests near the tree line (1800 m a.s.l.). These results show that it is necessary to use

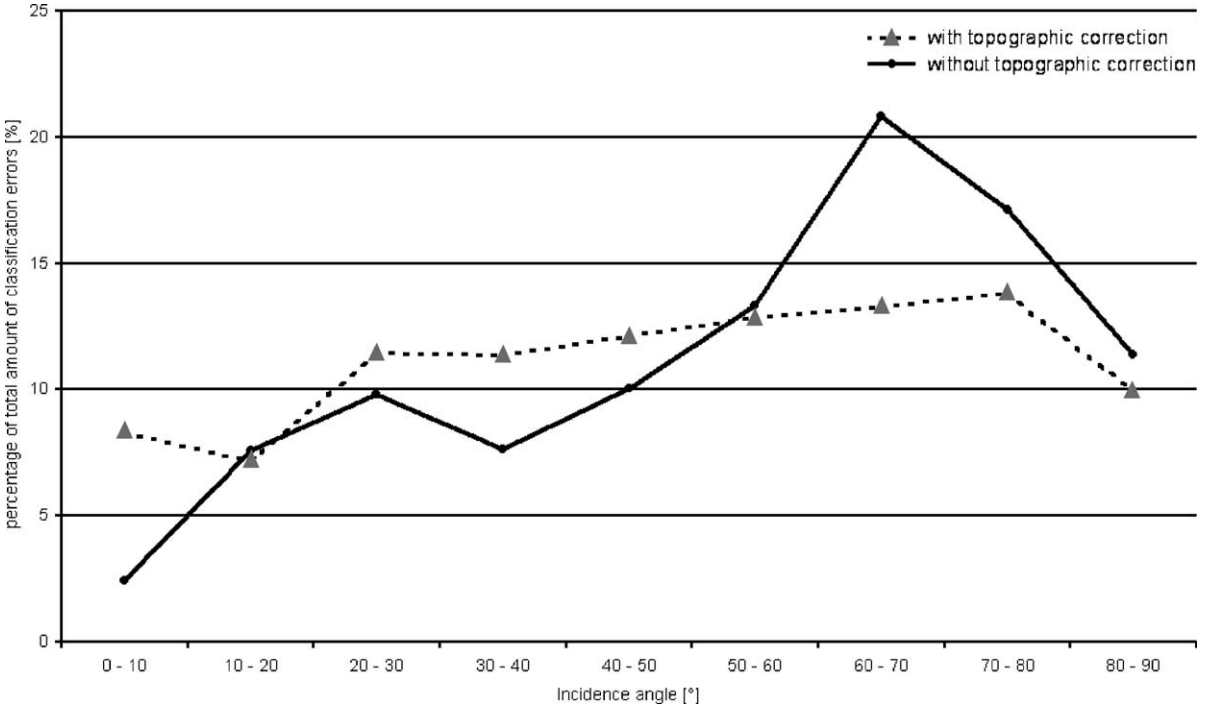

Fig. 4. The effect of topographic correction (SCS-correction) on the distribution of classification errors over different incidence angles. Percentages are normalised for the total number of pixels within each incidence angle class.

a DEM as additional band when classifying a Landsat TM image into forest stand types in steep mountainous terrain. For per-pixel classification of forest stand types in the study area, the combination of topographic corrected bands and the DEM provided the most optimal dataset tested. This dataset was also used for the object-based classification tests.

6. Object-based classification methodology

During the first step of the object-based classification method, image objects are created by means of segmentation. The segmentation process in the used software *eCognition* (Definiens-Imaging, 2001) is a bottom up region merging technique starting with randomly selected one-pixel objects, where a binary counter guarantees a regular spatial distribution and simultaneous growth of treated objects (Baatz and Schäpe, 2000). Throughout the segmentation process, an underlying optimisation procedure minimises the weighted heterogeneity ($n \times h$) of the created image objects, where n is the size of the object, defined by a scale parameter and h is a homogeneity criterion, defined by three sub-criteria: spectral heterogeneity, compactness and smoothness, which could be weighted as shown in Fig. 5. The algorithms of the segmentation process have been described in detail by Baatz and Schäpe (2000). Firstly, the forest mask, which was obtained with a supervised classification as described in Section 3.2, was segmented into forest and non-forest objects. These objects had an average size of 270 pixels. These objects were embedded in a data layer referred to as level 2. Level 2 ensured that during the eventual classification the same forest pixels according to the forest mask were classified as during the per-pixel classification. Subsequently, the large forest and non-forest objects were divided into smaller segments. The basis for this segmentation was the topographic corrected September 1998 TM band 4 and TM band 5. These smaller objects were embedded in a second data layer, hereafter referred to as level 1 (Fig. 6). Finally, all the objects in level 1 were classified using the same dataset that

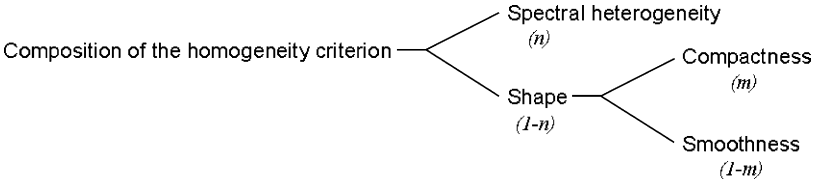

Fig. 5. Overview of the weighted sub-criteria of the homogeneity criterion, on the basis of which pixels are merged into objects.

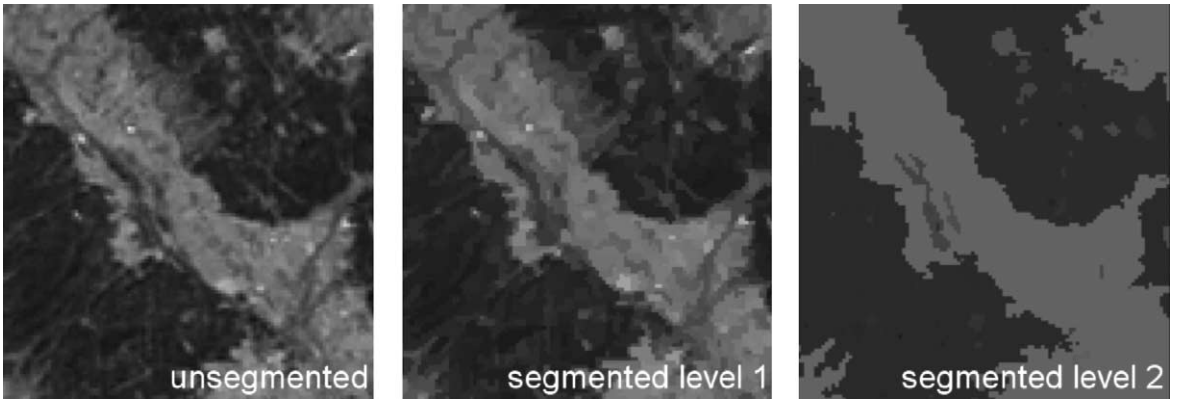

Fig. 6. Three hierarchical levels. Left image: pixel level (unsegmented TM band 4); centre image: segmented level 1 (small segmented objects, e.g. individual forest stands); right image: segmented level 2 (large forest/non-forest objects).

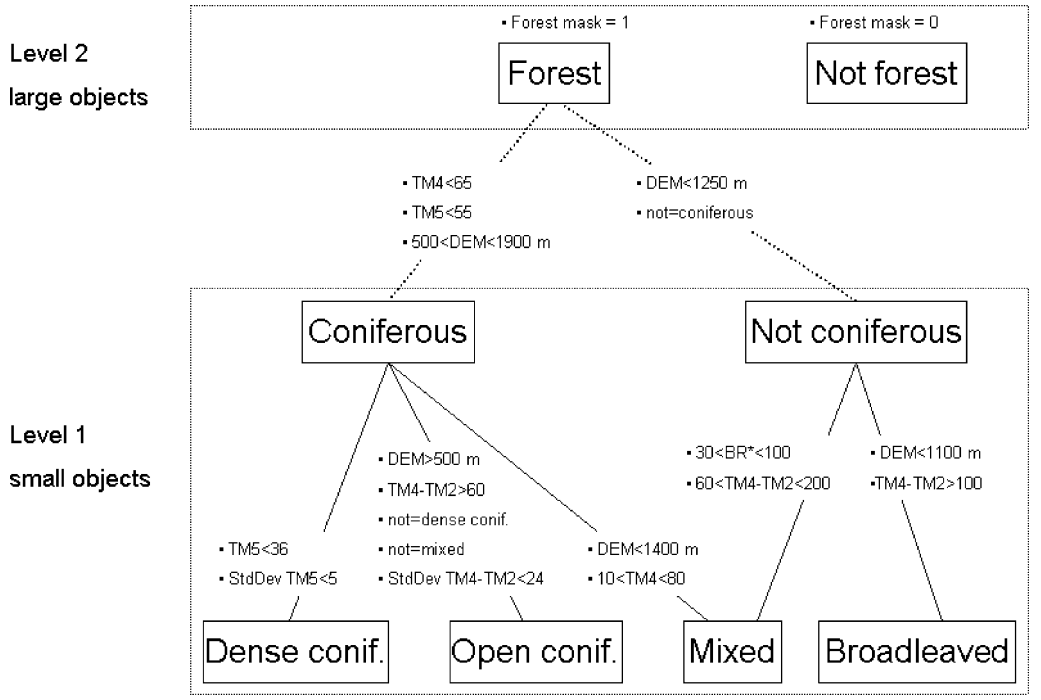

Fig. 7. Membership functions for all the classes. BR* denotes brightness = (mean TM5 + mean TM4 + mean (TM4 − TM2))/3. Objects that do not fall into any of the classes remain unclassified.

provided the best per-pixel classification result. This included TM band 4, TM band 5, the synthetic band TM4 − TM2, all corrected for topography and the DEM. The classification was based on a decision tree, which is presented in Fig. 7. This tree was built up from different membership functions for the four forest classes: mixed forest, dense coniferous forest, open coniferous forest and broadleaved forest. The membership functions for these forest classes were defined on the basis of the ground truth areas as depicted on the train map. This has effect on the portability of the developed method to other regions, since specific ground truth is needed to define the membership functions.

We tested 30 object-based classifications. For each of those different sizes of the objects in level 1 were used. The size of the objects ranged from an average object size of 2.0 pixels to 78.2 pixels. Eventually were applied. An overview of the object-based classification methodology is presented in Table 2.

Table 2
Outline of the object-based classification methodology

Step	Method	Average object size (pixels)	Used dataset	Segmentation weights
1	Segmentation of the forest mask—creation of level 2	270	Forest mask	Spectral heterogeneity: 0.6, shape: 0.4, compactness: 0.5, smoothness: 0.5
2	Creation of smaller segments—result is level 1	2.0 for the first classification to 78.2 for the last classification	TM5 and TM4	Spectral heterogeneity: 0.6, shape: 0.4, compactness: 0.5, smoothness: 0.5
3	Classification of level 1	2.0 for the first classification to 78.2 for the last classification	DEM, TM4, TM5 and TM4 − TM2	n.a.

7. Final results: object-based versus per-pixel classification

7.1. Object size and object-based classification accuracy

Fig. 8 shows the average object size versus the KHAT values and the overall accuracy of the object-based classifications. The overall accuracies and the KHAT values of the object-based classifications are dependent of the average object size. The best object-based was based on segmented objects with an average size of 21.6 pixels. The overall accuracy and the KHAT value of this classification were, respectively, 0.70 and 0.58 (Table 3). Smaller and larger objects produced lower classification accuracies. The overall accuracy and KHAT value when using smaller objects decreased to, respectively, 0.63 and 0.49 (average object size: 2.0 pixels). When using larger objects, the overall accuracy and KHAT value decreased to, respectively, 0.59 and 0.46 (average object size: 78.2 pixels).

7.2. Object-based versus per-pixel classification accuracy

Table 3 summarises the classification results of the best object-based and per-pixel classification.

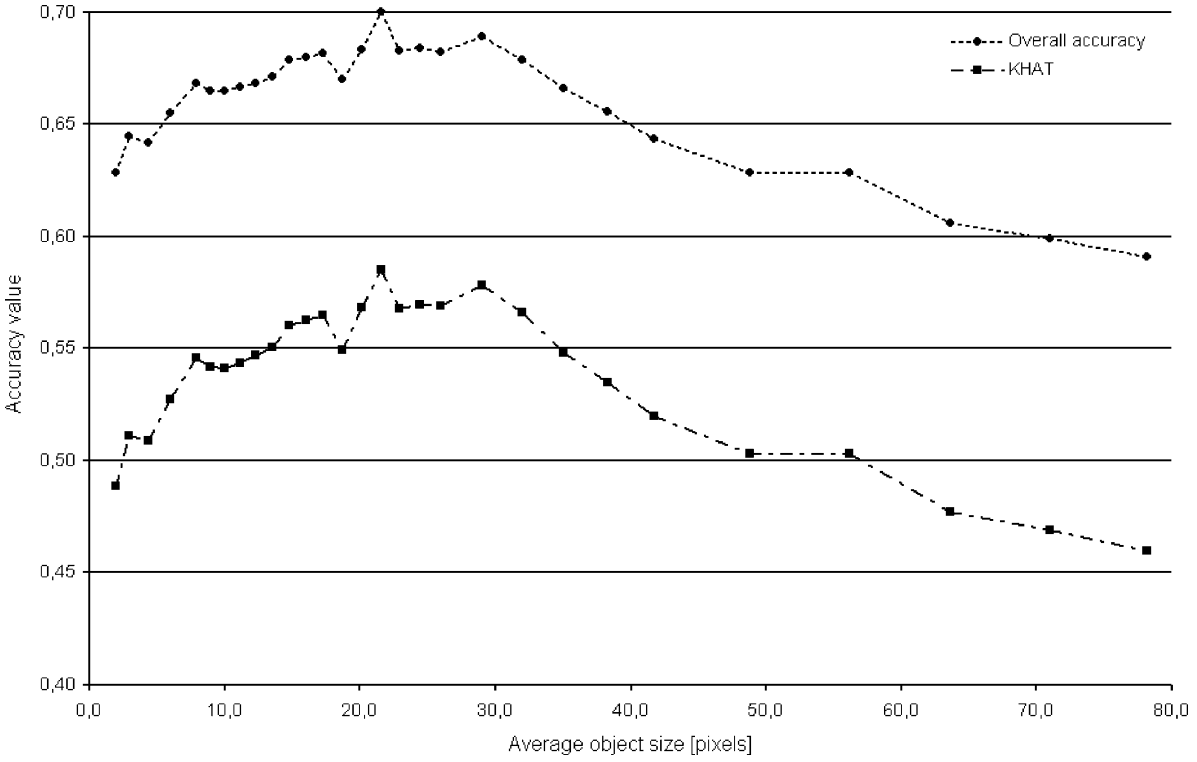

Fig. 8. Average object size versus the overall accuracy and the KHAT values of the object-based classifications.

Table 3
Summary of the best per-pixel and object-based classification results

Classification method	Average object size (pixels)	Used dataset	Overall accuracy ($n = 7411$)	KHAT ($n = 7411$)
Per-pixel	–	DEM, TM4 – TM2, TM5, TM4	0.73	0.63
Object-based	21.6	DEM, TM4 – TM2, TM5, TM4	0.70	0.58

Table 4
Confusion matrix of the best per-pixel and object-based classification, normalised for the amount of pixels within a ground truth class

	Validation/ground truth			
	Dense coniferous ($n = 2632$)	Mixed ($n = 2460$)	Broadleaved ($n = 1258$)	Open coniferous ($n = 1061$)
Per-pixel classification				
Dense coniferous	**0.814** NS[a]	0.133	0.002	0.132
Mixed	0.101	**0.720** NS	0.106	0
Broadleaved	0	0.125	**0.737** PP[b]	0
Open coniferous	0.079	0	0	**0.515** PP
Unclassified	0.006	0.022	0.155	0.353
Object-based classification				
Dense coniferous	**0.787** NS	0.103	0.006	0.155
Mixed	0.126	**0.732** NS	0.196	0.030
Broadleaved	0	0.143	**0.668** PP	0
Open coniferous	0.050	0	0	**0.415** PP
Unclassified	0.037	0.022	0.130	0.400

[a] No significant difference between both methods at the 0.01 confidence level.
[b] Per-pixel method significant better at the 0.01 confidence level.

This table shows that the overall accuracy and the KHAT value of the per-pixel classification are, respectively, 0.03 and 0.05 higher than the object-based classification. The difference in accuracy between the two methods is shown in more detail by the confusion matrices of the best per-pixel and object-based classification (Table 4). These confusion matrices tabulate the proportion of pixels matching ground truth pixels, normalised for the amount of pixels within a ground truth class. A binomial test on the class accuracies (applying normal approximation with test statistic Z) showed that no significant ($\alpha = 0.01$) differences exist between the two classification methods for the classes 'dense coniferous forest' and 'mixed forest'. The classes 'broadleaved' and 'open coniferous' were significantly better classified by the per-pixel method.

Both methods are able to classify at least 67% of the validation pixels of the classes 'dense coniferous', 'mixed' and 'broadleaved' forest correctly, but many pixels within the class 'open coniferous' could not be classified as such. This especially accounts for the object-based classification, where 40% of pixels of the class 'open coniferous' remained unclassified. Another striking difference between the two methods as shown in Table 4 is the confusion between broadleaved and mixed forest stands as classified by the object-based method, since almost 20% of the class 'broadleaved' forest has been classified as mixed forest stands.

7.3. Differences in forest stand type maps

During field visits a strong variation between open coniferous, dense coniferous, broadleaved and mixed forest stands was observed on the lower parts of the oversteepened valley slopes in the study area. Therefore, during the visual evaluation of the classification results the variation of forest stand types in this altitudinal zone was firstly evaluated. As shown in Fig. 9, the map produced by the object-based classification showed an alternation of mainly broadleaved and mixed forest stands within this zone, but also some dense and open coniferous forest stands. The per-pixel classification resulted in rather large homogeneous areas.

8. Final discussion

8.1. Accuracy and applicability of derived maps

Results in this study show that object-based classification compared to per-pixel classification does not improve the accuracy of Landsat TM derived forest stand type maps according to the overall accuracies and the KHAT values. At least, this accounts when using TM bands 4, 5, and a synthetic band (TM band 4 − TM band 2), all corrected for topography, in

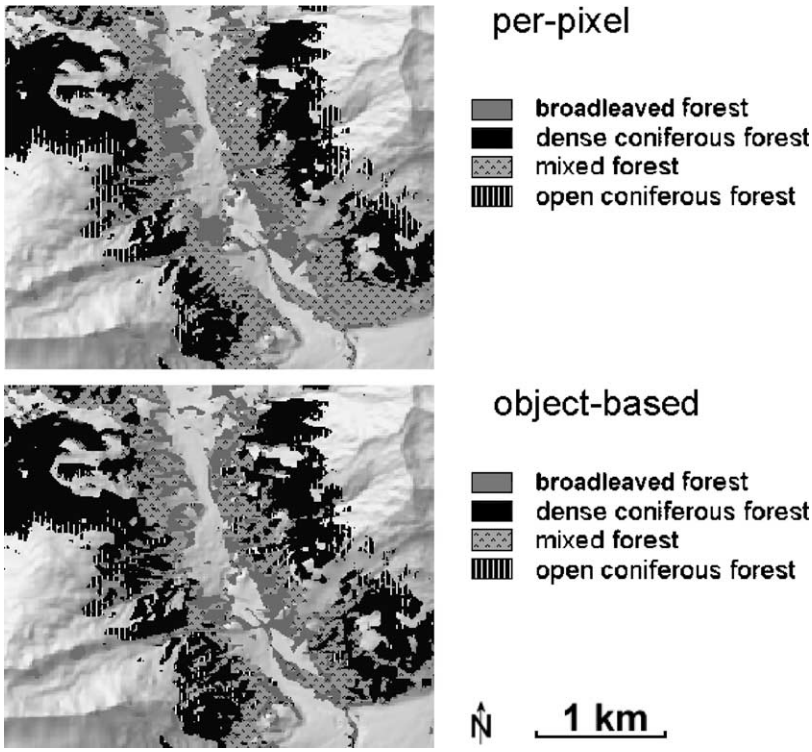

Fig. 9. Upper image: subset of the forest stand type map produced by the per-pixel classification. Lower image: subset of the forest stand type map produced by the object-based classification. The central part of the figure shows that per-pixel classification resulted in homogeneous altitudinal vegetation zones, whereas the object-based classification produced more heterogeneity, especially within the lowest forested altitudinal zone.

combination with a DEM. For forest stand type mapping in the study area using per-pixel classification, this dataset provided the most accurate results. This was not the case for object-based classification, since a large amount of open coniferous forest stands could not be classified as such (Table 4). Furthermore, the object-based classification encountered problems while separating broadleaved forest stands from mixed forest stands. This problem could probably be solved with a more specific training dataset, which is known to be required for an object-based classification as used in this study (Lobo and Chic, 1996). A specific training dataset for object-based classification might provide a better combination of TM bands and additional data. On the basis of such a dataset, the separation of the classes 'broadleaved' and 'mixed' forest could most likely be improved.

Despite lower accuracies, object-based classification compared to per-pixel classification does improve the applicability of Landsat TM-based forest stand type maps. More specifically, the forest stand type map produced by the object-based classification showed more agreement with the field situation than the map produced by the per-pixel classification. Local foresters working in the study area confirmed this. The object-based map produced a strong variation in forest stand types as observed on the lower parts of the oversteepened valley slopes in the study area. These slopes are transportation zones of snow avalanches, rockfalls, debris flows and landslides. Therefore, large variations in mixture, age and tree distribution exist (Maier, 1993; van Noord, 1996). On these slopes the per-pixel classification produced large homogeneous forested areas, which were mainly determined by altitude. The variation in forest stand types observed on the oversteepened slope in the study area was not produced.

Why, then, does object-based classification produce lower accuracies? One reason is the location of ground truth polygons in the study area. These polygons were

obtained by digitising large forest stands that are representative for each classified forest stand type. But large broadleaved forest stands are rare in the study area. Smaller broadleaved stands prevail, especially in areas where the stand type variability is high as a result of site, growth and disturbance factors. Therefore, a large set of validation data for the class 'broadleaved' was lacking. Using the DEM as additional band in the per-pixel classification resulted in a map showing strongly banded vegetation zones determined by altitude (Fig. 9). Therefore, the per-pixel classification performed very well according to the confusion matrix, since the validation pixels that were available were generally determined by altitude.

In the case of object-based classification, many different forest stand types, in particular within the lower altitudinal zones, were produced. As a result, the probability of classified forest stand objects mismatching ground truth pixels increased. It was observed that mainly the edges of delineated forest stands mismatched ground truth pixels (Fig. 10). Obviously, this led to an increase of the amount of classification errors and an according decrease of the classification accuracy. Classification errors were expressed as the amount of mismatching pixels. Instead, expressing classification errors as the amount of objects in which the majority of pixels mismatched ground truth pixels would be a more suitable method for assessing accuracy of object-based classifications. Regarding the pixel-based accuracy assessment used in this study, a random sampled ground truth set based on field inventories would probably have been better. However, in steep mountainous terrain it is very time and labour consuming to obtain such a set.

8.2. Knowledge-based classification of objects

The relation between the average object size and the accuracy of forest stand type maps derived with object-based classification, indicated that a per-pixel method is not a good basis for stand type classification. One reason is that per-pixel classification ignores the fact that there is a high probability that neighbouring pixels belong to the same class (Stuckens et al., 2000). As stated by Woodcock and Harward (1992) the ideal situation is reached when the image objects correspond with objects in the ground scene. This statement was confirmed by the results of the object-based classification of this study. These results showed that the classification accuracy decreased when using too small or too large objects. An explanation for this is that the standard deviation of objects, which was e.g. used for identifying open coniferous stands, was not suitable for classifying small objects. If objects were too large, the standard deviation within an object was too large to identify pure coniferous stands. Within this study, an average object size of 21.6 pixels resulted in the highest accuracy. This size is comparable to the average stand size in the study area, which confirms the statement of Woodcock and Harward (1992).

If segmentation provides image objects that correspond with objects in the ground scene, then an intelligent classifier should be able to classify the image objects perfectly. At present, segmentation

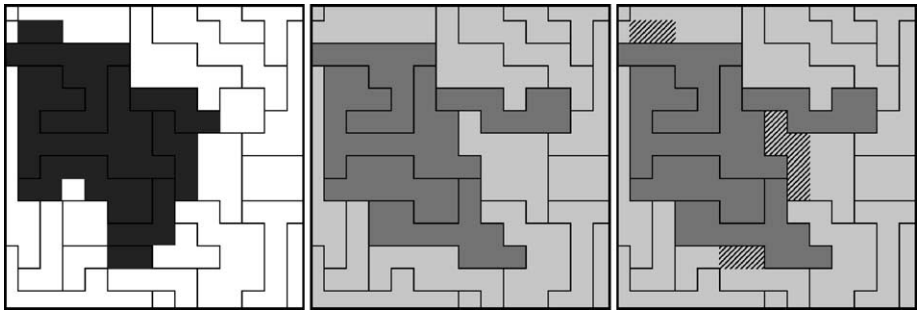

Fig. 10. Left image: hypothetical segmented image objects (black outlined) and a rasterized ground truth polygon (dark grey); centre image: hypothetical object-based classification result showing two classes (medium grey and light grey); right image: mismatched pixels (indicated with diagonal pattern) if medium grey class is identical to the ground truth class.

and object-based classification works very well for images with spatially discontinuous elements with abrupt borders such as large agricultural fields or metropolitan areas surrounded by forests (Stuckens et al., 2000). In case of segmenting Landsat TM images of steep forested mountainous terrain into forest stands, both a sophisticated segmentation procedure and an intelligent classifier are needed, since forest stands are elements with gradual and fuzzy transitions. A sophisticated segmentation procedure should be able to segment complex, gradual changing elements in images into realistic objects. Within current segmentation and pattern recognition research such procedures have already been developed and tested (Ojala and Pietikäinen, 1999). Intelligent classifiers could be based on membership functions in combination with an extensive hierarchical or multiple-scale knowledge base (Sowmya and Trinder, 2000) or on neural networks (Pax-Lenney et al., 2001).

9. Conclusions

For Landsat TM-based forest stand types mapping in the steep mountainous study area, a combination of a DEM and the TM bands 4, 5 and a synthetic band (TM band 4 − TM band 2), which were all corrected for topography, provided the most optimal dataset when using traditional per-pixel classification. Therefore, this dataset was used to compare traditional per-pixel classification with an object-based classification method. The results from this comparison suggest that an object-based approach fails to improve the accuracy of forest stand type maps in comparison with traditional per-pixel classification, according to the overall accuracies and KHAT values. Object-based classification, however, does improve the applicability of Landsat TM derived forest stand type maps. The forest stand type map derived with object-based classification agreed more with reality than the forest stand type map derived with per-pixel classification. The use of a more specific accuracy assessment method for the object-based classification could probably quantify this. Alternatively, the use of a random sampled ground truth set based on field inventories for accuracy assessment could have resulted in higher object-based classification accuracies as well. Unfortunately a random sampled ground truth set was not available, because in steep mountainous terrain it is very time and labour consuming to obtain such a set.

The relation between the average object size and the accuracy of forest stand type maps indicated that object-based classification has substantial potential for forest stand type mapping on the basis of Landsat TM. Though, techniques for segmentation and classification of spatially continuous elements with gradual transitions in images have to be improved.

Acknowledgements

This paper is a result of work carried out for the CARTESIAN project. The funding by the European Community (ENV4-CT98-0746) and the project coordination by J.C.J.H. Aerts are gratefully acknowledged. We thank Ger Bergkamp and Anton Imeson for suggestions for improvement.

References

Anderson, J.R., Hardy, E.E., Roach, J.T., Witmer, R.E., 1976. A land use and land cover classification system for use with remote sensor data. US Geological Survey, Professional Paper 964, Washington, DC, 28 pp.

Ardö, J., 1992. Volume quantification of coniferous forest compartments using spectral radiance recorded by Landsat Thematic Mapper. Int. J. Remote Sens. 13 (9), 1779–1786.

Baatz, M., Schäpe, A., 2000. Multiresolution segmentation—an optimization approach for high quality multi-scale image segmentation. In: Strobl, J., Blaschke, T., Griesebener, G. (Eds.), Angewandte Geographische Informationsverarbeitung XII. Beiträge zum AGIT-Symposium Salzburg 2000, Herbert Wichmann Verlag, Kahrlsruhe, pp. 12–23.

Bebi, P., Kienast, F., Schönenberger, W., 2001. Assessing structures in mountain forests as a basis for investigating the forests' dynamics and protective function. For. Ecol. Manage. 145 (1), 3–14.

BEV, 2001. Bundesamt für Eich- und Vermessungswesen, Produktinformation zum Geländehöhenmodell. http://www.bev.gv.at/prodinfo/dgm/dgm_3f.htm, as seen on 1 December 2001.

Carpenter, G.A., Gopal, S., Macomber, S., Martens, S., Woodcock, C.E., 1999. A neural network method for mixture estimation for vegetation mapping. Remote Sens. Environ. 70 (2), 138–152.

Carper, W.J., Lillesand, T.M., Kiefer, R.W., 1990. The use of Intensity-Hue-Saturation transformations for merging SPOT panchromatic and multispectral image data. Photogramm. Eng. Remote Sens. 56 (4), 459–467.

Civco, D.L., 1989. Topographic normalization of Landsat Thematic Mapper digital imagery. Photogramm. Eng. Remote Sens. 55 (9), 1303–1309.

Cohen, J., 1960. A coefficient of agreement of nominal scales. Educ. Psychol. Measur. 20 (1), 37–46.

Colby, J.D., 1991. Topographic normalization in rugged terrain. Photogramm. Eng. Remote Sens. 57 (5), 531–537.

Conese, C., Gilabert, M.A., Maselli, F., Bottai, L., 1993. Topographic normalization of TM scenes through the use of an atmospheric correction method and digital terrain models. Photogramm. Eng. Remote Sens. 59 (12), 1745–1753.

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

Congalton, R.G., Plourde, L.C., 2000. Sampling methodology, sample placement, and other important factors in assessing the accuracy of remotely sensed maps. In: Heuvelink, G.B.M., Lemmens, M.J.P.M. (Eds.), Accuracy 2000. Delft University Press, Amsterdam, pp. 117–124.

Congalton, R.G., Green, K., Teply, J., 1993. Mapping old growth forests on national forest and park lands in the Pacific Northwest from remotely sensed data. Photogramm. Eng. Remote Sens. 59 (4), 529–535.

Coops, N., Culvenor, D., 2000. Utilizing local variance of simulated high spatial resolution imagery to predict spatial pattern of forest stands. Remote Sens. Environ. 71 (3), 248–260.

Definiens-Imaging, 2001. eCognition, object oriented image analysis. http://www.definiens-imaging.com/ecognition/index.htm, as seen on 1 December 2001.

Degui, G., Gillespie, A.R., Adams, J.B., Weeks, R., 1999. Statistical approach for topographic correction of satellite images by using spatial context information. IEEE Trans. Geosci. Remote Sens. 37 (1), 236–246.

Ekstrand, S., 1994. Assessment of forest damage with Landsat TM, correction for varying forest stand characteristics. Remote Sens. Environ. 47, 291–302.

Ekstrand, S., 1996. Landsat TM based forest damage assessment Correction for topographic effects. Photogramm. Eng. Remote Sens. 62 (2), 151–161.

Fiorella, M., Ripple, W., 1993. Analysis of conifer forest regeneration using Landsat Thematic Mapper data. Photogramm. Eng. Remote Sens. 59 (9), 1383–1388.

Frank, T., 1988. Mapping dominant vegetation communities in the Colorado Rocky Mountain front range with Landsat Thematic Mapper and digital terrain data. Photogramm. Eng. Remote Sens. 54, 1727–1734.

Franklin, S.E., Giles, P.T., 1995. Radiometric processing of aerial and satellite remote-sensing imagery. Comput. Geosci. 21 (3), 413–423.

GBA, 1998. In: Oberhauser, R., Rataj, W. (Eds.), Geologisch-Tektonische Übersichtskarte von Vorarlberg (1:200.000) mit Erläuterungen. Published by the Geological Survey of Austria.

Gemmell, F.M., 1995. Effects of forest cover, terrain, and scale on timber volume estimation with thematic mapper data in a rocky mountain site. Remote Sens. Environ. 51 (2), 291–305.

Gu, D., Gillespie, A., 1998. Topographic normalization of Landsat TM images of forest based on subpixel sun-canopy-sensor geometry. Remote Sens. Environ. 64 (2), 166–175.

He, H.S., Mladenoff, D.J., Radeloff, V.C., 1998. Integration of GIS data and classified satellite imagery for regional forest assessment and landscape modeling. Ecol. Appl. 8, 1072–1083.

Holben, B.N., Justice, C.O., 1980. The topographic effects on spectral response from nadir pointing sensors. Photogramm. Eng. Remote Sens. 46 (9), 1191–1200.

Hudson, W.D., Ramm, C.W., 1987. Correct formulation of the KHAT value of agreement. Photogramm. Eng. Remote Sens. 53 (4), 421–422.

Hutchinson, C.F., 1982. Techniques for combining Landsat and ancillary data for digital classification improvement. Photogramm. Eng. Remote Sens. 48 (1), 123–130.

Hyyppä, J., Hyyppä, H., Inkinen, M., Engdahl, M., Linko, S., Zhu, Y., 2000. Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. For. Ecol. Manage. 128 (1), 109–120.

Itten, K.I., Meyer, P., 1993. Geometric and radiometric correction of TM-data of mountainous forested areas. IEEE Trans. Geosci. Remote Sens. 31 (4), 764–770.

Itten, K.I., Meyer, P., Kellenberger, T., Leu, R., Sandmeier, S., Bitter, P., Seidel, K., 1992. Correction of the impact of topography and atmosphere on Landsat-TM forest mapping of Alpine regions. Remote Sensing Series 18, University of Zurich, 50 pp.

Jansa, J., 1998. A global topographic normalisation algorithm for satellite images. In: Proceedings of the ISPRS Commission VII Symposium, Budapest, Hungary. Int. Arch. Photogramm. Remote Sens. XXXII (7), 8–15.

Kilpeläinen, P., Tokola, T., 1999. Gain to be achieved from stand delineation in LANDSAT TM image-based estimates of stand volume. For. Ecol. Manage. 124, 105–111.

Leprieur, C.E., Durand, J.M., Peyron, J.L., 1988. Influence of topography on forest reflectance using Landsat thematic mapper and digital terrain data. Photogramm. Eng. Remote Sens. 54 (4), 491–496.

Lobo, A., Chic, O., 1996. Classification of Mediterranean crops with multi sensor data: per-pixel versus per-object statistics and image segmentation. Int. J. Remote Sens. 17, 2385–2400.

Lucas, N.S., Curran, P.J., 1999. Forest ecosystem simulation modelling, the role of remote sensing. Progr. Phys. Geogr. 23 (3), 391–423.

Maier, B., 1993. Forstinventur Stand Montafon. Internal Report Stand Montafon, Schruns, 154 pp.

Martin, M.E., Newman, S.D., Aber, J.D., Congalton, R.G., 1998. Determining forest species using high spectral resolution remote sensing data. Remote Sens. Environ. 65, 249–254.

Meyer, P., Itten, K.I., Kellenberger, T., Sandmeier, S., Sandmeier, R., 1993. Radiometric correction of topographically induced effects on Landsat TM data in an alpine environment. ISPRS J. Photogramm. Remote Sens. 48 (4), 17–28.

Mickelson, J.G.J., Civco, D.L., Silander, J.A.J., 1998. Delineating forest canopy species in the Northeastern United States using multi-temporal TM data and GPS referenced data. Photogramm. Eng. Remote Sens. 64, 891–907.

Minnaert, M., 1941. The reciprocity principle in lunar photometry. Astrophys. J. 93, 403–410.

Ojala, T., Pietikäinen, M., 1999. Unsupervised texture segmentation using feature distributions. Pattern Recogn. 32, 744–786.

Pax-Lenney, M., Woodcock, C.E., Macomber, S.A., Song, C., 2001. Forest mapping with a generalized classifier and Landsat TM data. Remote Sens. Environ. 77, 241–250.

Peterson, D.L., Spanner, M.A., Running, S.W., Teuber, K.B., 1987. Relationship of thematic mapper simulator data to leaf area index of temperate coniferous forest. Remote Sens. Environ. 22, 323–341.

Richards, J.A., Jia, X., 1999. Remote Sensing Digital Image Analysis, An Introduction. Springer, Berlin, 363 pp.

Richter, R., 1997. Correction of atmospheric and topographic effects for high spatial resolution satellite imagery. Int. J. Remote Sens. 18 (5), 1099–1111.

Richter, R., 1998. Correction of satellite imagery over mountainous terrain. Appl. Opt. 37 (18), 4004–4015.

Salvador, R., Pons, X., 1998. On the reliability of Landsat TM for estimating forest variables by regression techniques, a methodological analysis. IEEE Trans. Geosci. Remote Sens. 36 (6), 1888–1897.

Sandmeier, S., Itten, K.I., 1997. A physically-based model to correct atmospheric and illumination effects in optical satellite data of rugged terrain. IEEE Trans. Geosci. Remote Sens. 35 (3), 708–717.

Smith, J.A., Lin, T.L., Ranson, K.J., 1980. The Lambertian assumption and Landsat data. Photogramm. Eng. Remote Sens. 46, 1183–1189.

Sowmya, A., Trinder, J., 2000. Modelling and representation issues in automated feature extraction from aerial and satellite images. ISPRS J. Photogramm. Remote Sens. 55, 34–47.

Strahler, A.H., Logan, T.L., Bryant, N.A., 1978. Improving forest cover classification accuracy from Landsat by incorporating topographic information. In: Proceedings of the 12th International Symposium on Remote Sensing of Environment, ERIM, Ann Harbor, MI, pp. 927–942.

Stuckens, J., Coppin, P.R., Bauer, M.E., 2000. Integrating contextual information with per-pixel classification for improved land cover classification. Remote Sens. Environ. 71, 282–296.

Teillet, P.M., Guidon, B., Goodenough, D.G., 1982. On the slope-aspect correction of multispectral scanner data. Can. J. Remote Sens. 8, 84–106.

Tokola, T., Sarkeala, J., Linden, M.V.D., 2001. Use of topographic correction in Landsat TM-based forest interpretation in Nepal. Int. J. Remote Sens. 22 (4), 551–563.

Ton, J., Sticklen, J., Jain, A.K., 1991. Knowledge-based segmentation of Landsat images. IEEE Trans. Geosci. Remote Sens. 29, 222–232.

Trotter, C.M., Dymond, J.R., Goulding, C.J., 1997. Estimation of timber volume in a coniferous plantation forest using Landsat TM. Int. J. Remote Sens. 18 (10), 2209–2223.

UN-ECE/FAO, 2000. Forest Resources of Europe, CIS, North America, Australia, Japan and New Zealand (industrialized temperate/boreal countries). UN-ECE/FAO Contribution to the Global Forest Resources Assessment 2000. Geneva Timber and Forest Study Papers, No. 17, United Nations, New York and Geneva.

van Noord, H., 1996. The role of geomorphological information in ecological forest site typology in mountainous areas, a methodological study in the E-Rätikon and NW-Montafon mountains (Vorarlberg, Austria). Ph.D. Thesis. University of Amsterdam, Amsterdam, 185 pp.

Woodcock, C., Harward, V.J., 1992. Nested-hierarchical scene models and image segmentation. Int. J. Remote Sens. 13 (16), 3167–3187.